



Universidade Nova de Lisboa
Faculdade de Ciências Médicas

SUSCEPTIBILIDADE GENÉTICA PARA CANCRO DA MAMA

SUSANA MARIA NUNES DA SILVA

**Dissertação apresentada para
obtenção do Grau de Doutor em
Ciências da Vida - ramo Genética,
pela Universidade Nova de Lisboa,
Faculdade de Ciências Médicas**

Esta dissertação teve orientação de:
Orientador: Professor Doutor Jorge Francisco Gaspar
Co-Orientador: Professor Doutor José Rueff

Lisboa
2010

Ao Rui, à Madalena e à pequena Marta
que chegou para nos unir ainda mais.
Todo o meu Amor!

AGRADECIMENTOS

Gostaria de agradecer a todas as pessoas e instituições que de alguma forma contribuíram para a realização do trabalho que possibilitou a elaboração desta dissertação. Existem no entanto aqueles que merecem o meu profundo agradecimento, não podendo deixar de o destacar.

Ao Professor Doutor José Rueff, Director do Departamento de Genética da Faculdade de Ciências Médicas da Universidade Nova de Lisboa, por me ter recebido quando era ainda uma *jovem* a terminar a licenciatura e ter permitido a minha progressão científica, facultando todas as condições necessárias à realização do trabalho aqui apresentado. Ainda pela amizade, disponibilidade, empenho e confiança que depositou em mim ao longo de todos estes anos. Muito Obrigada.

Ao Professor Doutor Jorge Gaspar por me ter dado a possibilidade de desempenhar um trabalho científico de qualidade, por ter sido muito mais do que um mentor, um amigo, nesta minha caminhada. Pela amizade, confiança e disponibilidade que sempre revelou ao longo de todos estes anos de trabalho conjunto. Mesmo nas horas menos boas, o que não nos derruba torna-nos mais fortes. Muito Obrigada.

Ao Professor Doutor Sebastião Rodrigues e ao Professor Doutor Nuno Oliveira por terem tido sempre uma palavra amiga nas mais diversas ocasiões.

Aos colegas e amigos do Departamento de Genética da Faculdade de Ciências Médicas por todo o apoio e companheirismo. Seriam necessárias muitas páginas para deixar a cada um uma palavra especial. Cada um de vós, à sua maneira, deu o seu precioso contributo, com um gesto, um olhar ou uma simples palavra. A todos muito obrigada.

Um agradecimento muito especial à minha família e Amigos, porque sempre acreditaram que seria possível. Pelo apoio incondicional nos melhores e nos piores momentos, especialmente quando achamos que não somos capazes. Muito Obrigada.

SUMÁRIO

O cancro da mama é a patologia oncológica mais frequente nas mulheres sendo o responsável pela maior taxa de mortalidade por cancro no sexo feminino. Contudo, as causas inerentes a esta patologia permanecem por esclarecer. Nos últimos anos tem-se verificado que o risco para patologia neoplásica depende de factores ambientais e genéticos, estando estes últimos associados à variabilidade genética inter-individual. Polimorfismos genéticos em genes envolvidos no metabolismo de hormonas sexuais, de cancerígenos ambientais e na reparação da lesão genética, são potenciais candidatos a estarem associados à susceptibilidade individual para esta patologia. Assim, neste trabalho desenvolveram-se estudos de associação caso-controlo na população Portuguesa, com vista a avaliar-se o papel atribuído aos polimorfismos na susceptibilidade para cancro da mama.

Foram seleccionados polimorfismos em genes envolvidos em diferentes vias mecanicistas: destoxificação de cancerígenos, metabolismo de estrogénios, reparação por excisão de bases, reparação por excisão de nucleótidos, reparação *mismatch* e reparação por recombinação homóloga. Os resultados obtidos revelaram associação entre os seguintes polimorfismos e a susceptibilidade individual para cancro da mama: os dois SNPs estudados no gene *XRCC1* (Arg194Trp e Arg399Gln) e o SNP no gene *XRCC3* (Thr241Met) após estratificação pelo *status* menopausico. Mediante estratificação por *status* de amamentação os SNPs identificados nos genes *MnSOD* (Val16Ala) e *XRCC2* (Arg118His); um SNP no gene *MLH3* (Leu844Pro), e por fim como resultado de interacção gene-gene as interacções descritas por *MSH3* Ala1045Thr/*MSH6* Gly39Glu e *MSH4* Ala977Thr/*MLH3* Leu844Pro.

Os resultados obtidos e apresentados na presente dissertação, revelam que o estudo de polimorfismos pode representar um papel determinante na etiologia do cancro da mama. No entanto, mais estudos envolvendo estes mesmos polimorfismos em populações casuisticamente superiores serão uma mais-valia nos estudos de associação para esta neoplasia. Adicionalmente, a utilização da metodologia de *Pools* de DNA, poderá ser uma ferramenta útil na pré-selecção dos polimorfismos mais relevantes a estudar, na medida em que permite estimar a frequência alélica de cada SNP numa determinada população.

ABSTRACT

Breast cancer is the most common form of cancer among women, being the responsible for the highest mortality rate from cancer among the female sex. However, the main causes related to this pathology remain unclear. The risk of neoplastic disease has been connected with genetic and environmental factors. In fact, genes and the environment share the stage for most, if not all, common non-familial cancers, and are related to individual susceptibility. Genetic polymorphisms identified in genes encoding enzymes involved in estrogen metabolism, xenobiotics and DNA repair pathways are believed to be candidates for associations with breast cancer. Therefore, it was our intention to develop case-control studies among the Portuguese population, in order to evaluate the potential role of several genetic polymorphisms in breast cancer susceptibility.

We selected polymorphisms in genes involved in different pathways: carcinogenic detoxification, estrogen metabolism, base excision repair, nucleotide excision repair, mismatch repair and double strand break repair by homologous recombination. The results obtained revealed potential associations between some polymorphisms studied and individual susceptibility to breast cancer. Regarding this fact, our results suggest the potential involvement of two *XRCC1* gene polymorphisms (Arg194Trp and Arg399Gln) and *XRCC3* gene polymorphism (Thr241Met) after stratification to menopausal *status* and after stratification to breastfeeding *status* an association of *MnSOD* gene polymorphism (Val16Ala) and *XRCC2* (Arg188His) with the disease. The SNP identified in *MLH3* gene (Leu844Pro), and the interaction gene-gene described by *MSH3* Ala1045Thr/*MSH6* Gly39Glu and *MSH4* Ala97Thr/*MLH3* Leu844Pro were also related to breast cancer susceptibility.

The results shown in the present dissertation have revealed the potential role of polymorphisms in breast cancer etiology. However, further studies will be needed with larger populations to confirm these results. Additionally, the use of DNA *pools* methodology, as a pre-selection tool, could allow the identification of the most relevant polymorphisms to be studied, estimating the allelic frequency of each SNPs in different populations.

PUBLICAÇÕES

Esta dissertação contém dados e/ou métodos descritos nas seguintes publicações internacionais:

- (1) Teixeira JP, Gaspar J, Coelho P, Costa C, Pinho-Silva S, Costa S, Da Silva S, Laffon B, Pásaro E, Rueff J, Farmer P. “Cytogenetic and DNA damage on workers exposed to styrene”. (2010) *Mutagenesis*. **IN PRESS**
- (2) Gomes BC, Silva SN, Azevedo AP, Manita I, Monteiro Gil O, Ferreira TC, Limbert E, Rueff J, Gaspar JF. “The role of common variants of non-homologous end-joining repair genes *XRCC4*, *LIG4* and *Ku80* in thyroid cancer risk”. (2010) *Oncology Reports*. **IN PRESS**
- (3) **Silva SN**, Tomar M, Paulo C, Gomes BC, Azevedo AP, Teixeira V, Esperança-Pina J, Rueff J, Gaspar JF. “Breast cancer risk and common single nucleotide polymorphisms in homologous recombination DNA repair pathway genes *XRCC2*, *XRCC3*, *NBS1* and *RAD51*.” (2010) *Cancer Epidemiology* **34**: 85-92.
- (4) Conde J, **Silva SN**, Azevedo AP, Teixeira V, Esperança-Pina J, Rueff J, Gaspar JF. “Association of common variants in mismatch repair genes and breast cancer susceptibility: a multigene study.” (2009) *BMC Cancer* **9**: 344.
- (5) Bastos HN, Antão MR, **Silva SN**, Azevedo AP, Manita I, Teixeira V, Esperança-Pina J, Monteiro Gil O, Ferreira TC, Limbert E, Rueff J, Gaspar JF. “Association of polymorphisms in genes of the homologous recombination DNA repair pathway and thyroid cancer risk.” (2009) *Thyroid* **19(10)**: 1067-1075.

- (6) **Silva SN**, Azevedo AP, Teixeira V, Esperança-Pina J, Rueff J and Gaspar JF. “The role of *GSTA2* polymorphisms and haplotypes in breast cancer susceptibility: a case-control study in the Portuguese Population”. (2009) *Oncology Reports* **22** (3): 593-598.
- (7) **Silva SN**, Moita R, Azevedo AP, Gouveia R, Manita I, Esperança-Pina J, Rueff J and Gaspar J. “Menopausal age and *XRCC1* gene polymorphisms: role in breast cancer risk”. (2007) *Cancer Detection and Prevention* **31** (4): 303-309.
- (8) **Silva SN**, Bezerra de Castro G, Faber A, Pires M, Oliveira VC, Azevedo AP, Cabral MN, Manita I, Esperança-Pina J, Rueff J and Gaspar J. “The role of *ERCC2* polymorphisms in breast cancer risk”. (2006) *Cancer Genetics and Cytogenetics* **170** (1): 86-88.
- (9) **Silva SN**, Cabral MN, Azevedo AP, Manita I, Esperança-Pina J, Rueff J and Gaspar J. “Breast cancer risk and polymorphisms in genes involved in metabolism of estrogens (*CYP17*, *HSD17β1*, *COMT* and *MnSOD*): Possible protective role of *MnSOD* gene polymorphism Val/Ala and Ala/Ala in women that never breast fed”. (2006) *Oncology Reports* **16** (4): 781-788.
- (10) **Silva SN**, Monteiro Gil O, Oliveira VC, Cabral MN, Azevedo AP, Faber A, Manita I, Ferreira TC, Limbert E, Esperança-Pina J, Rueff J and Gaspar J. “Association of *ERCC2* polymorphisms with non-familial thyroid cancer risk.” (2005) *Cancer Epidemiology Biomarkers and Prevention* **14** (10): 2407-2412.
- (11) **Silva SN**, Cabral MN, Bezerra G, Pires M, Azevedo AP, Manita I, Esperança-Pina J, Rueff J and Gaspar J. “The *MnSOD* polymorphisms and breast cancer risk”. (2005) *Cancer Epidemiology Biomarkers & Prevention* **14** (11 – Part 2): 2731s.
(Abstrat)

No âmbito desta dissertação foram igualmente apresentadas várias comunicações em congressos e reuniões científicas nacionais e internacionais, das quais se destacam:

- (1) Conde J, **Silva SN**, Azevedo AP, Manita I, Teixeira V, Esperança-Pina J, Rueff J and Gaspar JF. **Mismatch repair gene polymorphisms and breast cancer susceptibility.** XXXIV Jornadas Portuguesas de Genética, Lisboa, Portugal. (Abril 2009); Abstract aceite para comunicação oral.
- (2) **Silva SN**, Azevedo AP, Manita I, Esperança-Pina, Rueff J and Gaspar J. **GSTA2 Polymorphisms and breast cancer susceptibility.** 99th Annual Meeting of the American Association for Cancer Research – San Diego, CA, (Abril 2008).
- (3) Tomar M, **Silva SN**, Azevedo AP, Manita I, Esperança-Pina, Rueff J and Gaspar J. **Polymorphisms in genes of the homologous recombination DNA repair pathway, breast feeding and breast cancer risk.** 98th Annual Meeting of the American Association for Cancer Research – Los Angeles, CA (Abril 2007).
- (4) **Silva SN**, Moita R, Azevedo AP, Manita I, Esperança-Pina, Rueff J and Gaspar J. **Menopausal age and XRCC1 polymorphisms: role on breast cancer.** 97th Annual Meeting of the American Association for Cancer Research – Washington, DC (Abril 2006).
- (5) **Silva SN**, Cabral MN, Bezerra G, Pires M, Azevedo AP, Manita I, Esperança-Pina, Rueff J and Gaspar J. **The MnsOD polymorphisms and breast cancer risk.** Fourth Annual AACR International Conference - Frontiers in Cancer Prevention Research – Baltimore, Maryland (Novembro 2005).
- (6) **Silva SN**, Oliveira VC, Cabral MC, Azevedo AP, Faber A, Manita I, Esperança-Pina J, Rueff J and Gaspar J. **The role of ERCC2 polymorphisms in breast cancer risk.** 96th Annual Meeting of the American Association for Cancer Research – Anaheim, CA (Abril 2005).

- (7) Gomes M, Guerreiro D, **Silva SN**, Faber A, Azevedo AP, Manita I, Esperança-Pina J, Rueff J and Gaspar J. **DNA pooled samples strategies for identification of genetic polymorphisms associated with breast cancer risk.** 95th Annual Meeting of the American Association for Cancer Research – Orlando, Florida (Março 2004).
- (8) Paulo C, **Silva SN**, Faber A, Azevedo AP, Manita I, Esperança-Pina J, Rueff J and Gaspar J. **XRCC3 polymorphisms and breast cancer risk.** XXXI Jornadas de Genética. Oeiras, Portugal (Fevereiro 2004).
- (9) Gaspar J, **Silva S**, Faber A, Guerreiro D, Azevedo AP, Manita I, Ferreira G, Esperança-Pina J and Rueff J. **Polymorphisms in ERCC2 gene are not associated with breast cancer.** 94th Annual Meeting of the American Association for Cancer Research – Washington, USA (Julho 2003).

SIMBOLOGIA E ABREVIATURAS

AP	Local abásico
B(a)P	Benzo[a]pireno
BER	Reparação por Excisão de Bases
bp	Pares de bases
CYP	Citocromo P450
DCIS	Carcinoma ductal <i>in situ</i>
DNA	Ácido Desoxirribonucleico
DP	Desvio Padrão
DSB	Quebras de cadeia dupla
ER	Receptor de estrogénios
Fago γ	Fago <i>lambda</i>
GSH	Glutathione reduzido
GST	Glutathione-S-Transferase
GWAS	do inglês <i>Genome Wide Association Study</i>
HCA	Aminas aromáticas heterocíclicas
HRR	Reparação por recombinação homóloga
HWE	Equilíbrio de Hardy-Weinberg
IC	Intervalo de Confiança
IDC	Carcinoma ductal invasivo
IDL	<i>Loops</i> de Inserção/Delecção
ILC	Carcinoma lobular invasivo
IR	Radiação Ionizante
LCIS	Carcinoma Lobular <i>in situ</i>
LD	Desequilíbrio de <i>Linkage</i> do inglês <i>Linkage Disequilibrium</i>
LOH	Perda de heterozigotia
MAF	Frequência do alelo menos comum, do inglês <i>Minor Allele Frequency</i>
MMR	Reparação por excisão de erros na incorporação e emparelhamento de bases, do inglês <i>Mismatch Repair</i>
MN	Micronúcleos
MS	Microsatélites
MSI	Instabilidade de Microsatélites
NAT	N-acetiltransferases
NER	Reparação por Excisão de Nucleótidos
NHEJ	Reparação não-homóloga
OR	Factor de risco
PAH	Hidrocarbonetos aromáticos policíclicos
PCR	Reacção de amplificação em cadeia, do inglês <i>Polymerase Chain Reaction</i>
PR	Receptor de progesterona

RFLP	Análise de fragmentos de restrição, do inglês <i>Restriction Fragment Length Polymorphism</i>
ROS	Espécies reactivas de oxigénio
SCE	Troca entre cromátides irmãs, do inglês <i>Sister Chromatid Exchange</i>
SNP	Polimorfismo de base única, do inglês <i>Single Nucleotide Polymorphism</i>
SOD	Superóxido dismutase
SSB	Quebras de cadeia simples
TNM	Do inglês <i>Tumor-Node-Metastasis</i>
UTR	Região não traduzida, do inglês <i>Untranslated Region</i>
UV	Ultra-Violeta

Nota: algumas das abreviaturas que constam desta lista e presentes nesta dissertação resultam de estrangeirismos. A inclusão destas abreviaturas, ou das próprias expressões que as originaram, na sua língua original, resultou do facto de serem expressões técnicas amplamente utilizadas e reconhecidas pela comunidade científica internacional.

SÍMBOLOS DOS AMINOÁCIDOS

A	ou	Ala	Alanina	M	ou	Met	Metionina
C	ou	Cis	Cisteína	N	ou	Asn	Asparagina
D	ou	Asp	Ácido Aspártico	P	ou	Pro	Prolina
E	ou	Glu	Ácido Glutâmico	Q	ou	Gln	Glutamina
F	ou	Fen	Fenilalanina	R	ou	Arg	Arginina
G	ou	Gli	Glicina	S	ou	Ser	Serina
H	ou	His	Histidina	T	ou	Tre	Treonina
I	ou	Ile	Isoleucina	V	ou	Val	Valina
K	ou	Lis	Lisina	W	ou	Trp	Triptofano
L	ou	Leu	Leucina	Y	ou	Tir	Tirosina

ÍNDICE GERAL

Agradecimentos	v
Sumário	vii
Abstract	viii
Publicações.....	ix
Simbologia e Abreviaturas.....	xiii
Capítulo 1.....	1
1. Enquadramento Teórico.....	3
1.1 Genética do Cancro	3
1.2 Cancro de Mama	4
1.2.1 Histologia.....	4
1.3 Epidemiologia	6
1.4 Factores de Risco.....	8
1.4.1 Factores Genéticos	8
1.4.2 Factores Não-Genéticos	10
1.5 Susceptibilidade Individual.....	18
1.5.1 Polimorfismos Genéticos e Variabilidade Inter-individual.....	18
1.5.2 Estudos de Associação.....	20
1.6 Metabolização de cancerígenos.....	22
1.7 Reparação da Lesão Genética	25
1.7.1 Reparação Directa (DR)	26
1.7.2 Reparação de <i>Mismatch</i> (MMR).....	27
1.7.3 Reparação por Excisão de Bases (BER).....	30
1.7.4 Reparação por Excisão de Nucleótidos (NER)	33
1.7.5 Reparação de Quebras em Cadeia Dupla (DSB)	36
Capítulo 2.....	39
2. Objectivos.....	41
Capítulo 3.....	45
3. O papel de polimorfismos do gene <i>GSTA2</i> e haplótipos na susceptibilidade individual para cancro da mama: estudo caso-controlo na população portuguesa.....	47
3.1 Introdução	47
3.2 Materiais e Métodos	49

3.2.1	Descrição da População	49
3.2.2	Extracção de DNA.....	49
3.2.3	Genotipagem de polimorfismos do gene <i>GSTA2</i> : P110S, S112T e E210A	50
3.2.4	Análise Estatística.....	51
3.3	Resultados.....	52
3.4	Discussão.....	56
Capítulo 4	61
4.	Polimorfismos envolvidos no metabolismo de estrogénios (<i>CYP17</i> , <i>HSD17β1</i> , <i>COMT</i> e <i>MnSOD</i>) e o risco para cancro da mama.....	63
4.1	Introdução.....	63
4.2	Materiais e Métodos.....	65
4.2.1	Descrição da População	65
4.2.2	Extracção de DNA.....	66
4.2.3	Genotipagem dos polimorfismos <i>CYP17</i> (5'-UTR, T27C), <i>HSD17β1</i> (Gly313Ser), <i>COMT</i> (Val158Met) e <i>MnSOD</i> (Val16Ala).....	66
4.2.4	Análise Estatística.....	68
4.3	Resultados.....	68
4.4	Discussão.....	73
Capítulo 5	79
5.	A idade da Menopausa e Polimorfismos no gene <i>XRCC1</i> : o seu papel no cancro da mama. 81	
5.1	Introdução.....	81
5.2	Materiais e Métodos.....	83
5.2.1	Descrição da População	83
5.2.2	Extracção de DNA.....	83
5.2.3	Genotipagem dos polimorfismos <i>OGG1</i> (Ser326Cys), <i>XRCC1</i> (Arg194Trp) e (Arg399Gln).....	83
5.2.4	Análise Estatística.....	84
5.3	Resultados.....	84
5.4	Discussão.....	93
Capítulo 6	97
6.	O papel de polimorfismos do gene <i>ERCC2</i> no risco para cancro da mama	99
6.1	Introdução.....	99
6.2	Materiais e Métodos.....	101
6.2.1	Descrição da população	101
6.2.2	Extracção de DNA.....	101

6.2.3	Genotipagem dos polimorfismos presentes no gene <i>ERCC2</i> (Ile199Met), (His201Tyr), (Asp312Asn) e (Lys751Gln)	101
6.2.4	Análise Estatística	102
6.3	Resultados	104
6.4	Discussão	108
Capítulo 7.....		113
7.	Associação de variantes em genes envolvidos da via de reparação <i>mismatch</i> e susceptibilidade para cancro da mama: estudo multigénico.....	115
7.1	Introdução	115
7.2	Materiais e Métodos	116
7.2.1	Descrição da população.....	116
7.2.2	Extracção de DNA	116
7.2.3	Genotipagem dos polimorfismos	117
7.2.4	Análise Estatística	117
7.3	Resultados	118
Capítulo 8.....		133
8.	Risco para cancro da mama e SNPs em genes da via de reparação por recombinação homóloga <i>XRCC2</i> , <i>XRCC3</i> , <i>NBS1</i> e <i>RAD51</i>	135
8.1	Introdução	135
8.2	Materiais e Métodos	138
8.2.1	Seleccção da População	138
8.2.2	Extracção de DNA	138
8.2.3	Seleccção dos tagSNPs	138
8.2.4	Genotipagem dos polimorfismos	140
8.2.5	Análise Estatística	141
8.3	Resultados	141
8.4	Discussão	147
Capítulo 9.....		153
9.	Aplicabilidade da técnica de <i>POOLS</i> de DNA	155
9.1	Introdução	155
9.2	Materiais e Métodos	157
9.2.1	Quantificação das amostras de DNA e construção dos diferentes <i>pools</i>	157
9.2.2	Genotipagem individual	160
9.2.3	Frequências alélicas determinadas após análise dos <i>pools</i>	160
9.3	Resultados	161

9.3.1	Reproductibilidade da técnica	161
9.4	Discussão.....	163
Capítulo 10	169
10.	Considerações Finais.....	171

ÍNDICE DE TABELAS

TABELA 1.1 – LESÕES DE DNA E RESPECTIVAS VIAS DE REPARAÇÃO	26
TABELA 1.2 – COMPONENTES DE MMR E AS SUAS FUNÇÕES (JIRICNY, 2006)	28
TABELA 3.1 - ALGUNS DETALHES DA TÉCNICA PCR-RFLP APLICADOS AO POLIMORFISMO <i>GSTA2</i> E210A	50
TABELA 3.2 – CARACTERÍSTICAS GERAIS DO GRUPO DE DOENTES (N=291) E DO GRUPO DE CONTROLOS (N=547) ESTUDADOS.53	
TABELA 3.3 – DISTRIBUIÇÃO DE GENÓTIPOS E RISCO PARA CANCRO DA MAMA PARA TODOS OS POLIMORFISMOS ESTUDADOS (<i>GSTA2</i> P110S, S112T E E210A) EM AMBAS AS POPULAÇÕES, DOENTES E CONTROLOS.....	55
TABELA 3.4 – COEFICIENTE DE DESEQUILÍBIO DE <i>LINKAGE D'</i> ENTRE OS SNPs ESTUDADOS.	56
TABELA 3.5 – FREQUÊNCIAS DE HAPLÓTIPOS PARA O GENE <i>GSTA2</i> NA POPULAÇÃO PORTUGUESA OBTIDAS ATRAVÉS DO SOFTWARE SNPSTATS.	56
TABELA 4.1 – CONDIÇÕES DE PCR-RFLP PARA OS POLIMORFISMOS ESTUDADOS.....	67
TABELA 4.2 - DETERMINAÇÃO DAS FREQUÊNCIAS GENOTÍPICAS E ALÉLICAS PARA A POPULAÇÃO CONSTITUÍDA PELOS ÍNDIOS XAVANTE (N = 179), PARA OS QUATRO POLIMORFISMOS EM ESTUDO.	68
TABELA 4.3 – FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS DOS POLIMORFISMOS ESTUDADOS PUBLICADAS EM ESTUDOS DESENVOLVIDOS EM DIFERENTES POPULAÇÕES.	70
TABELA 4.4 – CARACTERÍSTICAS GERAIS DAS POPULAÇÕES DE DOENTES E CONTROLOS ESTUDADAS PARA O POLIMORFISMO NO GENE <i>MNSOD</i>	71
TABELA 4.5 – VALORES DE OR (IC 95%) PARA CANCRO DA MAMA DETERMINADO PARA OS GENÓTIPOS GERADOS PELO POLIMORFISMO PRESENTE NO GENE <i>MNSOD</i> (VAL16ALA).	73
TABELA 5.1 - CONDIÇÕES DE PCR-RFLP PARA OS POLIMORFISMOS ESTUDADOS.	86
TABELA 5.2 - DETERMINAÇÃO DAS FREQUÊNCIAS GENOTÍPICAS E ALÉLICAS PARA A POPULAÇÃO CONSTITUÍDA PELOS ÍNDIOS XAVANTE (N = 179), PARA OS POLIMORFISMOS EM ESTUDO.	87
TABELA 5.3 – FREQUÊNCIAS GENOTÍPICAS E ALÉLICAS DOS POLIMORFISMOS ESTUDADOS PUBLICADAS EM ESTUDOS DESENVOLVIDOS EM DIFERENTES POPULAÇÕES.	88
TABELA 5.4 - CARACTERÍSTICAS GERAIS DO GRUPO DE DOENTES (N=241) E DO GRUPO DE CONTROLOS (N=457) ESTUDADOS. 89	
TABELA 5.5 – VALORES DE OR (IC 95%) PARA O POLIMORFISMO <i>XRCC1</i> ARG194TRP E RESPECTIVA ASSOCIAÇÃO COM CANCRO DA MAMA.....	91
TABELA 5.6 - VALORES DE OR (IC 95%) PARA O POLIMORFISMO <i>XRCC1</i> ARG399GLN E RESPECTIVA ASSOCIAÇÃO COM CANCRO DA MAMA.....	92
TABELA 5.7 – COMBINAÇÃO DE GENÓTIPOS DOS POLIMORFISMOS DO GENE <i>XRCC1</i> ESTUDADOS (ARG194TRP E ARG399GLN) NA POPULAÇÃO PORTUGUESA.	93
TABELA 6.1 - CONDIÇÕES DE PCR-RFLP PARA OS POLIMORFISMOS DO GENE <i>ERCC2</i> ESTUDADOS.	103
TABELA 6.2 - CARACTERÍSTICAS GERAIS DO GRUPO DE DOENTES (N=241) E DO GRUPO DE CONTROLOS (N=452) ESTUDADOS.	105
TABELA 6.3 - VALORES DE OR (IC 95%) PARA OS POLIMORFISMOS <i>ERCC2</i> ASP312ASN E LYS751GLN E RESPECTIVA ASSOCIAÇÃO COM CANCRO DA MAMA.	107
TABELA 6.4 – COMBINAÇÃO DOS GENÓTIPOS OBTIDOS PARA OS POLIMORFISMOS ASP312ASN E LYS751GLN DO GENE <i>ERCC2</i> NAS POPULAÇÕES ESTUDADAS.	107
TABELA 6.5 - VALORES DE OR (IC 95%) OBTIDOS POR REGRESSÃO LOGÍSTICA PARA AS COMBINAÇÕES DE GENÓTIPOS DOS POLIMORFISMOS <i>ERCC2</i> ASP312ASN E LYS751GLN E ASSOCIAÇÃO COM CANCRO DA MAMA.	108
TABELA 7.1 – CARACTERÍSTICAS GERAIS DAS POPULAÇÕES DE DOENTES (N=287) E CONTROLOS (N=547) ENVOLVIDAS NO ESTUDO.....	118
TABELA 7.2 – FREQUÊNCIAS GENOTÍPICAS E FREQUÊNCIAS DO MENOR ALELO (MAF) NA POPULAÇÃO CONTROLO (N=547) E NA POPULAÇÃO DE DOENTES (N=287) E ANÁLISE DE ASSOCIAÇÃO DOS SNPs ESTUDADOS INDIVIDUALMENTE COM RISCO PARA CANCRO DA MAMA.	121
TABELA 7.3 – EFEITO DA COMBINAÇÃO DE GENÓTIPOS ENTRE DOIS POLIMORFISMOS DE UM MESMO GENE NA SUSCEPTIBILIDADE PARA CANCRO DA MAMA.	123
TABELA 7.4 – EFEITOS DAS INTERACÇÕES SNP-SNP NO CANCRO DA MAMA.	124
TABELA 7.5 – INTERACÇÕES ENTRE SNPs DE DIFERENTES GENES COM PLAUSIBILIDADE BIOLÓGICA E RISCO PARA CANCRO DA MAMA.	129
TABELA 8.1 – RELAÇÃO DOS TAGSNPs SELECIONADOS PARA OS GENES EM ESTUDO (FONTE: GENOME VARIATION SERVER).139	
TABELA 8.2 - CONDIÇÕES DE PCR-RFLP PARA OS POLIMORFISMOS <i>XRCC2</i> R188H E <i>XRCC3</i> T241M.	141

TABELA 8.3 – CARACTERÍSTICAS GERAIS PARA AS POPULAÇÕES DE DOENTES COM CANCRO DA MAMA (N=289) E CONTROLO (N=548).	142
TABELA 8.4 – DISTRIBUIÇÃO DE GENÓTIPOS E RISCO PARA CANCRO DA MAMA ASSOCIADO AOS POLIMORFISMOS <i>XRCC2</i> R188H, <i>NBS1</i> E185Q, <i>XRCC3</i> T241M E <i>RAD51</i> 5'UTR NAS POPULAÇÕES DE DOENTES (N = 289) E CONTROLOS (N = 548).	143
TABELA 8.5 – ASSOCIAÇÃO DOS DIFERENTES POLIMORFISMOS <i>XRCC2</i> R188H, <i>NBS1</i> E185Q, <i>RAD51</i> 5'UTR E <i>XRCC3</i> T241M E RISCO PARA CANCRO DA MAMA NA POPULAÇÃO DE DOENTES (N=289), DE ACORDO COM O <i>STATUS</i> DE AMAMENTAÇÃO.....	144
TABELA 8.6 - ASSOCIAÇÃO DOS DIFERENTES POLIMORFISMOS <i>XRCC2</i> R188H, <i>NBS1</i> E185Q, <i>RAD51</i> 5'UTR E <i>XRCC3</i> T241M E RISCO PARA CANCRO DA MAMA NA POPULAÇÃO DE DOENTES (N=289), DE ACORDO COM O <i>STATUS</i> MENOPAUSICO.	146
TABELA 9.1 – DESCRIÇÃO DA COMPOSIÇÃO DE TODOS OS <i>POOLS</i> CONSTRUÍDOS.....	158
TABELA 9.2 - RELAÇÃO ENTRE A ALTURA RELATIVA DOS PICOS E A PERCENTAGEM DAS FREQUÊNCIAS ALÉLICAS PARA O ALELO MAIS COMUM, PARA CADA UM DOS <i>SNPs</i> , OBTIDAS PARA OS DIFERENTES <i>POOLS</i> (Δ CORRESPONDE À DIFERENÇA ENTRE O VALOR DA FREQUÊNCIA ALÉLICA OBTIDA POR GENOTIPAGEM INDIVIDUAL E O VALOR OBTIDO ATRAVÉS DOS <i>POOLS</i>).	164

ÍNDICE DE FIGURAS

FIGURA 1.1 - SUSCEPTIBILIDADE GENÉTICA PARA O CANCRO DA MAMA (ADAPTADO DE OLOPADE <i>ET AL.</i> , 2008).....	10
FIGURA 1.2 - ALGUNS FACTORES GENÉTICOS E NÃO-GENÉTICOS ASSOCIADOS AO RISCO PARA CANCRO DA MAMA (ADAPTADO DE LEVY-LAHAD E PLON, 2003).....	11
FIGURA 1.3 – APRESENTAÇÃO ESQUEMÁTICA DA BIOSÍNTESE E INACTIVAÇÃO METABÓLICA DE ESTROGÉNIOS (ADAPTADO DE MITRUNEN E HIRVONEN, 2003).....	17
FIGURA 1.4 – DISTRIBUIÇÃO PERCENTUAL DAS DIFERENTES VARIAÇÕES POLIMÓRFICAS NO GENOMA HUMANO, DE ACORDO COM A BASE DE DADOS HGMD (DADOS DE 2009).	19
FIGURA 1.5 - DIAGRAMA ESQUEMÁTICO DO MECANISMO ATRAVÉS DO QUAL O BENZO[A]PIRENO PROMOVE O DESENVOLVIMENTO DO CANCRO (ADAPTADO DE RUNDLE, 2006).	24
FIGURA 1.6 - MODELO PARA A VIA DE REPARAÇÃO DE <i>MISMATCH</i> (MMR) (ADAPTADO DE JASCUR E BOLAND, 2006).	28
FIGURA 1.7 - VIA DE REPARAÇÃO POR EXCIÇÃO DE BASES (BER) (ADAPTADO DE HOEIJMAKERS, 2001).	32
FIGURA 1.8 - MECANISMO DE REPARAÇÃO POR EXCIÇÃO DE NUCLEÓTIDOS (NER). SÃO REPRESENTADAS AS DUAS SUB-VIAS QUE CARACTERIZAM A VIA NER (GLOBAL-NER E TC-NER) (ADAPTADO DE CLEAVER <i>ET AL.</i> , 2009).	35
FIGURA 1.9 - VIAS DE REPARAÇÃO ASSOCIADAS COM REPARAÇÃO DE LESÕES DE QUEBRAS EM CADEIA DUPLA (ADAPTADO DE BRISTOW E HILL, 2008).....	37
FIGURA 9.1 – ESTRATÉGIA UTILIZADA PARA ANÁLISE E PREPARAÇÃO DE <i>POOLS</i> DE DNA. APÓS QUANTIFICAÇÃO, O <i>POOL</i> É PREPARADO COM IGUAL QUANTIDADE DE CADA AMOSTRA QUE O CONSTITUI, SENDO POSTERIORMENTE UTILIZADO COMO ÚNICA AMOSTRA NAS REACÇÕES DE PCR/RFLP. OS PRODUTOS DE REACÇÃO SÃO SEPARADOS EM GEL DE POLIACRILAMIDA E QUANTIFICADOS.	159

CAPÍTULO 1

ENQUADRAMENTO TEÓRICO

1. ENQUADRAMENTO TEÓRICO

1.1 GENÉTICA DO CANCRO

O cancro é essencialmente uma doença genética, e foi neste pressuposto que durante anos a investigação nesta área se baseou, procurando eventos mutacionais que o justificassem. Com o contínuo desenvolvimento dos estudos na área da cancerigénese aumentaram as evidências associadas ao envolvimento de diversos factores na progressão do cancro dos quais se salientam a importância dos mecanismos genéticos e epigenéticos e das interacções celulares (Ponder, 2001). O estudo destes fenómenos conduziu à caracterização fenotípica de uma célula cancerígena: capacidade de ignorar sinais indicadores de paragem da proliferação celular e de sinais que indiquem diferenciação celular, potencial replicativo ilimitado, evasão de apoptose, invasão de tecidos (metastização) e angiogénese (Hanahan e Weinberg, 2000; Peters *et al.*, 2001; Ponder, 2001).

A doença oncológica é uma doença complexa na qual ocorre perturbação da regulação de crescimento e maturação normal das células (Peters *et al.*, 2001). Existem diferentes tipos de cancro, e mesmo alguns do mesmo tipo podem apresentar comportamentos distintos (Peters *et al.*, 2001) normalmente associados à combinação de diferentes eventos genéticos responsáveis por originarem diferentes características (Ponder, 2001). Apesar desta diversidade, existem alterações fundamentais comuns a todos os tipos de cancro, nomeadamente a instabilidade genómica (Peters *et al.*, 2001). Existe uma grande variedade de formas de provocar lesões e desencadear esta instabilidade. Destacam-se, processos que perturbem mecanismos de replicação e reparação, que ocorrem naturalmente com a idade; bem como exposição ocupacional, estilo de vida, exposição a cancerígenos ambientais; alterações hereditárias ou mesmo o acaso (Peters *et al.*, 2001).

Após três décadas de rápidos avanços da investigação na área do cancro, e desta investigação ter gerado um vasto e complexo conhecimento, tem vindo a ser demonstrado que o cancro é uma doença que envolve alterações no genoma. Contudo, muitos autores defendem que a origem e o tratamento desta doença

continuarão a ser alvo de estudo durante os próximos anos, acrescentando sucessivos níveis de complexidade a uma investigação que já é, por si só, complexa (Hanahan e Weinberg, 2000).

Da compreensão das causas subjacentes à aquisição da lesão genética responsáveis pelo fenómeno da cancerigénese, é possível prosseguir no sentido de redução da sua incidência (Bertram, 2000).

1.2 CANCRO DE MAMA

1.2.1 HISTOLOGIA

O cancro da mama é uma doença altamente heterogénea, caracterizada por ser, não uma única doença, mas sim um conjunto de doenças mamárias com diferentes histopatologias associadas, variabilidade genética e, conseqüentemente, diversos prognósticos clínicos de acordo com o tipo de tumor (ductal, lobular, invasivo, etc.) (Vargo-Gogola e Rosen, 2007).

A organização anatómico-funcional da glândula mamária é fundamental para uma eficaz classificação quanto ao tipo de tumor. Assim sendo, é um órgão que apresenta um sistema tubulo-glandular, assente em tecido conectivo estromal e tecido adiposo mamário que se estende do mamilo até aos tubulos. As células epiteliais mamárias estão organizadas em estruturas tubulares tridimensionais fortemente dependentes da morfologia polarizada, da interacção especializada célula-célula e de anexos específicos subjacentes à membrana basal, sendo a sua organização celular igualmente influenciada por forças mecânicas e de sinalização de células vizinhas. As estruturas tubulares são alinhadas por duas camadas celulares: a camada epitelial localizada no lúmen (células luminais) e a camada mioepitelial, com propriedades contrácteis, que se encontra em contacto com a membrana basal, que separa o tecido periglandular estromal (Hergueta-Redondo et al., 2008).

A grande maioria dos tumores malignos da mama são carcinomas, ou seja tumores com origem em células epiteliais, têm origem nos ductos ou nos lóbulos da mama, sendo mais frequente o carcinoma ductal (cerca de 75%) do que o carcinoma lobular (7%). Menos de 1% dos tumores malignos são sarcomas que se desenvolvem a

partir dos tecidos conectivo, ósseo, muscular ou adiposo. Existem ainda outros tipos raros de tumor da mama conhecidos, de que são exemplo o carcinoma medular, o carcinoma mucinoso, o carcinoma tubular e o tumor filóide, entre outros. Quando o tumor se encontra limitado aos ductos ou aos lóbulos, isto é, sem invasão dos tecidos mamários adjacentes ou de outros órgãos, é denominado carcinoma ductal *in situ* (DCIS) ou carcinoma lobular *in situ* (LCIS). Quando ocorre invasão do tecido mamário adjacente é utilizada a designação de carcinoma ductal invasivo (IDC) ou carcinoma lobular invasivo (ILC) (Wajapeyee e Somasundaram, 2004). Estima-se que aproximadamente 70% dos casos de cancro de mama sejam IDC e 10% ILC (Hergueta-Redondo *et al.*, 2008). É utilizado um sistema de classificação do tumor maligno desenvolvido pela *International Union Against Cancer* (UICC) e pela *American Joint Committee on Cancer* (AJCC), o sistema TNM (do inglês *Tumor-Node-Metastasis*), que pretende avaliar o estágio do tumor tendo em consideração os seus atributos morfológicos, nomeadamente: o tamanho do tumor primário (T), extensão de nódulos linfáticos afectados (N) e a presença de metástases distantes (M) (Singletary e Connolly, 2006). A utilização deste sistema de classificação na caracterização dos tumores é essencial para a selecção da abordagem clínico/terapêutica a aplicar.

Contudo, e devido à grande heterogeneidade atribuída a esta neoplasia, especialmente no que respeita aos IDCs, estes tumores são ainda classificados molecularmente de acordo com a expressão dos seus receptores hormonais: receptor de estrogénio (ER), receptor de progesterona (PR), e na sob-expressão/amplificação do oncogene *ERBB2* (*HER2*). A classificação molecular permite classificar o cancro da mama em vários sub-tipos de acordo com o perfil de expressão destes genes (Huber *et al.*, 2009), estabelecendo uma correlação entre a sua expressão e os diferentes resultados clínicos, nomeadamente resposta terapêutica, tempo de sobrevida à doença e, até mesmo, sobrevida global inerente ao fenótipo que caracteriza cada um. Molecularmente os casos de cancro da mama dividem-se nos seguintes sub-tipos: tipo Luminal A e B, tipo Basal, HER2-positivo e tipo normal (em que a percentagem de células normais está sob-representada na amostra tumoral) (Badve e Nakshatri, 2009; Hergueta-Redondo *et al.*, 2008). Esta nomenclatura é baseada no facto dos genes serem expressos em dois dos componentes da glândula mamária normal (células epiteliais e mioepiteliais) (Hergueta-Redondo *et al.*, 2008; Vargo-Gogola e Rosen,

2007). Os fenótipos característicos para cada um destes sub-tipos são: Luminal A (ER⁺, HER2⁺), e representa o quadro clínico mais favorável, Luminal B (ER⁺, HER2⁻), HER2 (ER⁻, HER2⁺) e Basal (ER⁻, HER2⁻), representativo do quadro clínico mais agressivo (Cianfrocca e Gradishar, 2009; Huber *et al.*, 2009). Ocorre ainda em cerca de 10 – 17% dos casos de cancro da mama um fenótipo conhecido como triplo negativo (ER⁻, PR⁻ e HER2⁻), são menos frequentes, embora muito agressivos do ponto de vista clínico-patológico, apresentam fraco prognóstico, e alvo terapêutico limitado, sendo a quimioterapia a única opção (Irvin, Jr. e Carey, 2008). A grande agressividade deste tipo de tumor está ainda relacionada com o pico de recorrência entre o primeiro e o terceiro ano, sendo que a maioria das mortes ocorre nos primeiros 5 anos, após a terapia (Irvin, Jr. e Carey, 2008; Reis-Filho e Tutt, 2008).

Na realidade estes marcadores moleculares são de grande importância no prognóstico clínico, no entanto doentes com um quadro clínico-patológico idêntico podem apresentar respostas terapêuticas diferentes. Um dos grandes desafios para os próximos anos reside na tentativa de integrar a classificação molecular atribuída para o cancro da mama com a classificação morfológica e com as alterações moleculares já estabelecidas.

1.3 EPIDEMIOLOGIA

Entre 1951 e 1990 assistiu-se a um aumento significativo na taxa de mortalidade por cancro da mama na grande maioria dos países Europeus. Passadas cerca de duas décadas, a tendência tem-se vindo a alterar, diminuindo gradualmente com o contínuo e constante desenvolvimento de novas terapêuticas. No entanto, é ainda a doença oncológica mais frequente entre as mulheres e a primeira causa de morte por patologia neoplásica no sexo feminino. Mesmo sendo um tumor de grande incidência, é também o que apresenta um prognóstico mais favorável devido à precocidade de detecção, o que conduz à diminuição na taxa de mortalidade de mulheres com neoplasia mamária aumentando a taxa de sobreviventes (Bastos *et al.*, 2007).

Com base em estudos epidemiológicos efectuados em diferentes populações, foram estabelecidos prováveis factores de risco para o cancro da mama, nos quais se

incluem: idade, localização geográfica, *status* socioeconómico, eventos reprodutivos, terapia de substituição hormonal, factores associados a estilo de vida, exposição à radiação ionizante, bem como a predisposição genética individual (Dumitrescu e Cotarla, 2005; McPherson *et al.*, 2000).

Existe uma diferença significativa nas taxas de incidência e mortalidade por cancro da mama que permitem separar as regiões geográficas por áreas de alto e baixo risco. Na realidade, a variação na taxa de incidência deste tipo de cancro entre diferentes regiões geográficas pode ser atribuída a diferenças genéticas entre populações e/ou diferenças significativas quanto ao estilo de vida, incluindo dieta e exposição ambiental (Dumitrescu e Cotarla, 2005). Os estudos efectuados em populações migrantes representam uma sólida evidência da componente ambiental (antes mesmo da componente genética) na incidência do cancro da mama (Bray *et al.*, 2004). A incidência desta neoplasia aumenta em pessoas que tenham emigrado de um país de baixa incidência (por exemplo, países asiáticos) para outro de elevada incidência (por exemplo, Estados Unidos da América), afectando igualmente a sua descendência com um aumento de risco nas sucessivas gerações (Bray *et al.*, 2004; Dumitrescu e Cotarla, 2005; McPherson *et al.*, 2000), o que revela, uma vez mais, uma componente ambiental muito forte no desenvolvimento de cancro da mama.

A incidência de cancro da mama em mulheres com menos de 25 anos é muito baixa aumentando cerca de 100x a partir dos 45 anos, o que coincide com a idade média de aparecimento da menopausa (Bray *et al.*, 2004; Dumitrescu e Cotarla, 2005), padrão que sugere o envolvimento das hormonas reprodutivas na etiologia do cancro da mama. De facto, menos de 5% dos casos são descobertos antes dos 35 anos sendo a maioria diagnosticados em mulheres acima dos 50 anos (Keen e Davidson, 2003; Schedin, 2006). No entanto, há que considerar que uma taxa de incidência elevada para uma determinada faixa etária numa dada etnia, pode não o ser quando se considera outra população distinta. Tomando como exemplo a população americana, sabe-se que a incidência é maior em mulheres caucasianas do que em mulheres afro-americanas. Contudo, esta tendência é contrariada quando se considera a faixa etária dos 25-29 anos, revelado por um aumento de incidência de 1,5x em mulheres afro-americanas relativamente a mulheres caucasianas (Masi e Olopade, 2005).

O *status* socioeconómico parece relacionar-se, igualmente, com um ligeiro aumento de risco em mulheres das classes económicas mais altas, bem como o facto de viverem em centros urbanos quando comparado com mulheres que vivam no campo (Bray *et al.*, 2004; Dumitrescu e Cotarla, 2005; Masi e Olopade, 2005). Este aumento de risco é, no entanto, compensado por uma taxa de mortalidade inferior devido à capacidade de diagnóstico precoce e à terapêutica adequada disponível (Bray *et al.*, 2004). As diferenças socioculturais entre grupos étnicos e raciais, são também um factor epidemiológico preponderante a considerar, não obstante as diferenças no perfil genético de cada grupo que alteram a susceptibilidade para o cancro devido a factores genéticos, como por exemplo mutações em genes que confirmam elevado risco para a doença (Mor e Oberle, 2008; Neuhausen, 1999).

1.4 FACTORES DE RISCO

1.4.1 FACTORES GENÉTICOS

A última década testemunhou uma grande evolução na compreensão dos mecanismos moleculares envolvidos na progressão do cancro da mama, e uma evolução no sentido das alterações genéticas poderem exercer um papel importante na detecção precoce, diagnóstico e tratamento desta neoplasia.

Sabe-se que aproximadamente 5 a 10 % dos casos de cancro da mama, em todas as populações, surge em indivíduos que tenham herdado mutações em genes de alta penetrância, como sendo os genes *BRCA1* e *BRCA2* (Olopade *et al.*, 2008). Na realidade, o risco aumenta significativamente em mulheres que tenham história familiar da doença, especialmente em casos de familiares de primeiro e segundo grau (Dumitrescu e Cotarla, 2005). Os indivíduos portadores de mutações num destes genes apresentam 40 a 80% de probabilidade de desenvolverem esta neoplasia (Fackenthal e Olopade, 2007).

De salientar, que foram identificadas mutações cujo efeito fundador reporta a muitas gerações anteriores em grupos étnicos isolados nos quais se incluem os judeus Ashkenazi, Islandeses e Finlandeses (Bradbury e Olopade, 2007; Fackenthal e Olopade, 2007; Olopade *et al.*, 2008). No entanto, a variabilidade genética limitada nestes

grupos permite uma avaliação mais fidedigna da importância destes genes neste tipo de cancro.

Estão descritas algumas mutações, mesmo que raras, em genes associados a síndromes hereditárias, dos quais se salientam Síndrome de Li-Fraumeni (*TP53*), Síndrome de Cowden (*PTEN*), Síndrome de Peutz-Jeghers (*STK11/LKB1*) e Ataxia-telangiectasia (*ATM*). Estes genes parecem correlacionar-se, igualmente, com o aumento de risco para cancro da mama (Figura 1.1) (Bradbury e Olopade, 2007; Olopade *et al.*, 2008), embora os genes *BRCA1* e *BRCA2* sejam os mais fortemente relacionados com a vertente hereditária desta neoplasia. Os genes *BRCA* desempenham um papel importante na manutenção da integridade do genoma e no controlo da reparação de quebras em cadeia dupla por recombinação homóloga (Bradbury e Olopade, 2007; Nagaraju e Scully, 2007; Venkitaraman, 2002). Muitos autores sugerem, ainda, a existência de um novo gene de alta penetrância associado com o cancro da mama, denominado *BRCAx* ou *BRCA1/2* negativo, pelo facto de estar associado com a componente familiar da doença, sem no entanto revelar qualquer mutação nos genes *BRCA1* e *BRCA2*. No entanto, a opinião da comunidade científica é controversa (Hopper, 2001; Lacroix e Leclercq, 2005; Mangia *et al.*, 2008; Oldenburg *et al.*, 2007). Há autores que defendem que a incidência de tumores *BRCAx* está associada com síndromes raras, nos quais o cancro da mama é apenas uma componente, outros há que preconizam que este tipo de tumores resulta da expressão de mutações, em vários genes, de elevada prevalência mas fracamente penetrantes (Lacroix e Leclercq, 2005). Muitos estudos têm sido realizados para tentar caracterizar os tumores tipo *BRCAx* na tentativa de se encontrarem genes associados, no entanto, os resultados têm-se revelado inconclusivos.

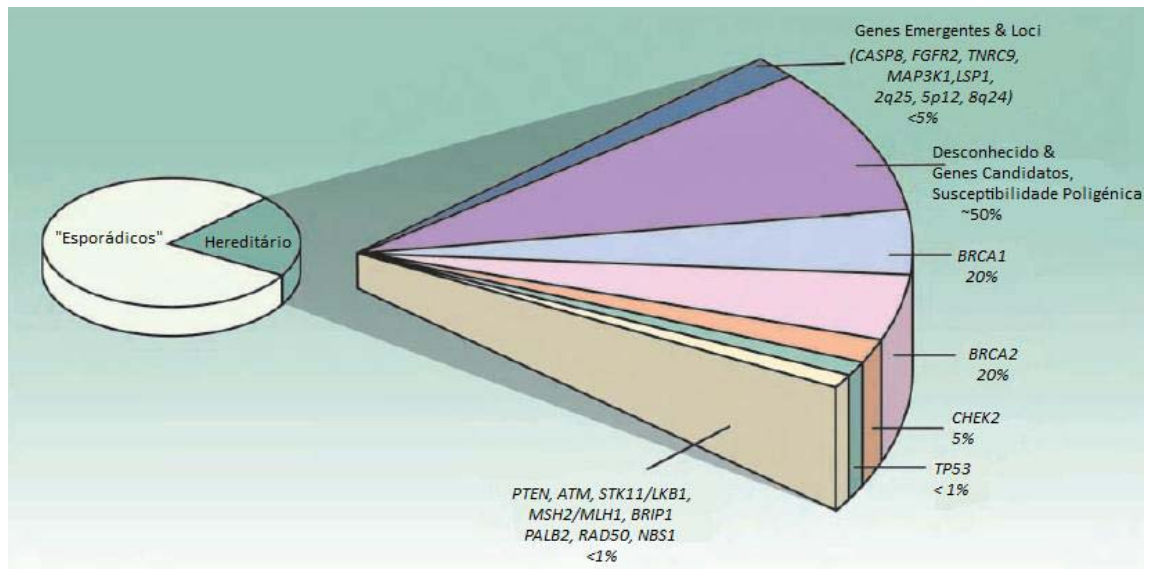


Figura 1.1 - Susceptibilidade genética para o cancro da mama (Adaptado de(Olopade *et al.*, 2008).

Na realidade, 90 a 95% dos casos de cancro da mama são esporádicos e ocorrem sem qualquer mutação nos genes referidos anteriormente. As formas esporádicas poderão estar relacionadas com polimorfismos em genes associados ao metabolismo de cancerígenos, de hormonas esteróides e à reparação de lesões genéticas no DNA (Kolonel *et al.*, 2004). Normalmente são formas variantes destes genes muito comuns na população, e que conferem um risco baixo a moderado na susceptibilidade individual para cancro da mama. No entanto, a combinação dos diferentes polimorfismos pode conduzir a um efeito aditivo, pelo que tudo indica que as formas esporádicas resultem da complexa interacção entre a expressão dos genes de baixa penetrância e factores ambientais (Costa *et al.*, 2007; Lacroix e Leclercq, 2005; Oldenburg *et al.*, 2007).

1.4.2 FACTORES NÃO-GENÉTICOS

Factores não-genéticos (Figura 1.2) são todos aqueles factores que estão relacionados com eventos reprodutivos (menarca, menopausa, gravidez e amamentação), hormonas externas (terapia de substituição hormonal e contraceptivos orais), factores associados com estilo de vida (dieta, exercício físico, consumo de

bebidas alcoólicas), exposição ambiental e ocupacional (por exemplo, exposição a radiação ionizante).

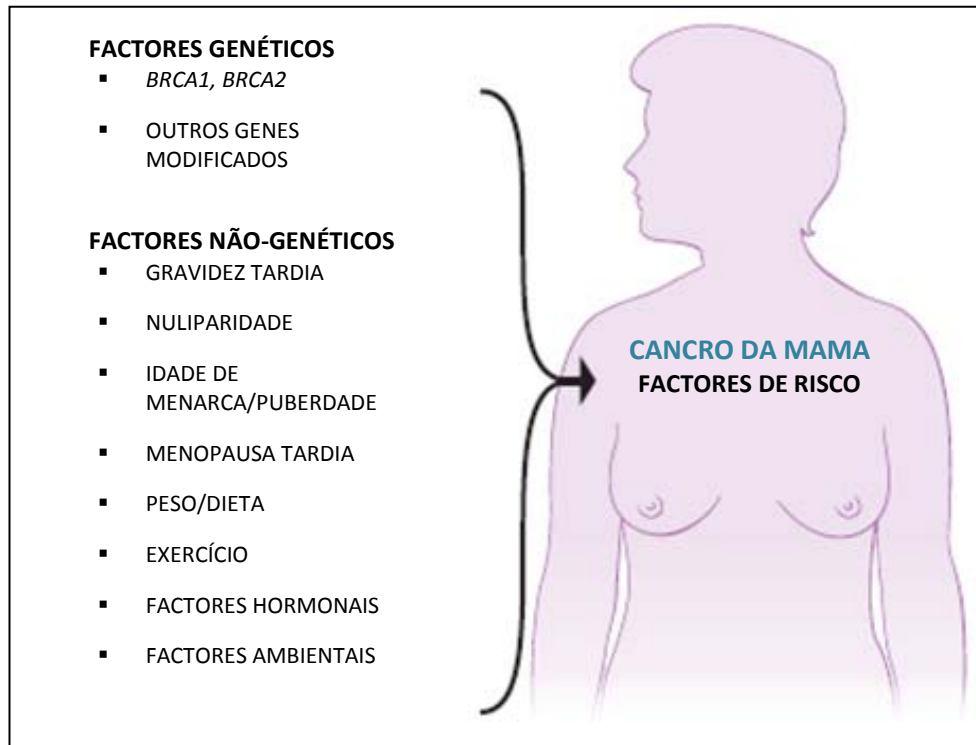


Figura 1.2 - Alguns fatores genéticos e não-genéticos associados ao risco para cancro da mama (Adaptado de (Levy-Lahad e Plon, 2003).

Os fatores relacionados com estilo de vida e exposição ambiental têm sido fonte de muitos estudos epidemiológicos dada a sua preponderância na progressão do cancro da mama.

Terapia de Substituição Hormonal

A utilização de hormonas exógenas, tais como as utilizadas na terapia de substituição hormonal (HRT) (especialmente a combinação estrogénio/progestinas) e nos contraceptivos orais, induz um aumento de risco para neoplasia mamária. A terapia de substituição hormonal começou a ser utilizada em mulheres pós-menopausicas como terapêutica no controlo de sintomas associados à menopausa e na prevenção da osteoporose (Singletary, 2003). No entanto, verificou-se que a utilização desta terapia estaria a aumentar o risco no desenvolvimento de cancro da mama, dependendo da duração e tipo de terapêutica (Collins *et al.*, 2005; Conner *et al.*, 2008; Dumitrescu e Cotarla, 2005). Fatores inatos e hormonais são importantes indicadores de resposta individual à terapia hormonal, ambos consideram a

proliferação celular e as alterações na densidade mamográfica (um dos factores de controlo de progressão mamária efectuado no decorrer da terapêutica). Parece evidente que o risco associado ao desenvolvimento de neoplasia mamária deva depender da dosagem de hormonas administradas na terapêutica, não podendo ser esquecido o facto de cada mulher reagir de forma diferente à dose e ao tipo de tratamento (Conner *et al.*, 2008).

Contraceptivos Orais

Quanto à utilização de contraceptivos orais, existe um pequeno aumento no risco relativo para cancro da mama entre mulheres que os utilizam mas, e uma vez que a sua utilização tem início numa idade mais jovem, quando o risco é muito baixo, o referido aumento tem pouco efeito na taxa de incidência global (Oldenburg *et al.*, 2007).

Dieta

Muitos estudos epidemiológicos sobre o papel da dieta no desenvolvimento do cancro da mama têm sido realizados, revelando que existem compostos derivados do processamento alimentar que possam estar associados a um aumento de risco no desenvolvimento de neoplasia mamária, assim como outros que promovem uma diminuição de risco. A dieta expõe o indivíduo a um grande e variável conjunto de cancerígenos químicos e naturais, alguns dos quais são responsáveis por provocar danos no DNA, e por compostos conhecidos pelas suas características anti-cancerígenas. Alguns destes compostos podem actuar através da formação de espécies reactivas de oxigénio (ROS), conhecidas pela sua capacidade de lesar o DNA, ou ainda outros compostos lesivos, nos quais se incluem as aminas aromáticas heterocíclicas (HCA) (produzidas durante o processamento de alimentos como a carne ou o peixe) (Weyandt *et al.*, 2008). As HCA, são formadas pela pirólise de aminoácidos e proteínas durante o processamento dos alimentos, e são responsáveis pela formação de aductos de DNA. Por sua vez, a exposição a hidrocarbonetos aromáticos policíclicos (PAHs) proveniente do fumo do tabaco, poluição do ar e escapes de automóveis, podendo estar também presentes em alguns alimentos, têm a capacidade de se ligarem ao DNA formando aductos mutagénicos (Gammon e Santella, 2008; Weyandt *et al.*, 2008).

Outro factor de risco associado com a dieta e que se encontra bem documentado em estudos epidemiológicos é o excesso de peso. A obesidade parece relacionar-se com o *status* menopausico estando associada com um significativo aumento de risco para esta neoplasia, sendo incontestável em mulheres pós-menopausicas, possivelmente por estar associado com o aumento dos níveis de estrogénios endógenos em mulheres obesas, já que o tecido adiposo é por si só uma fonte importante destas hormonas (Carmichael *et al.*, 2009; Dumitrescu e Cotarla, 2005; Vogel, 2008).

O consumo de álcool, por sua vez, tem sido o factor associado à dieta mais estudado, estando relacionado com um comprovado aumento de risco para esta neoplasia associado a um moderado aumento dos estrogénios endógenos (Boyle, 2005; Conner *et al.*, 2008; Dumitrescu e Cotarla, 2005).

Radiação Ionizante

A radiação ionizante provoca lesões no DNA quer através de interacções biofísicas quer físico-químicas, sendo a extensão das lesões celulares proporcional à quantidade de radiação recebida. Existe, assim, uma evidente relação entre lesão de DNA, mutações e desenvolvimento de cancro (Mullenders *et al.*, 2009). A lesão induzida por radiação ionizante constitui uma ampla série de lesões de bases e quebras de cadeia substancialmente coincidente com as lesões de DNA produzidas pelo metabolismo oxidativo endógeno (Mullenders *et al.*, 2009). Normalmente, a deposição da radiação que origina a lesão do DNA é manifestada por quebras de cadeia simples (SSBs), lesões pré-mutagénicas que podem ser reparadas através de mecanismos específicos, e de cadeia dupla (DSBs), as mais frequentes e que conduzem a mutações genéticas que desempenham um papel crucial no processo de cancerigénese (MacMahon, 2006; Mullenders *et al.*, 2009). Este tipo de radiação é um factor exógeno muito bem documentado como estando associado a um aumento de risco para neoplasia mamária, derivado da exposição a doses elevadas de radiação ionizante (MacMahon, 2006). Exemplos de exposição a elevados níveis de radiação ionizante são os que incluem mulheres expostas às explosões atómicas, consecutivos fluoroscopias para tratamento da tuberculose e tratamento para a doença de Hodgkin's. O risco associado depende da dose de exposição e diminui gradualmente ao

longo do tempo (Dumitrescu e Cotarla, 2005; Oldenburg *et al.*, 2007). Sabe-se que o tecido mamário é altamente sensível aos efeitos cancerígenos da radiação, em particular quando a exposição ocorre durante a infância/puberdade (Golubicic *et al.*, 2008; MacMahon, 2006). Na realidade, todas as mulheres estão expostas a doses baixas de radiação (por razões ocupacionais, procedimentos de diagnóstico médico ou mesmo exposição ambiental) cujos efeitos no desenvolvimento de cancro da mama não são adequadamente documentados (Golubicic *et al.*, 2008). Actualmente, pensa-se que a exposição terapêutica à radiação seja a causa mais significativa de cancerigénese induzida por radiação. No entanto, a razão risco/benefício para a exposição derivada da radioterapia não está totalmente clara, na medida em que as doses de radiação utilizadas na terapêutica representam claramente um efeito benéfico que compensa o risco de exposição (Mullenders *et al.*, 2009).

Factores Hormonais

Todas as evidências experimentais e epidemiológicas parecem sugerir que a exposição a estrogénios promove a cancerigénese mamária, podendo ser justificado por duas vias distintas. Por um lado, os produtos derivados do metabolismo de estrogénios podem induzir lesões no DNA através da formação de aductos e de bases oxidadas, provocando mutações em oncogenes e genes supressores de tumor, responsáveis pelo controlo do crescimento e proliferação celular. Por outro, estas hormonas podem alterar a expressão de genes especificamente envolvidos na estimulação do crescimento e proliferação das células epiteliais da mama (Weyandt *et al.*, 2008).

A grande maioria dos factores de risco descritos anteriormente está associada com a exposição a compostos químicos que alteram a normal regulação dos processos biológicos. Para além de metabolitos endogenamente formados, existem ainda aqueles que derivam da exposição ambiental e ocupacional, muitos dos quais actuam por apresentarem uma estrutura que se assemelha à dos estrogénios naturais, ligando-se muitas vezes aos receptores de estrogénios existentes nas células, exemplo disto são as HCA e os PHA.

Por outro lado, os níveis hormonais, especialmente hormonas sexuais, alteram-se naturalmente no decorrer da vida de uma mulher. Factores como: menarca precoce

(antes dos 12 anos), menopausa tardia (depois dos 55 anos) e nuliparidade estão correlacionadas com aumento de risco para cancro da mama, o que sugere que uma exposição prolongada a elevados níveis de estrogénios e progesterona, devido a um elevado número de ciclos ovulatórios, contribua para o desenvolvimento da doença (Dumitrescu e Cotarla, 2005; Keen e Davidson, 2003). No entanto, o papel atribuído aos estrogénios no desenvolvimento da neoplasia mamária é incerto, uma vez que estas hormonas actuam como promotores de proliferação dos epitélios normal (principalmente durante a puberdade e a gravidez) (Hilakivi-Clarke, 2000) e neoplásico da mama (Russo *et al.*, 2003).

Estudos epidemiológicos evidenciam que a menopausa cirurgicamente induzida (ovariectomia ou histerectomia) entre os 35 e os 45 anos diminui o risco de cancro da mama em cerca de 60% relativamente a mulheres que atinjam a menopausa naturalmente (Dumitrescu e Cotarla, 2005; Oldenburg *et al.*, 2007). A idade a que ocorre a primeira gravidez influencia também o risco de desenvolvimento de cancro da mama. Estudos epidemiológicos indicam que uma gravidez em mulheres jovens (menos de 20 anos) induz uma diminuição de risco no desenvolvimento desta neoplasia (Cuzick, 2008; Dumitrescu e Cotarla, 2005; Singletary, 2003), possivelmente associado ao desenvolvimento e diferenciação da mama característicos da gravidez (Russo *et al.*, 2005b; Russo *et al.*, 2005a). Multiparidade é outro factor e, está associada com uma diminuição de risco no desenvolvimento da neoplasia. Em contraste, nuliparidade e idade tardia de uma primeira gravidez contribuem para um significativo aumento de risco na progressão da doença, aumento este que é ainda maior em mulheres cuja primeira gravidez ocorra depois dos 35 anos (Dumitrescu e Cotarla, 2005). A amamentação prolongada parece exercer um efeito protector no desenvolvimento do cancro da mama, o que poderá estar associado com a esfoliação das células ductais como consequência da amamentação e que poderá ser responsável pela eliminação de um número significativo de células com lesão, prevenindo assim a sua transformação em células neoplásicas (Silva *et al.*, 2006b; Thompson *et al.*, 2002).

A razão benefício/risco do papel dos estrogénios resume-se então ao facto de, para além de serem necessários para o bom funcionamento de alguns sistemas biológicos, podem também aumentar a proliferação e a instabilidade genética, possivelmente pela indução de radicais livres que dão origem a lesões e mutações no

DNA. A instabilidade genética aumenta a probabilidade das células normais se transformarem progressivamente em células malignas. Deste modo, é crucial garantir que ao mesmo tempo que os estrogénios são de extrema importância para o funcionamento metabólico, também existam outros mecanismos que activam as vias de reparação como resposta às lesões provocadas por estas hormonas (Hilakivi-Clarke, 2000).

A existência de metabolitos de estrogénio oxidativos, nomeadamente quinonas de catecol-estrogénio, que reagem com o DNA dão ênfase à acção destes metabolitos como cancerígenos químicos endógenos capazes de lesar o DNA e dessa forma serem os responsáveis pela iniciação do processo de cancerigénese (Gaikwad *et al.*, 2008).

Como se pode observar na figura 1.3, a biossíntese de estrogénios envolve uma série de passos enzimáticos desde o colesterol até à formação do estradiol e da estrona, onde estão presentes enzimas da superfamília do Citocromo P450 (CYPs), sendo os de maior importância o CYP11A, o CYP17 e o CYP19 e as enzimas hidroxisteróides desidrogenases entre elas a 17β -HSD (Mitrunen e Hirvonen, 2003). A divisão da cadeia de colesterol pelo CYP11A para formar pregnenolona e progesterona é considerado um passo limitante na biossíntese de todos os esteróides. Os mineralcorticóides, glucocorticóides e os androgénios (androstenediona e testosterona) são sintetizados a partir da progesterona. O CYP19 cataliza os passos finais de formação de estrogénios a partir dos androgénios, onde a androstenediona dá origem à estrona e a testosterona dá origem ao estradiol (Mitrunen e Hirvonen, 2003).

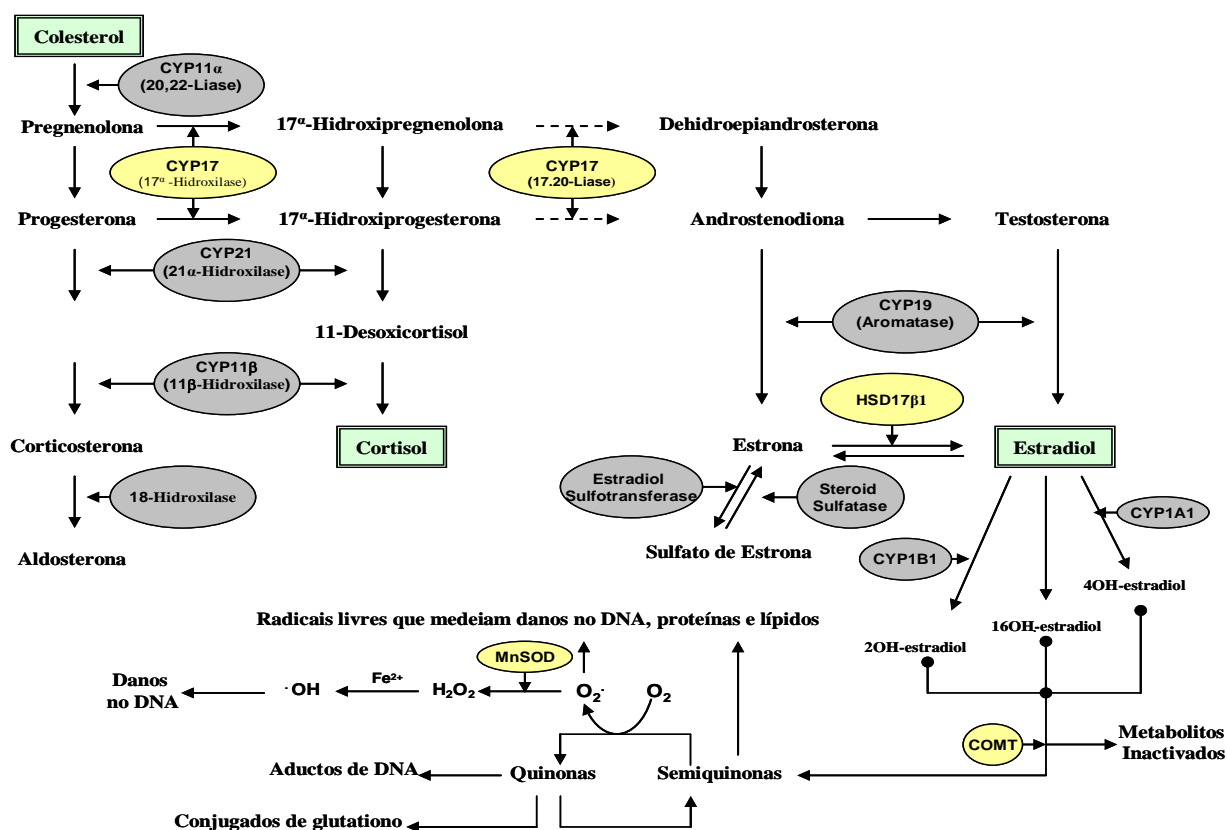


Figura 1.3 – Apresentação esquemática da biossíntese e inativação metabólica de estrogénios (Adaptado de(Mitrunen e Hirvonen, 2003).

Após a biossíntese de estrogénios, estes vão ser hidroxilados, levando à formação de catecol-estrogénios, nomeadamente 2-hidroxiestradiol (2OH-E₂), 4-hidroxiestradiol (4OH-E₂) e 16-hidroxiestradiol (16OH-E₂). Para que os catecol-estrogénios possam ser convertidos em metabolitos inactivados para posterior excreção, é necessária a sua inactivação por metilação mediada pela enzima COMT (Catecol-*O*-Metil Transferase), prevenindo a formação de quinonas e do ciclo redox (Santen *et al.*, 2009; Silva *et al.*, 2006b). Apesar da via mais activa de conjugação dos metabolitos ser a metilação, podem também ser conjugados por sulfatação (conjugação do estrogénio com sulfato) e glucuronidação (conjugação do estrogénio com ácido glucorónico), de modo a formarem compostos excretáveis pela urina e bÍlis. Como referido anteriormente, os catecol-estrogénios para além de poderem dar origem a metabolitos inactivos, podem levar à formação de semi-quinonas e consequentemente quinonas, que reagem com o DNA formando aductos que podem induzir a fixação de mutações (Gaikwad *et al.*, 2008). Durante o ciclo redox decorrente da conversão de semi-quinonas a quinonas pode haver a formação de espécies

reactivas de oxigénio (ROS) (Santen *et al.*, 2009), que por sua vez são responsáveis pela formação de lesões oxidativas em lípidos, proteínas e mesmo no DNA. Quando ocorre a formação de quinonas, estas podem ser destoxificadas por glutathione S-transferases (GSTs) e dar origem a conjugados de glutathione ou podem ser reduzidas a derivados de catecol-estrogénios. As semi-quinonas podem reagir com o oxigénio molecular e formar radicais superóxido que por sua vez podem ser reduzidos a peróxido de hidrogénio (H₂O₂) espontaneamente ou catalizados pela superóxido dismutase (SOD) (Silva *et al.*, 2006b). O peróxido de hidrogénio, na presença de iões metálicos reduzidos (como por exemplo Fe²⁺), pode levar à formação de um oxidante mais potente, o radical hidroxilo (*OH) (Maynard *et al.*, 2009; Mitrunen e Hirvonen, 2003; Russo *et al.*, 2003). Deste modo, será possível intuir que o risco de desenvolvimento de qualquer tipo de cancro, está relacionado com a capacidade individual para destoxificar metabolitos cancerígenos e/ou reparar lesões no DNA provocadas por todos estes factores.

Assim sendo, é correcto afirmar-se que, a susceptibilidade para cancro da mama pode ser mediada, em parte, pela variabilidade genética de genes envolvidos nas vias de metabolização e destoxificação de metabolitos reactivos e na reparação da lesão genética.

1.5 SUSCEPTIBILIDADE INDIVIDUAL

1.5.1 POLIMORFISMOS GENÉTICOS E VARIABILIDADE INTER-INDIVIDUAL

A variabilidade genética individual é, há muito, um importante factor que influencia a susceptibilidade individual de desenvolvimento de doenças, nomeadamente do cancro. Com o desenvolvimento do projecto de sequenciação do genoma humano, foram identificadas variações polimórficas na sequência de DNA de inúmeros genes (Figura 1.4), a maioria das quais alterações pontuais que ocorrem numa única base designando-se assim de *Single Nucleotide Polymorphism* (SNP). Porém, apesar de existirem milhões de SNPs no genoma humano, pensa-se que apenas cerca de 30.000 tenham efeitos fenotipicamente visíveis em termos clínicos (Kruglyak e Nickerson, 2001; Sachidanandam *et al.*, 2001; Tempfer *et al.*, 2006). Os SNPs representam a principal fonte de variabilidade genética e fenotípica (Feuk *et al.*, 2006;

Ford *et al.*, 2000), e podem ser encontrados em todas as regiões genómicas, desde exões, intrões e regiões promotoras/reguladoras e regiões intergénicas. A grande variedade de localização dos SNPs torna este tipo de variações polimórficas mais susceptíveis de provocar relevância alélica funcional ou fisiológica face a outros tipos de polimorfismos. A maioria dos estudos epidemiológicos, que tenham como expressa finalidade a identificação de SNPs associados com a susceptibilidade individual para patologia multifactorial, incide essencialmente sobre SNPs localizados em regiões codificantes ou em regiões reguladoras do genoma. Alterações nestas regiões genómicas são passíveis de apresentar qualquer consequência fenotípica dada a forte possibilidade de afectarem a função ou expressão de uma determinada proteína (Calladine *et al.*, 2004; Feuk *et al.*, 2006; Gaspar *et al.*, 2006).

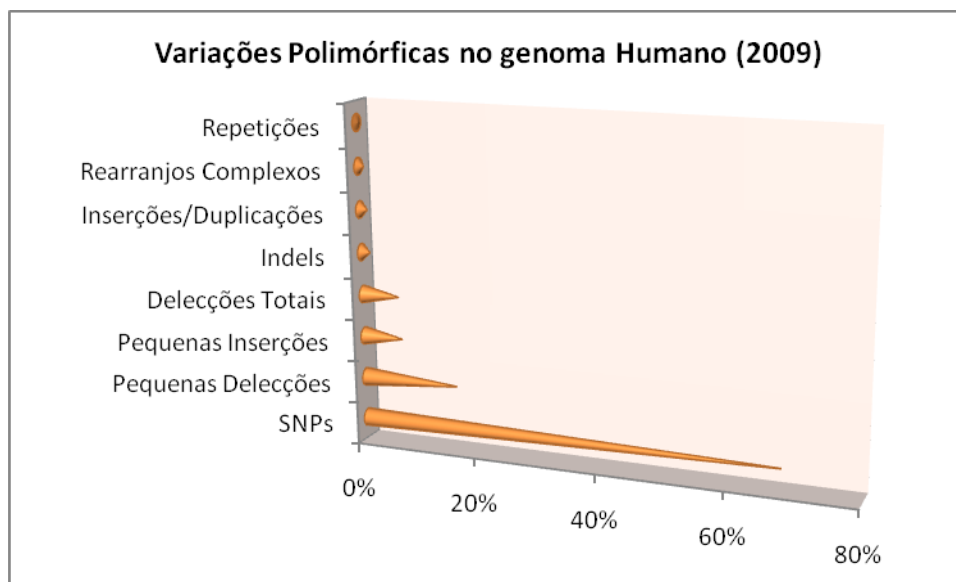


Figura 1.4 – Distribuição percentual das diferentes variações polimórficas no genoma humano, de acordo com a base de dados HGMD¹ (dados de 2009).

Como já foi referido, o cancro da mama é considerado uma doença multifactorial na qual exposições múltiplas, a factores endógenos e a cancerígenos provenientes de exposição exógena, interagem com o património genético do indivíduo de forma complexa e resultam na modulação do risco para contrair a doença. O estudo de polimorfismos em genes de baixa penetrância, normalmente

¹ <http://www.hgmd.cf.ac.uk>

associados a mecanismos bioquímicos ou fisiológicos, pretende avaliar e identificar alguns genes de susceptibilidade para a cancerígenese mamária (Dumitrescu e Cotarla, 2005). Os genes candidatos são, normalmente, aqueles que codificam para enzimas envolvidas no metabolismo de estrogénios (CYP17, COMT) ou de vários cancerígenos (CYPs, GSTs, NATs), destoxificação de ROS produzidos por estas vias (SOD), enzimas envolvidas na reparação da lesão genética (XRCC1 – 3) ou em processos de sinalização celular. Os efeitos fenotípicos dos polimorfismos são baseados em efeitos genéticos directos e em interacções gene-gene e gene-ambiente (Tempfer *et al.*, 2006).

1.5.2 ESTUDOS DE ASSOCIAÇÃO

Até há algum tempo as metodologias disponíveis para estudos de associação permitiam apenas a avaliação individual de um pequeno número de polimorfismos, apresentando-se como uma ferramenta importante na detecção de alelos de baixa penetrância (Easton e Eeles, 2008). Gradualmente e com o desenvolvimento de novas plataformas que permitem analisar centenas de milhares de SNPs foi possível aumentar a previsibilidade dos estudos de associação numa abordagem que conduziu à criação de haplótipos (arranjos específicos de alelos). Embora o efeito individual de um SNP seja diminuto, o resultado genético de combinações entre SNPs e/ou alelos funcionais, próximos entre si e herdados em bloco (haplótipos), poderão contribuir, de forma aditiva, para o aumento de risco no desenvolvimento do cancro, nomeadamente no cancro da mama (Onay *et al.*, 2006; Pharoah *et al.*, 2007). Felizmente, não é necessário proceder-se à genotipagem de todas as variantes alélicas possíveis para detectar uma associação entre os diferentes SNPs potencialmente envolvidos numa patologia, uma vez que existem SNPs que se encontram fisicamente próximos entre si que podem ser correlacionados. A este fenómeno dá-se o nome de Desequilíbrio de *linkage* (LD), ou seja, o facto de determinados alelos poderem co-ocorrer no mesmo haplótipo mais frequentemente do que o expectável definido pelo acaso (Pharoah *et al.*, 2004).

Estudos de associação genética têm sido desde sempre muito utilizados com a perspectiva de se identificarem alelos de susceptibilidade para o cancro. O mais importante neste tipo de estudos é a determinação de uma associação entre o facto

de se ser portador de um alelo ou genótipo e a aquisição de uma determinada doença (Tempfer *et al.*, 2006). Normalmente este tipo de estudos concentra-se em genes candidatos ou mecanismos que se suspeitem importantes no desenvolvimento da doença (Easton e Eeles, 2008). Com o aumento do número de estudos de associação com o intuito de se estabelecer uma relação entre genótipos polimórficos e alterações fenotípicas, desenvolveu-se uma nova abordagem que permite aumentar a credibilidade deste tipo de estudos, e que se designa por estudos de meta-análise. Estes estudos baseiam-se na análise da compilação de estudos de associação previamente publicados e confirmam ou não as associações genótipo/fenótipo para a doença em causa (Tempfer *et al.*, 2006). Para o cancro de mama alguns estudos de meta-análise foram já publicados que confirmam ou refutam a associação de vários SNPs e a vertente esporádica desta neoplasia (Bag e Bag, 2008; Wang *et al.*, 2009).

O Projecto de Sequenciação do Genoma Humano permitiu identificar milhões de marcadores genéticos que podem ser utilizados em estudos genéticos de associação. Por sua vez, o desenvolvimento de novas tecnologias moleculares de genotipagem e sequenciação, acrescido de métodos estatísticos sofisticados, tem permitido novas abordagens científicas com o desígnio de definir o papel das variações genéticas no desenvolvimento de doenças, nomeadamente na etiologia do cancro (Zhao *et al.*, 2003). A sequenciação de genes e as iniciativas de genotipagem, tais como o projecto *HapMap* (The International HapMap Consortium, 2003), o *SeattleSNP Variation Discovery Resource*² e o projecto do *SNP500* (Packer *et al.*, 2006), entre outros, conduziram a uma melhor compreensão das diferenças entre populações, das variações genéticas específicas e da estrutura de haplótipos descritos para o genoma humano. Por outro lado, a utilização de plataformas de genotipagem que permitem processar centenas de milhares de SNPs simultaneamente, tem possibilitado desenvolver amplos estudos de associação utilizando conjuntos de SNPs que estejam relacionados com a maioria das variações genéticas mais comuns do genoma (*tagSNPs*), e desta forma percorrer a sua totalidade na busca incessante de *locus* de associação (Easton e Eeles, 2008). Esta estratégia permite recorrer à genotipagem de um menor número de SNPs fornecendo, no entanto, uma descrição exaustiva da

² <http://pga.gs.washington.edu/>

variação genética de um gene, num *locus* específico ou em todo o genoma (Savage, 2008).

Recentemente, surgiram estudos de associação *genome-wide* (GWAS), que funcionam como uma poderosa ferramenta na detecção de *locus* de susceptibilidade para diferentes tipos de cancro, não sendo excepção o cancro de mama (Ahmed *et al.*, 2009; Easton e Eeles, 2008; Thomas *et al.*, 2009). Os estudos GWAS caracterizam-se, porém, em serem “agnósticos”, isto é, por procurarem associações com significâncias estatísticas elevadas para as quais não existe, ou não é colocada, qualquer hipótese prévia de plausibilidade biológica, ao contrário do que acontece nos estudos de gene candidato.

1.6 METABOLIZAÇÃO DE CANCERÍGENOS

Todos os organismos estão continuamente expostos a agentes lesivos provenientes do ambiente circundante. Como resultado, têm desenvolvido mecanismos celulares com vista a minimizar as consequências biológicas provocadas por estes agentes, nomeadamente mecanismos de destoxificação, de reparação, de controlo do ciclo celular, entre outros. Com a sequenciação do genoma humano tornou-se ainda mais relevante o estudo de polimorfismos no âmbito da toxicologia (Singh *et al.*, 2008), ou seja polimorfismos em genes envolvidos na maquinaria de resposta a lesões induzidas por agentes xenobióticos.

As enzimas envolvidas nas vias de metabolização de xenobióticos podem ser divididas em enzimas de fase I (família de citocromos P450), responsáveis pela activação metabólica de cancerígenos, e em enzimas de fase II (família de GSTs) responsáveis pela inactivação de cancerígenos e dos metabolitos intermediários gerados, normalmente, gerados por enzimas de fase I (por exemplo, CYP1A1, CYP2D6 e CYP2E1) (Dumitrescu e Cotarla, 2005; Ketterer, 1998). As enzimas de fase II actuam no sentido de facilitar a conjugação desses metabolitos, lesivos face ao DNA, nomeadamente com o glutationo promovendo a sua subsequente excreção (Ketterer, 1998). Também enzimas de fase II, as N-acetiltransferases (NAT) estão especialmente envolvidas na destoxificação de arilaminas, de alguns compostos presentes no fumo do tabaco, e também de aminas provenientes do processamento alimentar. Contudo, a

acção destas enzimas nestes compostos cancerígenos origina, por vezes, metabolitos electrofílicos responsáveis pela indução de mutações pontuais no DNA.

Polimorfismos presentes em alguns destes genes têm sido descritos como estando associados a diferentes tipos de cancro, nomeadamente no cancro da mama (Torresan *et al.*, 2008). O envolvimento da família dos CYP e da família das GSTs deve-se ao importante papel que estas enzimas desempenham na protecção contra agentes ambientais e metabolitos endógenos, nomeadamente, os produzidos pelo metabolismo de estrogénios, potencialmente envolvidos na cancerigénese mamária (Egan *et al.*, 2004; Han *et al.*, 2004b). Dentro destas famílias, os mais referenciados são os polimorfismos nos genes *GSTM1* e *GSTT1*, ambos deleções de parte dos genes o que conduz à perda de actividade das enzimas influenciando a sua capacidade de destoxificação de agentes cancerígenos, aumentando, deste modo, a possibilidade de desenvolvimento tumoral (Strange *et al.*, 2000; Torresan *et al.*, 2008). O gene *GSTP1* é também amplamente estudado no âmbito da cancerigénese devido às funções que desempenha na destoxificação de compostos cancerígenos, na metabolização de agentes quimioterapêuticos e ainda devido ao seu envolvimento no ciclo celular e na regulação da apoptose (Moyer *et al.*, 2008; Torresan *et al.*, 2008; Van Emburgh *et al.*, 2008). Os genes envolvidos no metabolismo de activação de cancerígenos (*CYP1A1*, *CYP1B1*, *CYP2D6*, entre outros) são responsáveis pela formação de aductos no DNA, de ROS, inclusive provenientes do metabolismo de estrogénios (Cavalieri *et al.*, 2000), entre outros agentes cancerígenos e desencadeiam a resposta que será mediada pelas enzimas de fase II. Deste modo, alguns estudos têm relacionado o efeito combinado de polimorfismos descritos em genes de fase I e fase II, nomeadamente em estudos de cancro da mama, já que são reconhecidos factores de risco para esta neoplasia a exposição a cancerígenos ambientais e a exposição a estrogénios (Cavalieri *et al.*, 2000; Ketterer, 1998; Torresan *et al.*, 2008; Van Emburgh *et al.*, 2008), como foi anteriormente referido.

A acção das enzimas de fase I e de fase II possibilitam, então, a transformação de compostos cancerígenos noutros mais solúveis e, portanto, mais fáceis de excretar favorecendo uma redução no risco. No entanto, os produtos mais solúveis de alguns compostos são cancerígenos muito mais potentes, deste modo, a alteração genética associada a um aumento de expressão de um gene ou na actividade de uma proteína

poderá aumentar a quantidade de metabolitos reactivos formados, aumentando o risco de lesão. Um exemplo clássico é a exposição ao Benzo[a]pireno (B(a)P), esquematizado na figura 1.5.

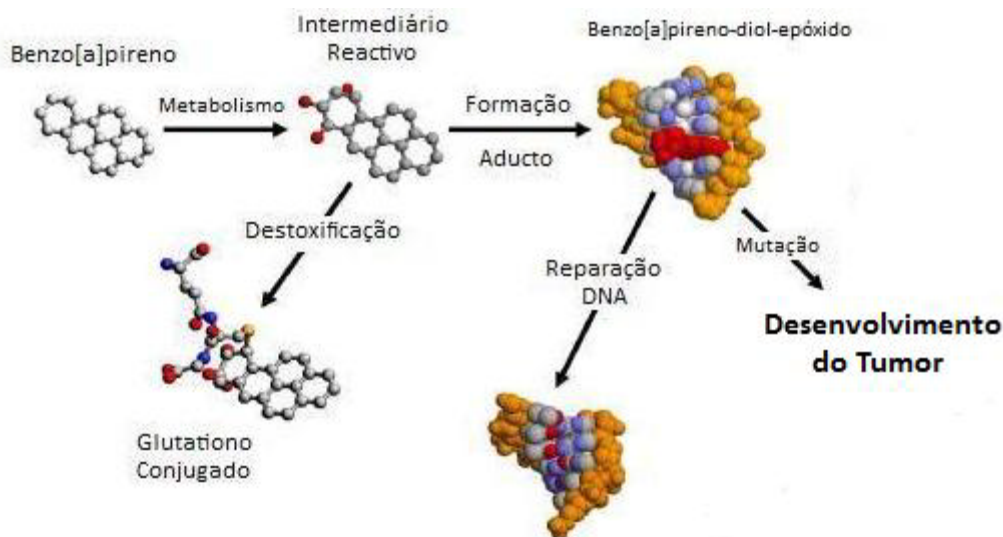


Figura 1.5 - Diagrama esquemático do mecanismo através do qual o Benzo[a]pireno promove o desenvolvimento do cancro (Adaptado de(Rundle, 2006).

A figura 1.5 ilustra, então, os mecanismos que poderão ocorrer na metabolização do Benzo[a]pireno. Após exposição o composto é absorvido e metabolizado num metabolito mais reactivo, processo normalmente desencadeado por enzimas de fase I, que vão originar um intermediário electrofilico com capacidade de reacção com o DNA por ligação covalente formando um aducto (benzo[a]pireno-diol-epóxido). O metabolito intermediário pode também ser destoxificado por enzimas de fase II, normalmente através de um processo de conjugação com glutaciono dando origem a um composto hidrosolúvel que é excretado. Pode, ainda, ser removido através de enzimas envolvidos nas vias de reparação, também estas enzimas polimórficas, caso não o seja dará lugar à fixação de uma mutação. A mutação ao ocorrer num gene crítico, como por exemplo num gene supressor de tumor ou num oncogene, poderá dar origem ao desenvolvimento do tumor (Rundle, 2006). Polimorfismos em enzimas intervenientes em qualquer destas vias comprometem a sua actividade catalítica, e logo todos os processos inerentes à biotransformação do composto. Polimorfismos nos genes que codificam para enzimas envolvidas nos

processos de reparação de DNA podem levar a uma diminuição na capacidade de reparação, especialmente após exposição a agentes endógenos ou exógenos, contribuindo para o desenvolvimento de cancro da mama.

Alterações genéticas na sequência dos genes que codificam para estas enzimas poderão alterar a sua actividade enzimática ou via biossintética conduzindo a deficiência nos processos de destoxificação de cancerígenos. SNPs presentes nestes genes podem ser responsáveis pela alteração da resposta individual a fármacos, especialmente agentes quimioterapêuticos, podendo modificar a magnitude do efeito farmacológico, a efectividade terapêutica, o limite de toxicidade, as interacções medicamentosas ou os efeitos secundários que provocam (Wajapeyee e Somasundaram, 2004), podendo preponderar o risco individual de desenvolvimento de cancro, nomeadamente de cancro da mama (Dumitrescu e Cotarla, 2005; Singh *et al.*, 2008).

1.7 REPARAÇÃO DA LESÃO GENÉTICA

As lesões de DNA podem perturbar o estado de equilíbrio celular e activar ou amplificar determinados mecanismos bioquímicos responsáveis pela regulação do crescimento e divisão celular (Sancar *et al.*, 2004). A resposta celular à lesão de DNA e a sua capacidade na manutenção da estabilidade genómica através da reparação de DNA podem ser fundamentais para prevenir a iniciação e a progressão de fenómenos de cancerigénese. Para além disto, a variabilidade genética nos genes de reparação de DNA pode afectar o produto final da sua expressão influenciando a susceptibilidade individual para desenvolvimento de uma neoplasia. As células de mamíferos desenvolveram vias distintas para reparar diferentes tipos de lesão no DNA de forma a manter a sua integridade genómica. Assim sendo, a capacidade celular de reparação de DNA pode de algum modo influenciar a susceptibilidade individual a compostos endógenos e exógenos (Costa *et al.*, 2007; Synowiec *et al.*, 2008).

Quando a célula detecta erros no DNA acciona uma série de mecanismos de sobrevivência, essenciais para a manutenção da integridade do genoma e prevenção na ocorrência de mutações (Hoeijmakers, 2001; Shimada e Nakanishi, 2006). Alguns destes mecanismos alteram a progressão do ciclo celular de forma a dar tempo para

que a maquinaria de reparação de DNA ocorra, impedindo que informações genéticas incorrectas passem para a descendência (Houtgraaf *et al.*, 2006). Os processos de reparação de DNA estão organizados em cinco vias distintas, descritos sumariamente na Tabela 1.1, altamente conservadas, podendo ocorrer subdivisões em algumas destas vias. Cada via tem a capacidade de reparar uma classe específica de lesão de DNA, no entanto pode ocorrer alguma complementaridade entre as diferentes vias (Mohrenweiser *et al.*, 2003).

Tabela 1.1 – Lesões de DNA e respectivas vias de Reparação

AGENTES LESIVOS	TIPO DE LESÃO	VIA DE REPARAÇÃO
Agentes alquilantes	Aductos DNA	Reparação Directa
Erros de replicação	<i>Mismatch</i> base-base, Inserções e Delecções	MMR
ROS, IR, Agentes alquilantes	Base alterada, Local abásico (AP), SSBs	BER
UV, Aductos de cisplatina, PAHs	<i>Cross-links</i> entre bases da mesma cadeia de DNA, Aductos de elevado peso molecular	NER
IR, Agentes alquilantes, Cisplatina	<i>Cross-links</i> entre as cadeias de DNA, DSBs	Reparação DSB (HR e NHEJ)

1.7.1 REPARAÇÃO DIRECTA (DR)

A reparação directa está envolvida na reparação de lesões provocadas por agentes alquilantes. É a única via que envolve uma única proteína, denominada O^6 -metilguanina (O^6 -MeG) DNA metiltransferase (MGMT), que codifica para uma proteína de reparação de DNA O^6 -alquilguanina (O^6 -AG) DNA alquiltransferase (AGT). É a única proteína envolvida em reparação de DNA que actua apenas na remoção de aductos de DNA, já que todas as outras proteínas envolvidas nas diferentes vias de reparação actuam como complexos multi-enzimáticos na remoção das diferentes lesões (Gerson, 2004). Esta proteína é responsável pela transferência do grupo alquil da posição O^6 da guanina para um resíduo de cisteína no seu centro activo, originando a reposição da

composição química do DNA sem haver necessidade de qualquer reconstrução genómica. No entanto, este evento provoca a inactivação irreversível da proteína MGMT, reacção vulgarmente referida como “suicida”, estando associada a um aumento do risco de cancerigénese e aumento da sensibilidade a agentes metilantes (Gerson, 2004; Mohrenweiser *et al.*, 2003).

1.7.2 REPARAÇÃO DE *MISMATCH* (MMR)

A via de reparação *mismatch* é um processo que ocorre durante a replicação de DNA de forma a corrigir os erros de emparelhamento provocados pela DNA polimerase (Helleday *et al.*, 2008). São normalmente erros esporádicos, tais como *mismatch* base-base e/ou inserção/delecção de *loops* que podem ser gerados durante os processos de replicação e recombinação (Mohrenweiser *et al.*, 2003), normalmente em regiões de sequências repetitivas, microsatélites (MS). A maquinaria envolvida neste processo de reparação tem necessariamente que cumprir dois requisitos: primeiro o reconhecimento das lesões e, segundo direccionar a maquinaria de reparação no sentido de reparar todas as informações genéticas erróneas contidas na cadeia sintetizada *de novo* (Jiricny, 2006). Defeitos nesta via de reparação são evidenciados pela instabilidade de microsatélites (MSI) e pela tolerância a efeitos citotóxicos de agentes alquilantes, o que resulta numa elevada taxa mutacional no genoma, bem como ao aumento da resistência à apoptose (Schroering *et al.*, 2007).

A principal componente do sistema geral de reparação do DNA é a DNA polimerase δ que corrige cerca de 99% dos erros que ocorrem durante a replicação. Porém, alguns nucleótidos mal emparelhados são deixados para trás, originando a criação de mutações. Assim é fundamental a existência de outras proteínas (proteínas MMR) capazes de reconhecer e corrigir estes erros (Jascur e Boland, 2006).

Os membros da família das proteínas MMR humanas responsáveis pelo reconhecimento de emparelhamentos erróneos são designados de hMSH, ou Homólogos de MutS (uma vez que advém da designação utilizada primeiramente em bactérias de MutS). O que diferencia cada uma destas proteínas, descritas na Tabela 1.2, é a forma como interagem no reconhecimento e no tipo de lesão *mismatch*.

Tabela 1.2 – Componentes de MMR e as suas funções (Jiricny, 2006)

COMPLEXOS	PROTEÍNAS	FUNÇÃO
MutS α	hMSH2/hMSH6	Reconhecimento de <i>mismatches</i> base-base e pequenos IDLs
MutS β	hMSH2/hMSH3	Reconhecimento de IDLs
MutL α	hMLH1/hPMS2	Forma complexo ternário com o <i>mismatch</i> e o MutS α ; Recombinação meiótica
MutL β	hMLH1/hPMS1	?
MutL γ	hMLH1/hMLH3	Recombinação meiótica; Apoio ao MutL α na reparação de <i>mismatches</i> base-base e pequenos IDLs

IDLs – Inserção/delecção de *loops*; MLH – MutL homologue; MSH – MutS homologue; PMS – Post-meiotic segregation protein.

A via MMR, ver figura 1.6, é constituída por quatro passos fundamentais: (I) reconhecimento do emparelhamento erróneo – (a) *mismatch* de um único nucleótido; (b) inserção/delecção de *loops* em sequências de microsatélites, (II) recrutamento das enzimas de reparação, (III) excisão da sequência incorrecta e (IV) síntese *de novo* do DNA pela DNA polimerase utilizando a cadeia-mãe como molde (Jascur e Boland, 2006).

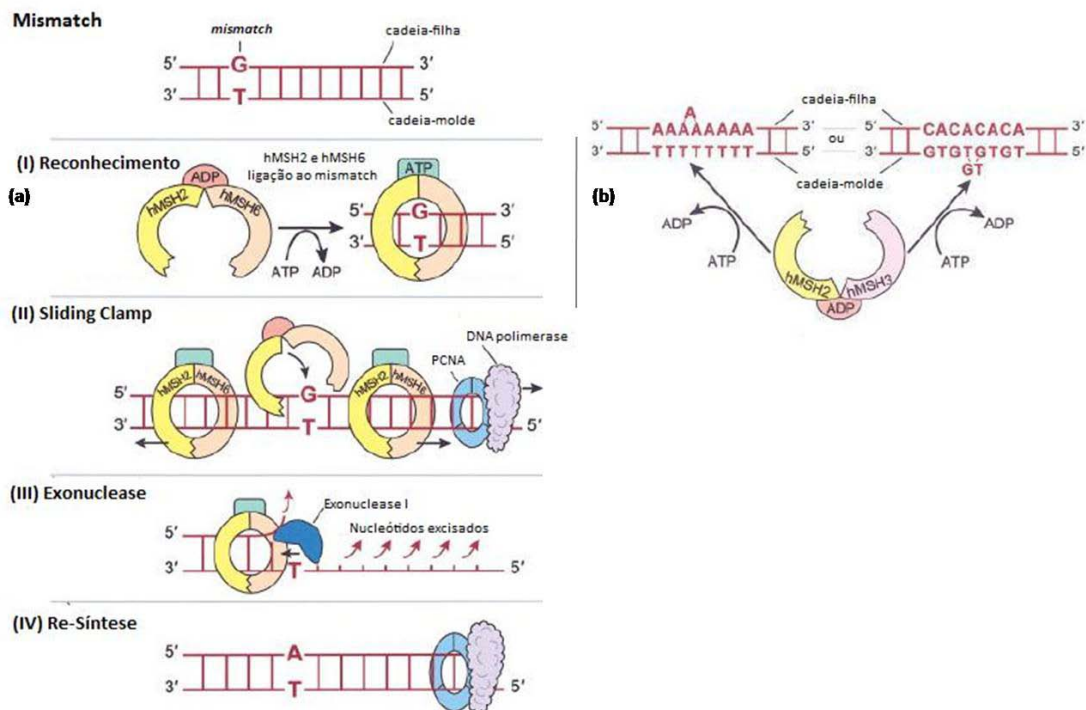


Figura 1.6 - Modelo para a via de reparação de *mismatch* (MMR) (Adaptado de (Jascur e Boland, 2006).

O primeiro passo envolvido na via MMR corresponde ao reconhecimento eficiente de distorções na cadeia de DNA, provocadas pelo incorrecto emparelhamento das bases, bem como por pequenas inserções/delecções que ocorrem, normalmente, em regiões repetitivas. Durante esta fase, a proteína hMSH2 complexa-se com a hMSH6 ou com a hMSH3 para formar os complexos MutS α ou MutS β , respectivamente. Estes dois complexos desempenham um papel importante no reconhecimento do *mismatch* e na iniciação do processo de reparação (Li, 2008). Estas proteínas MSH são ATPases, que contêm um local de ligação do ATP, que é altamente conservado de entre os polipéptidos envolvidos na reparação do DNA (Jiricny, 2006). Quando um *mismatch* é detectado ocorre a desfosforilação de ATP a ADP, através das subunidades dos complexos MutS, o que provoca uma alteração conformacional deste complexo, originando um complexo proteico alterado que rodeia a cadeia de DNA que contém o *mismatch*, permitindo-lhe deslizar livremente pela cadeia (designado de *sliding-clamp*). Ou seja, na presença de ADP o complexo liga-se fortemente ao *mismatch*, enquanto na presença de ATP actua como um *sliding-clam* (Jascur e Boland, 2006; Jiricny, 2006).

De seguida o complexo DNA:MutS:ATP recruta o complexo MutL α (hMLH1/hPMS2), que interage com o MutS no local do *mismatch*, funcionando como promotor molecular, reunindo outras proteínas, como a DNA polimerase, o PCNA (*proliferating cell nuclear antigen*) e a exonuclease I (EXO1) (Jascur e Boland, 2006). O PCNA interage com as proteínas MSH2, MSH3, MSH6 e MLH1 e parece estar envolvido na iniciação das etapas da nova síntese de DNA (Li, 2008). Neste momento, a nova cadeia de DNA sintetizada que contém o *mismatch* terá que ser selectivamente marcada para remoção e nova síntese. A remoção do emparelhamento incorrecto é realizada pela EXO1, tanto no sentido 5' \rightarrow 3' como no sentido 3' \rightarrow 5'. Quando o *mismatch* é, finalmente, removido, a actividade da EXO1 deixa de ser estimulada pelo complexo MutS α e é activamente inibida pelo MutL α . Por fim, a DNA polimerase (DNA polimerase δ e ϵ) preenche a falha inserindo correctamente um novo nucleótido ou a sequência em falta (Houtgraaf *et al.*, 2006; Jiricny, 2006).

Alterações neste sistema de reparação aumentam significativamente as taxas de mutação, acelerando o processo de cancerigénese, nomeadamente um aumento da amplificação genética e de uma redução nos processos de reparação. A perda da

actividade da via MMR, contribui para a iniciação e promoção de processos de cancerigénese, bem como para a instabilidade de sequências repetitivas de DNA, como por exemplo os microsátélites. As proteínas envolvidas nesta via de reparação estão também envolvidas noutras vias de reparação, alternando entre a sinalização da lesão de DNA e processos recombinogénicos e mutagénicos (Jiricny, 2006; Schroering *et al.*, 2007).

1.7.3 REPARAÇÃO POR EXCISÃO DE BASES (BER)

A via de reparação por excisão de bases (BER) é conhecida como a *guardiã* do genoma, e também a mais predominante, contra as lesões derivadas do metabolismo celular, incluindo as resultantes de ROS, metilação, desaminação e hidroxilação (Hoeijmakers, 2001). Esta via reconhece e repara bases modificadas, bem como locais abásicos (APs) e quebras de DNA em cadeia simples (SSBs), maioritariamente produzidas por ROS (Maynard *et al.*, 2009).

As espécies reactivas de oxigénio produzidas durante o metabolismo celular podem reagir aleatoriamente com lípidos, proteínas e ácidos nucleicos originando stress oxidativo e lesões nessas macromoléculas. A grande consequência do stress oxidativo são, de facto, as lesões de DNA responsáveis por grande instabilidade genómica. As bases do DNA são muito sensíveis à oxidação por ROS, especialmente a guanina devido ao seu baixo potencial redox, pelo que uma das lesões melhor caracterizada gerada por ROS é a 8-hidroxi guanina (8-oxoG), lesão altamente mutagénica que origina mutações por transversão (Maynard *et al.*, 2009).

A via BER é uma via crítica na manutenção da integridade genómica (ver figura 1.7), e apresenta duas sub-vias dependendo da lesão e das enzimas envolvidas nessa resposta: via curta e mais frequente (*Short-patch* BER) e via longa, normalmente para reparar uma lesão entre 2 a 6 nucleótidos (*Long-patch* BER) (Hoeijmakers, 2001; Maynard *et al.*, 2009; Sancar *et al.*, 2004).

O passo inicial da via de reparação por excisão de bases é o reconhecimento da lesão que é efectuado por DNA glicosilases. Estas enzimas provocam a quebra da ligação N-glicosil entre o açúcar e a base azotada, libertando a base lesada o que provoca um local abásico na cadeia de DNA. De acordo com o tipo de lesão a ser

reparada é recrutada uma glicosilase específica para essa lesão. Existem DNA glicosilases que reconhecem bases oxidadas/reduzidas, bases alquiladas (normalmente, metiladas), bases desaminadas e bases *mismatch* (Maynard *et al.*, 2009; Sancar *et al.*, 2004). A sua classificação é baseada nos seus mecanismos de reacção, algumas actuam como simples glicosilases catalizando a reacção de remoção da base formando um local AP o que vai implicar a acção de uma outra enzima com função liase (normalmente APE1) responsável pela incisão da cadeia, enquanto outras desempenham a dupla função de glicosilase e de AP liase. Quando a base lesada é removida por uma glicosilase/AP liase, responsáveis pela quebra da ligação 3'-fosfodiéster do local abásico, a endonuclease APE1 quebra a ligação 5' e recruta a DNA Polimerase β (Pol β) que vai preencher o *gap* formado, sendo a ligação da cadeia efectuada pelo complexo Ligase 3/XRCC1. Quando o local AP é gerado por hidrólise espontânea, o grupo 5'-fosfato fica modificado e torna-se resistente à eliminação pela Pol β , prosseguindo a reparação normalmente pela via *long-patch*. Após reconhecimento da lesão é recrutado um complexo PCNA/DNA polimerase δ/ϵ (Pol δ/ϵ) responsável pela adição de mais nucleótidos à extremidade 3' do local AP formando um *loop* que é removido por uma endonuclease (FEN1), seguindo-se posteriormente a ligação das extremidades do DNA pela DNA ligase I (Dianov *et al.*, 2003; Zhang e Dianov, 2005).

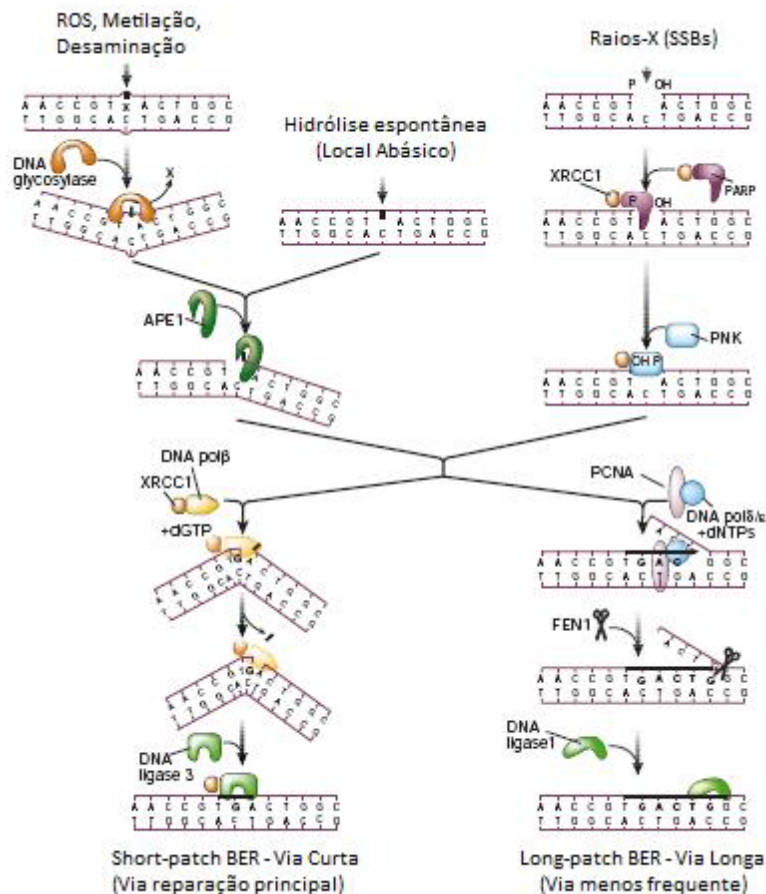


Figura 1.7 - Via de Reparação por Excisão de Bases (BER) (Adaptado de (Hoeijmakers, 2001).

Para além das enzimas referidas como participantes nesta via de reparação, outros factores proteicos têm sido identificados como moduladores da actividade da via BER. Estas proteínas auxiliares podem interagir com as proteínas principais da via e/ou com o DNA de forma a aumentar a actividade enzimática ou a eficiência das reacções (Fan e Wilson III, 2005; Maynard *et al.*, 2009). Do vasto conjunto de proteínas salienta-se a acção da p53, proteína supressora de tumor, que se presume que facilite a actividade de ligação da Polβ ao DNA; até mesmo do complexo MSH2/MSH6 de reconhecimento de *mismatch* que parece promover a reparação da base lesada, e que deverá ser reparada pela via BER, entre muitas outras (Fan e Wilson III, 2005).

Doenças genéticas provocadas por mutações em genes envolvidos na via BER, conhece-se uma mutação que afecta a linha germinal do gene MUTYH (MYH) associada com a predisposição para múltiplos adenomas colorectais, sendo conhecida

como polipose MYH, está descrita como uma doença autossómica recessiva (Al-Tassan *et al.*, 2002; Isidro *et al.*, 2004; Wilson III e Bohr, 2007). Não são conhecidas outras patologias, pelo que se especula que deficiências nesta via sejam incompatíveis com a vida (Maynard *et al.*, 2009; Mohrenweiser *et al.*, 2003).

1.7.4 REPARAÇÃO POR EXCISÃO DE NUCLEÓTIDOS (NER)

A reparação por excisão de nucleótidos é um sistema de reparação com capacidade de reparar uma ampla classe de lesões, nas quais se incluem, distorções na hélice de DNA como as provocadas por foto-produtos derivados de exposição à radiação UV e aductos químicos de elevado peso molecular, por exemplo, aductos de cisplatina. É uma via que pode detectar as lesões de duas formas, ou seja, apresenta duas sub-vias: (1) reparação global (do inglês *Global-Genome NER*: GG-NER), que detecta lesões associadas a distorções na hélice ao longo do genoma; e (2) reparação associada à transcrição (do inglês *Transcription- Coupled NER*: TC-NER) específica para lesões que bloqueiem a maquinaria associada ao desenvolvimento da transcrição, nomeadamente no bloqueio da RNA polimerase. Ambos os mecanismos detectam lesões, embora de formas diferentes, utilizando a mesma abordagem de reparação, que consiste em: abrir a dupla hélice, excisar 22-30 bases contendo a lesão, preencher o *gap* criado utilizando como molde a cadeia complementar e promover a ligação das extremidades do novo fragmento sintetizado (Garinis *et al.*, 2008). É uma via que envolve a acção concertada de cerca de 25 proteínas.

Na figura 1.8 estão esquematizadas as duas sub-vias associadas com a reparação NER. O que distingue estas duas sub-vias são os complexos de reconhecimento da lesão. A reparação global (GG-NER) é responsável por detectar e corrigir lesões ao longo de todo o genoma, o complexo proteico XPC-hHR23B é o responsável pela detecção da lesão, enquanto na sub-via TC-NER esta função de reconhecimento é desempenhada pelos factores CSA e CSB que são recrutadas quando ocorre o bloqueamento da RNA polimerase II, todos os outros passos serão idênticos. Após reconhecimento estão criadas as condições que favorecem o recrutamento de outros factores de reparação, nomeadamente de algumas helicases (XPB e XPD), que são sub-unidades do complexo factor de transcrição basal TFIIH. Ocorre o desenrolamento da estrutura de DNA em torno da lesão. A cadeia simples é

estabilizada devido à acção conjunta de duas proteínas, XPA e da proteína de replicação A (RPA), sendo que esta última confere especificidade à cadeia ao funcionar como sinalizadora da cadeia lesada inibindo qualquer incisão na cadeia intacta. Ocorre a ligação da proteína XPG e do complexo ERCC1-XPF, responsáveis pela quebra da cadeia nas extremidades 3' e 5' que rodeiam a lesão. A descontinuidade formada após excisão da cadeia lesada é sintetizada pela polimerase δ/ϵ e pelos factores de replicação e finalmente a união do fragmento é concluído pela acção de uma DNA ligase (Friedberg, 2001; Hoeijmakers, 2001; van Brabant *et al.*, 2000).

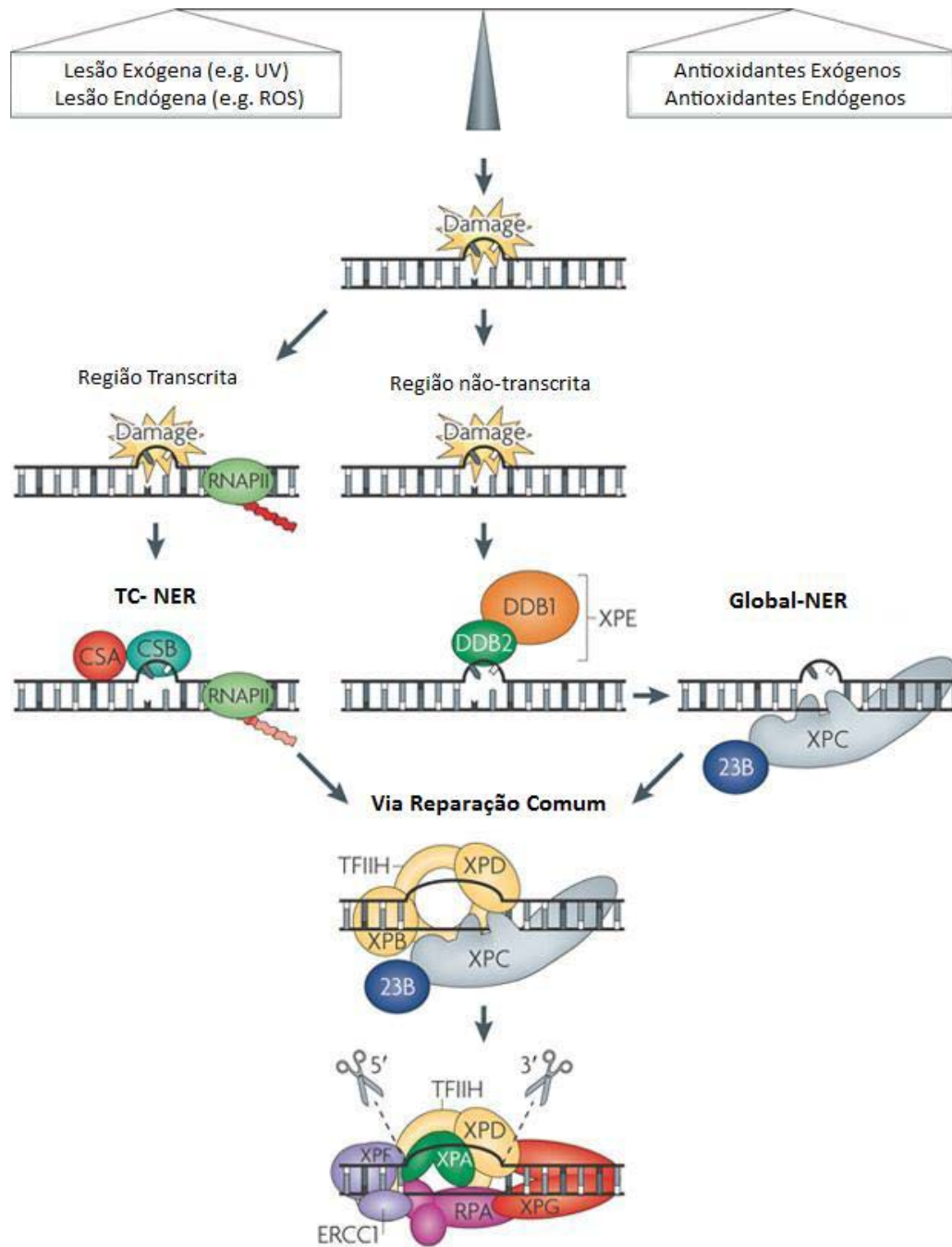


Figura 1.8 - Mecanismo de reparação por excisão de nucleótidos (NER). São representadas as duas sub-vias que caracterizam a via NER (Global-NER e TC-NER) (Adaptado de (Cleaver *et al.*, 2009).

Pelo menos três doenças hereditárias, estão relacionadas com alterações nos genes que codificam para as proteínas da via NER: XP (*Xeroderma Pigmentosum*), caracterizada por apresentar um risco 1000x superior no desenvolvimento do cancro de pele (mutações em sete genes *XPA-XPG*); síndrome de Cockayne (mutações genes *CSA* e *CSB*) e Tricotodistrofia (TTD) (Houtgraaf *et al.*, 2006; Mohrenweiser *et al.*, 2003).

1.7.5 REPARAÇÃO DE QUEBRAS EM CADEIA DUPLA (DSB)

Consideradas como as lesões mais críticas do genoma, as quebras em cadeia dupla, podem surgir por erros espontâneos que ocorram durante o processo de replicação ou devido à exposição directa a agentes lesivos, tais como a radiação ionizante, ROS e agentes quimioterapêuticos. Alguns estudos indicam que este tipo de lesão é potencialmente citotóxica e mutagénica para as células (Houtgraaf *et al.*, 2006; Mohrenweiser *et al.*, 2003). Estas quebras de cadeia dupla poderão dar origem a fragmentação cromossomal, translocações e deleções que se não forem atempadamente detectadas e reparadas podem ser responsáveis por aumentar a instabilidade genómica celular, sendo crucial a eficácia do processo de reparação (Mohrenweiser *et al.*, 2003).

Existem duas vias associadas à reparação de quebras em cadeia dupla: (1) reparação por recombinação homóloga (HRR), que ocorre predominantemente nas fases S e G₂ do ciclo celular e utiliza a sequência complementar da cromátide intacta para reparar a lesão que se encontra na cromátide que lhe é homóloga; (2) reparação não homóloga (NHEJ), não necessita de qualquer homologia já que envolve junção directa das extremidades (McKinnon e Caldecott, 2007; Mohrenweiser *et al.*, 2003) e normalmente ocorre durante as fases G₀ e G₁ do ciclo celular (Houtgraaf *et al.*, 2006).

O reconhecimento da lesão DSB é função desempenhada pelo complexo proteico MRE11-RAD50-NBS1 (MRN), com actividade de helicase e de endonuclease. Após reconhecimento, ocorre a activação e o recrutamento de proteínas de sinalização da lesão, nomeadamente as proteínas ATM, a subunidade catalítica de DNA-PKs (DNA-PK_{CS}) e uma histona fosforilada (H2AX). Seguidamente são recrutadas proteínas envolvidas na reparação da lesão, de acordo com figura 1.9 (Bristow e Hill, 2008).

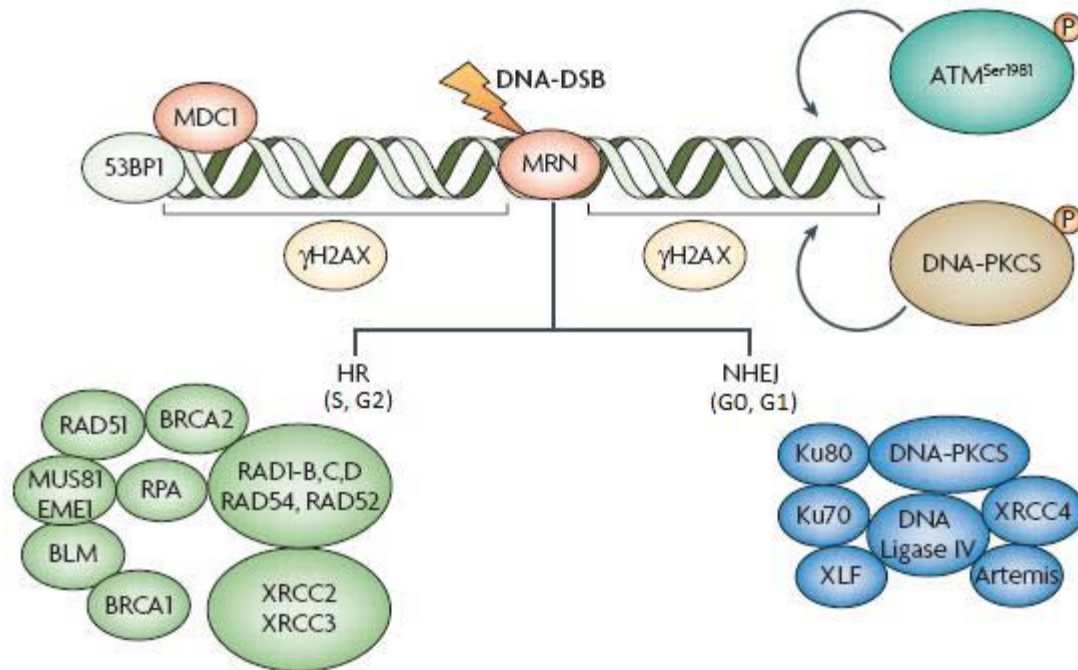


Figura 1.9 - Vias de reparação associadas com reparação de lesões de quebras em cadeia dupla (Adaptado de (Bristow e Hill, 2008))

Para que o processo de reparação NHEJ ocorra as proteínas KU70 e KU80 formam um heterodímero que se liga às extremidades da DSB e recruta a DNA-PK_{CS}. Este complexo por sua vez é responsável pela activação de uma outra proteína (Artemis, com actividade de exonuclease) cuja função é o processamento das extremidades de DNA, antes das proteínas XRCC4 e DNA Ligase IV finalizarem a ligação das cadeias. Os resíduos perdidos não são restaurados pelo que este tipo de reparação acarreta erros que podem ser mutagénicos (Lord *et al.*, 2006).

O processo de reparação HR é mais complexo do que a reparação NHEJ. Num primeiro passo ambas as extremidades 5' da quebra de dupla cadeia são eliminadas pela acção de uma nucleasa específica de forma a gerar extremidades 3' de DNA em cadeia simples (ssDNA). Subsequentemente, uma das extremidades 3'-ssDNA invade a cadeia dupla homóloga intacta e gera uma estrutura em *D-loop*. Para promover a invasão de sequências homólogas por parte das cadeias quebradas, a actividade de exonuclease 5' - 3' do complexo RAD50/MRE11/NBS1 expõe ambas as extremidades 3', para que a proteína RPA facilite a montagem de um filamento nucleoproteico de RAD51 constituído por conjunto de parálogos de RAD51 (XRCC2, XRCC3, RAD51B, C e

D). As proteínas RAD52 e RAD54 facilitam a montagem do filamento, enquanto a proteína RAD52 promove a ligação dos segmentos de DNA complementar, promovendo a síntese *de novo* utilizando a cadeia homóloga como molde (Khanna e Jackson, 2001; Mohrenweiser *et al.*, 2003). Este processo, de invasão da cadeia homóloga leva à formação de junções de *Holliday* resultantes do *crossover* de DNA. Durante o último passo da recombinação homóloga, as junções de *Holliday* são clivadas por uma resolvase dando origem a duas moléculas de DNA intactas (Khanna e Jackson, 2001). Esta via de reparação é uma via livre de erros pelo simples facto de utilizar a cromátide homóloga intacta ocorrendo reparação fidedigna da lesão.

CAPÍTULO 2

OBJECTIVOS

2. OBJECTIVOS

O cancro da mama é a patologia oncológica mais frequente nas mulheres, sendo a primeira causa de morte por cancro no sexo feminino. Mesmo sendo uma patologia com uma incidência elevada, as causas exactas que contribuem para o desenvolvimento desta neoplasia permanecem desconhecidas. Uma pequena percentagem (5 a 10%) deste tipo de cancro está relacionada com a vertente hereditária da doença, maioritariamente caracterizada por mutações nos genes *BRCA1* e *BRCA2*. Os restantes 90 a 95% são casos esporádicos e ocorrem na ausência de mutações nestes genes críticos. Alguns dos potenciais factores de risco associados com a etiologia do cancro da mama estão descritos e neles se incluem: factores associados com história reprodutiva e exposição hormonal como menarca precoce/menopausa tardia, idade tardia para a primeira gravidez e nuliparidade, e também a idade de diagnóstico. Existem ainda factores ambientais/ocupacionais (exposição a radiação ionizante e cancerígenos químicos) e outros relacionados com o estilo de vida (dieta, consumo de álcool e tabaco, actividade física, obesidade) igualmente relacionados com um aumento de risco para cancro da mama.

O risco associado ao desenvolvimento do cancro da mama, poderá estar relacionado com a capacidade individual para destoxificar metabolitos cancerígenos, provenientes quer do metabolismo endógeno quer da exposição ambiental, e/ou reparar lesões no DNA induzidas directa ou indirectamente pelos factores de risco referidos. Assim, a susceptibilidade para cancro da mama pode ser mediada, em parte, pela variabilidade genética de genes envolvidos nestas vias. A grande maioria das proteínas descritas como estando envolvidas nas diferentes vias quer de reparação de lesão genética quer de metabolização/destoxificação, são enzimas polimórficas, pelo que alterações nos genes que as codificam podem comprometer o seu desempenho, contribuindo assim para alterações na capacidade de reparação e destoxificação, e concomitantemente criar instabilidade genómica às células.

A utilização de estudos populacionais de associação, conhecidos como estudos caso-controlo, através do estudo de genes candidatos, permite estudar a potencial associação dos genes polimórficos a uma determinada doença numa população. No

presente trabalho foram efectuados estudos caso-controlo na população Portuguesa em que se escolheram genes candidatos que codificam para proteínas com potencial envolvimento na cancerigénese mamária. Esta dissertação está dividida em 10 capítulos, sendo cada um destes capítulos dedicado a uma via mecanicista específica e a sua associação com a susceptibilidade individual para cancro da mama.

Endogenamente são produzidos metabolitos com capacidade de actuarem como agentes cancerígenos, nomeadamente as espécies reactivas de oxigénio, que podem ser formadas por diferentes vias, ou exogenamente a exposição a compostos genotóxicos. O estudo destes agentes na cancerigénese não é novo e muito trabalho tem sido desenvolvido, inclusive na cancerigénese mamária. No entanto, e mesmo neste âmbito há ainda algum estudo para desenvolver.

Diversos factores alteram a exposição do organismo a hormonas endógenas. Deste modo, polimorfismos em genes envolvidos no metabolismo e mecanismos de acção de hormonas sexuais são fortes candidatos a estarem associados com susceptibilidade individual na cancerigénese mamária.

A reparação de DNA é um processo ubíquo em todos os sistemas vivos. A sua universalidade reflecte a constante pressão que conduz à alteração da integridade do genoma, resultante da instabilidade intrínseca do material genético e das limitações na fidelidade da replicação do DNA. Vários agentes induzem diferentes tipos de lesões no DNA e de acordo com a lesão que vão criar activam um conjunto de proteínas responsáveis pela eficiente reparação da lesão. A via de reparação por excisão de bases (BER) actua em lesões espontâneas ou induzidas por agentes lesivos principalmente os que advém do metabolismo celular, de que são exemplo as espécies reactivas de oxigénio (ROS). A via de reparação por excisão de nucleótidos (NER) repara um outro tipo de lesões, como por exemplo as criadas pela excessiva exposição à radiação UV ou aductos de elevado peso molecular. Também os erros de replicação originados pelo incorrecto emparelhamento provocado pela DNA polimerase, têm um mecanismo de reparação específico conhecido como reparação *mismatch* (MMR). O tipo de lesão conhecido como sendo mais agressivo para o genoma é, a quebra da dupla cadeia. Um reconhecido factor de risco para cancro da mama é a exposição à radiação ionizante, agente este que é responsável pela indução de quebras na dupla cadeia.

O desenvolvimento de estudos de associação requer um dispêndio económico-laboral significativo, pelo que o aperfeiçoamento de uma técnica que permita comparar as diferentes populações envolvidas nos estudos de associação de forma rápida e eficaz pode ser uma mais-valia para os estudos epidemiológicos. A técnica conhecida como *Pools* de DNA permite criar uma única amostra na qual se incluem, aleatoriamente, indivíduos da mesma população. Estes *Pools* populacionais são depois utilizados para comparar as frequências alélicas de diferentes SNPs nas diferentes populações permitindo perceber se aqueles SNPs poderão estar envolvidos na neoplasia mamária, neste caso, funcionando como pré-selecção dos SNPs mais relevantes.

As conclusões que foram sendo recolhidas no decorrer dos vários capítulos deste trabalho encontram-se resumidas no final da dissertação e serão apresentadas como “Considerações Finais” do conjunto de dados obtidos.

CAPÍTULO 3

**O PAPEL DOS POLIMORFISMOS DO GENE GSTA2 E
HAPLÓTIPOS NA SUSCEPTIBILIDADE INDIVIDUAL PARA
CANCRO DA MAMA: ESTUDO CASO-CONTROLO NA
POPULAÇÃO PORTUGUESA**

3. O PAPEL DE POLIMORFISMOS DO GENE *GSTA2* E HAPLÓTIPOS NA SUSCEPTIBILIDADE INDIVIDUAL PARA CANCRO DA MAMA: ESTUDO CASO-CONTROLO NA POPULAÇÃO PORTUGUESA

Publicado em Silva *et al.*, 2009, *Oncology Reports*, 22, 593-598

3.1 INTRODUÇÃO

A maioria dos factores de risco, especialmente associados com a vertente esporádica da neoplasia mamária, aponta para uma associação com exposição prolongada a níveis elevados de estrogénios (Cheng *et al.*, 2005; Silva *et al.*, 2006b). Diversos estudos epidemiológicos têm sido publicados com referência ao facto dos estrogénios em circulação poderem contribuir para um aumento de risco de desenvolvimento de cancro da mama. Para além disso, durante o metabolismo dos estrogénios podem ser produzidos alguns metabolitos cancerígenos, tais como os catecol-estrogénios (Crooke *et al.*, 2006; Torresan *et al.*, 2008; Yager, 2000). Estes metabolitos podem ser responsáveis por um aumento da produção de ROS, provocando uma disfunção da homeostasia celular redox (Crooke *et al.*, 2006; Yager, 2000), e lesões no DNA. É postulado que as consequências genotóxicas provocadas pelo metabolismo oxidativo dos estrogénios sejam mitigadas por reacções de conjugação dos metabolitos com enzimas de fase II (Crooke *et al.*, 2006). Diversas enzimas estão envolvidas no metabolismo oxidativo e de conjugação de estrogénios (Yager, 2000), sugerindo o potencial envolvimento de formas polimórficas dos seus genes na susceptibilidade para cancro da mama.

Uma das famílias enzimáticas envolvidas no metabolismo de fase II (McIlwain *et al.*, 2006; Torresan *et al.*, 2008; Unlu *et al.*, 2008; Yager, 2000) é a super-família das Glutathione S-Transferase (GSTs). É uma família altamente polimórfica, responsável pela destoxificação de cancerígenos por conjugação de diversos metabolitos reactivos com glutathione reduzido (GSH), incluindo derivados do stress oxidativo (Hayes *et al.*, 2005) e metabolitos gerados através do metabolismo oxidativo de estrogénios (Crooke *et al.*, 2006; Torresan *et al.*, 2008). Alguns destes metabolitos incluem estrogénio quinonas

que são submetidos a conjugação com GSH através da acção catalítica da GSTP1. A formação de conjugados de GSH-estrogénios reduz o nível de quinonas de estrogénio diminuindo assim o potencial de criar lesões no DNA (Crooke *et al.*, 2006).

A classe *Alpha* GSTs (α GSTs) é altamente expressa no fígado, rins e glândula supra-renal, e tem sido descrita como uma das classes de GSTs mais versátil, uma vez que é responsável pela conjugação com GSH de compostos como a bilirubina, ácidos biliares e penicilina, hormonas esteróides e tiróides, permitindo a sua solubilização e armazenamento no fígado (Tetlow e Board, 2004). Adicionalmente, a conjugação de medicamentos alquilantes anti-cancerígenos do grupo mostarda nitrogenada, de algumas aminas heterocíclicas e aldeídos α,β -insaturados também conduz à detoxificação (Tetlow *et al.*, 2001; Tetlow e Board, 2004). As enzimas GSTA1 e GSTA2 são as *Alpha* GSTs mais expressas pelo fígado (Ahn *et al.*, 2006; Coles e Kadlubar, 2005), sendo também expressas pelo tecido mamário (Ahn *et al.*, 2006). Existem alguns estudos epidemiológicos, embora poucos, em que tenham sido estudados a função de polimorfismos em genes da classe *Alpha* GST na susceptibilidade para o cancro (Gemignani *et al.*, 2007; Ning *et al.*, 2004), especialmente para cancro da mama (Ahn *et al.*, 2006). Estando a classe *Alpha* GST envolvida numa ampla variedade de funções, nas quais se incluem biossíntese de esteróides e funções de protecção contra agentes alquilantes, variações polimórficas nos genes desta classe poderão ser responsáveis por consequências fisiológicas que poderão alterar a susceptibilidade para a doença e até mesmo para a resposta a fármacos (Tetlow e Board, 2004). Dois dos membros desta classe, os genes *GSTA1* e *GSTA2* (normalmente co-expressos), catalisam a conjugação com GSH de uma grande variedade de electrófilos, possuem actividade de isomerase esteróide e actividade de peroxidase dependente de glutationo (Coles e Kadlubar, 2005). Pensa-se que o gene *GSTA2* represente a maior linha de defesa contra o stress oxidativo (Tetlow *et al.*, 2004).

Até à data, foram efectuados alguns estudos epidemiológicos, baseados na susceptibilidade para cancro da mama, considerando diversas famílias polimórficas de GSTs, sendo os genes mais frequentemente estudados os *GSTT1*, *GSTP1* e *GSTM1* (Lee *et al.*, 2008; Maugard *et al.*, 1998; Saxena *et al.*, 2008; Steck *et al.*, 2007; Syamala *et al.*, 2008; Torresan *et al.*, 2008; Vogl *et al.*, 2004; Zheng *et al.*, 2003; Zheng *et al.*, 2002). A frequência de estudo de genes pertencentes a outras classes é muito menor,

encontrando-se uma escassez de resultados quando se limita a pesquisa à influência de genes da classe *Alpha* na susceptibilidade individual para cancro da mama. No entanto, e desde que o gene *GSTA2* está envolvido na destoxificação de metabolitos associados com aumento de risco para cancro da mama, estudou-se o envolvimento de variantes polimórficas deste gene na susceptibilidade individual para cancro da mama.

3.2 MATERIAIS E MÉTODOS

3.2.1 DESCRIÇÃO DA POPULAÇÃO

Este estudo incluiu 291 mulheres Caucasianas com cancro da mama, recrutadas no Hospital São Francisco Xavier (Departamento de Medicina Laboratorial), sem historial prévio de doenças neoplásicas, sem patologia da tiróide, e que não tenham realizado transfusões sanguíneas. O diagnóstico histológico foi confirmado para todos os casos. Por sua vez, a população controlo é constituída por 547 indivíduos Caucasianos do sexo feminino, sem qualquer registo de patologia neoplásica prévia ou doença maligna concomitante, tendo sido recrutada na mesma unidade hospitalar que a população de doentes. A todos os indivíduos incluídos no estudo foi facultado o objectivo do estudo, tendo sido obtido o seu consentimento por escrito e garantido o anonimato de todos os que constituem cada uma das populações. Informações sobre característica demográficas, antecedentes familiares de cancro, estilo de vida (hábitos tabágicos e alcoólicos) foram registadas num inquérito epidemiológico administrado por pessoal especializado. Foram considerados ex-fumadores aqueles que deixaram de fumar dois anos antes do diagnóstico de cancro ou, para controlos dois anos antes da data de inclusão no estudo. A taxa de resposta à participação foi superior a 95% em ambas as populações.

3.2.2 EXTRACÇÃO DE DNA

As amostras de sangue de todos os indivíduos foram recolhidas em tubos heparinizados de 10 ml e armazenadas a -20°C até posterior utilização. O DNA

genómico foi obtido a partir de 250 µl de sangue total utilizando um kit comercial tendo sido o método de extração utilizado o indicado pelo fabricante (QIAamp DNA mini kit; Qiagen, Hilden, Germany). Cada amostra de DNA foi então devidamente identificada e guardada a -20°C até ao seu processamento.

3.2.3 GENOTIPAGEM DE POLIMORFISMOS DO GENE *GSTA2*: P110S, S112T E E210A

A genotipagem do polimorfismo E210A foi efectuada pela técnica de reacção em cadeia da polimerase (PCR, do inglês *Polymerase Chain Reaction*) seguida da análise de fragmentos de restrição pela técnica RFLP (do inglês *Restriction Fragment Length Polymorphism*). Os *primers* (sequências iniciadoras a montante e a jusante da sequência que se pretende amplificar) utilizados na análise deste SNP estão apresentados na Tabela 3.1, e foram especificamente desenhados para a amplificação de um fragmento de 163 pb. A amplificação do fragmento pretendido realizou-se num volume final de 50 µl contendo 100-150 ng de DNA genómico, 1 µM de cada *primer*, 0,75 U de *Immolase* (Bioline), 2,5 mM de MgCl₂ e 0,8 mM de dNTPs, em tampão específico da polimerase à concentração final de 1,3x. As condições de PCR aplicadas consistiram num período inicial de activação de 7 minutos a 95 °C, 35 ciclos de amplificação incluindo desnaturação a 94 °C durante 30 segundos, emparelhamento a 62 °C durante 30 segundos e extensão a 72 °C durante 30 segundos, e um período de extensão final de 10 minutos a 72 °C. De forma a confirmar a amplificação do segmento de 163 pb, os produtos de PCR foram submetidos a electroforese em gel de agarose.

Tabela 3.1 - Alguns detalhes da técnica PCR-RFLP aplicados ao polimorfismo *GSTA2* E210A

<i>GSTA2</i>	SEQUÊNCIA DOS <i>PRIMERS</i>	EFEITO DA ENZIMA DE RESTRIÇÃO	PERFIS APÓS ANÁLISE DE RESTRIÇÃO
E210A	Forward: 5'-CCA AGG AAG CCT CCC ATG CAT G-3'	A → C, origina um local de reconhecimento para <i>PaeI</i>	AA: 163bp; AC: 163, 140, 23bp; CC: 140, 23bp
	Reverse : 5'-GCT TCA CAA CAG GCA CAA TCA ACA C-3'		

Este polimorfismo origina ganho ou perda de locais de restrição, os quais permitirão discriminar por RFLP, após digestão com enzima de restrição, os alelos comum e variante que caracterizam este SNP. Deste modo, os fragmentos amplificados (10 µl) foram digeridos com a enzima de restrição *PaeI* (Fermentas) durante 2 horas, em reacções de hidrólise de 12,5 µl de volume final. Os perfis de restrição característicos estão apresentados na Tabela 3.1. As experiências foram realizadas em duplicado em ensaios independentes, sendo os resultados inconclusivos reanalisados.

Os polimorfismos P110S e S112T foram genotipados por PCR em tempo real (AB7300) utilizando os ensaios da Applied Biosystems *TaqMan*[®] *SNP Genotyping Assays* (as referências dos ensaios são C_12027714_50 e C_22275149_30, respectivamente para cada SNP) segundo instruções fornecidas pelo fabricante. Para uma eficiente discriminação alélica as amostras de DNA foram quantificadas utilizando o Reagente de Quantificação Quant-iT[™] PicoGreen dsDNA Assay (Molecular Probes, Invitrogen) de acordo com as instruções fornecidas pelo fabricante.

3.2.4 ANÁLISE ESTATÍSTICA

A análise das frequências de Hardy-Weinberg dos alelos de todos os SNPs estudados, para a população de doentes e para os controlos, foi realizada através de testes de probabilidade exacta disponíveis no software Mendel (versão 8.0.1) (Lange *et al.*, 2001).

O grau de desequilíbrio de *linkage* entre os diferentes SNPs expresso como D' , e a estimativa de haplótipo para a combinação dos diferentes SNPs estudados foram calculados utilizando o software SNPStats (Sole *et al.*, 2006).

Com o intuito de verificar a normalidade das variáveis contínuas (idade) utilizou-se o Teste de Kolmogorov-Smirnov e para a análise da homogeneidade das variâncias populacionais foi utilizado o Teste de Levene. Estes testes de hipóteses são utilizados para averiguar qual o modelo estatístico de onde são provenientes os dados, ou seja, para averiguar se as amostras foram retiradas de populações com distribuição Normal. Uma vez verificados os pressupostos da normalidade das variáveis e da

homogeneidade das variâncias utilizou-se o teste t-*Student* para a análise de médias da variável idade.

De modo a aferir a relação entre a susceptibilidade para cancro da mama e as variantes (SNPs) estudadas foram realizados múltiplos cálculos estatísticos de modo a testar a hipótese nula (H_0). A H_0 corresponde à não-associação dos SNPs com a susceptibilidade para cancro da mama. A hipótese alternativa (H_1) corresponde à associação dos SNPs com a susceptibilidade. Ou seja, para um nível de significância de 0,05 (intervalo de confiança a 95% - 95% IC), se o valor de $p < 0,05$ rejeitamos H_0 , enquanto que para um valor de $p > 0,05$ não existem razões para rejeitar H_0 . Assim, apenas os valores de p inferiores a 0.05 foram considerados como significativos.

A análise estatística das frequências genóticas da distribuição de classes referentes a variáveis não contínuas entre as duas populações foi realizada recorrendo ao teste de χ^2 . O valor de OR (factor de risco) e respectivo intervalo de confiança a 95% (95% IC), associados a cada um dos diferentes genótipos foi determinado por regressão logística, tendo este valor sido igualmente ajustado para potenciais factores confundentes: hábitos tabágicos (fumadores e não fumadores) alcoólicos (nunca, social e regular) e idade (≤ 30 , 31-49, 50-69 e ≥ 70 anos). Os valores de OR foram calculados usando o genótipo dos indivíduos homocigóticos mais comuns como referência. Todos os cálculos foram realizados no programa SPSS (versão 15.0, SPSS Inc Chicago, IL).

3.3 RESULTADOS

Foram incluídos neste estudo caso-controlo, para a população Portuguesa, 291 doentes com cancro da mama e 547 controlos saudáveis. Na Tabela 3.2 estão apresentadas as características gerais tais como as classes de idade, hábitos tabágicos e hábitos alcoólicos para cada uma das populações. Após análise, pelo teste do χ^2 , das variáveis quantitativa contínua (classe de idades) e qualitativas (hábitos tabágicos e hábitos alcoólicos) verificou-se que não existem diferenças significativas ($p > 0,05$) para a distribuição das idades e hábitos tabágicos entre casos e controlos, enquanto que para o consumo de álcool existem diferenças significativas ($p < 0,001$) entre as duas

populações, sendo mais prevalentes os consumidores de bebidas alcoólicas na população de casos do que na população controlo.

Foram estudados 3 SNPs, todos não-sinónimos, presentes no gene *GSTA2* (P110S; S112T e E210A). As frequências de Hardy-Weinberg dos alelos de todos os SNPs estudados foram calculadas, tendo-se verificado, que a distribuição dos génotipos para todos os polimorfismos estão em acordo com o equilíbrio de Hardy-Weinberg ($p > 0,05$). A Tabela 3.3 apresenta as frequências do alelo menos comum (MAF, do inglês *Minor Allele Frequency*) e as frequências genóticas determinadas para ambas as populações. Os valores de MAF obtidos estão em concordância com os descritos no projecto HapMap³. Os resultados disponíveis na literatura que se refiram a populações Caucásicas, e as frequências genóticas para o polimorfismo

Tabela 3.2 – Características gerais do grupo de doentes (n=291) e do grupo de controlos (n=547) estudados.

CARACTERÍSTICAS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P *
Idade^{a, b}			
≤ 30	1 (0,3)	2 (0,4)	0,921 ^c
31 – 49	68 (23,4)	120 (21,9)	
50 – 69	161 (55,3)	300 (54,8)	
≥ 70	61 (21,0)	125 (22,9)	
Hábitos Tabágicos			
Nunca ou Ex-F.	254 (87,9)	490 (91,6)	0,087 ^c
Fumadores	35 (12,1)	45 (8,4)	
N.D.	2	12	
Hábitos Alcoólicos			
Nunca	221 (76,2)	441 (82,6)	< 0,0001 ^c
Social	25 (8,6)	59 (11,0)	
Regular	44 (15,2)	34 (6,4)	
N.D.	1	13	

^a Idade de diagnóstico de doentes. ^b Idade da população controlo na altura do diagnóstico do doente correspondente. ^c Grupo casos *versus* Grupo controlos. * Ver Materiais e Métodos. N.D. – Não Determinado

³ <http://www.hapmap.org/>

S112T do gene *GSTA2* nessas mesmas populações revelaram-se ambíguos. Dois dos estudos são concordantes com os resultados obtidos no presente estudo (Gemignani *et al.*, 2007; Landi *et al.*, 2005), no entanto existem outros dois estudos que reportaram frequências genotípicas díspares (Landi *et al.*, 2007; Ning *et al.*, 2004). As frequências genotípicas obtidas em ambas as populações, doentes e controlos, e os valores de OR não-ajustados e ajustados para os diferentes factores confundentes associados a cada polimorfismo, não são significativamente diferentes (Tabela 3.3), o que sugere que os polimorfismos estudados não estão associados com susceptibilidade individual para cancro da mama. Após estratificação para os diferentes factores relacionados com cancro da mama, que poderiam ser relevantes como potenciais factores de risco (*e.g.* amamentação, idade de menopausa, idade de menarca, paridade, idade da primeira gravidez), foram determinados os valores de OR associados aos diferentes SNPs estudados, não se tendo encontrado, igualmente, nenhuma potencial associação com a doença (dados não apresentados). A análise de desequilíbrio de *linkage* foi efectuada através do software SNPStats, o qual revelou que todos os SNPs incluídos no estudo se encontram em desequilíbrio de *linkage*, valores dados por D' como apresentado na Tabela 3.4.

Da combinação dos diferentes alelos gerados pelos SNPs estudados do gene *GSTA2* poderão ser potencialmente gerados oito haplótipos diferentes, no entanto, e de acordo com os resultados obtidos neste estudo, só foi possível encontrar seis haplótipos diferentes (Tabela 3.5), quatro deles foram previamente descritos por Tetlow *et al.* em diferentes populações (Australiana, Bantu, Africana Criola e Chinesa (Tetlow *et al.*, 2001; Tetlow e Board, 2004). A partir dos diferentes haplótipos gerados pelos polimorfismos estudados, dois deles (P110; S112; E210 e P110; T112; E210) (ver Tabela 3.5) apresentam uma frequência cumulativa superior a 90% nas populações estudadas, apresentando cada um uma frequência similar em ambas as populações. Consequentemente, os resultados obtidos considerando o papel deste haplótipos na susceptibilidade para cancro da mama não revela qualquer associação com a doença (resultados não apresentados).

Tabela 3.3 – Distribuição de genótipos e risco para cancro da mama para todos os polimorfismos estudados (*GSTA2* P110S, S112T e E210A) em ambas as populações, doentes e controlos.

GENÓTIPOS	DOENTES N (%)	CONTROLOS N (%)	MAF		OR _{NÃO-AJUSTADO} (IC A 95%)	OR _{AJUSTADO} (IC A 95%) ^a
			DOENTES	CONTROLOS		
P110S (Ex5+56C>T)						
CC	265 (91,7)	505 (92,8)			1 (Referência)	1 (Referência)
CT	24 (8,3)	39 (7,2)	T: 0,0415±0,0083	T: 0,0358±0,0056	1,173 (0,690 – 1,992)	1,218 (0,705 – 2,104)
TT	-	-				
N.D.	2	3				
S112T (Ex5+63G>C)						
GG	92 (31,8)	155 (28,5)			1 (Referência)	1 (Referência)
GC	146 (50,5)	280 (51,6)	C: 0,4291±0,0206	C: 0,4567±0,0151	0,878 (0,634 – 1,218)	0,873 (0,625 – 1,219)
CC	51 (17,6)	108 (19,9)				
N.D.	2	4				
E210A (Ex7+83A>C)						
AA	257 (88,6)	485 (88,7)			1 (Referência)	1 (Referência)
AC	33 (11,4)	61 (11,2)			1,017 (0,649 – 1,594)	1,008 (0,635 – 1,601)
CC	0 (0)	1 (0,2)	C: 0,0569±0,0096	C: 0,0576±0,007	ND	ND
AC + CC	33	62				
N.D.	-	-				

^a OR's foram ajustados para: idade de diagnóstico (≤30, 31 -49, 50-69 and ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência. N.D. – Não Determinado

Tabela 3.4 – Coeficiente de desequilíbrio de *linkage D'* entre os SNPs estudados.

SNP	P110S	S112T	E210A
P110S	-	0,8992	0,9665
S112T	-	-	0,9696
E210A	-	-	-

Tabela 3.5 – Frequências de haplótipos para o gene *GSTA2* na população Portuguesa obtidas através do software SNPStats.

HAPLÓTIPO ESTIMADO	FREQUÊNCIA DE HAPLÓTIPOS	
	DOENTES	CONTROLOS
GSTA2 (P110; S112; E210) *	0,5713	0,5387
GSTA2 (P110; S112; A210) *	NA	0,00140
GSTA2 (P110; T112; E210) *	0,3303	0,3679
GSTA2 (S110; S112; E210) *	NA	0,003
GSTA2 (P110; T112; A210)	0,0567	0,0562
GSTA2 (S110; T112; E210)	0,0418	0,0328

* Haplótipos identificados por (Tetlow *et al.*, 2001)

3.4 DISCUSSÃO

Existe uma grande lacuna de resultados que relacionem os polimorfismos descritos para o gene *GSTA2* e os diferentes tipos de cancro. Os resultados publicados até à data têm incidido maioritariamente na organização genómica do *GSTA2*, e a sua expressão nos diferentes tecidos (Coles *et al.*, 2000; Coles *et al.*, 2001; Kelley *et al.*, 1994) e a caracterização funcional dos polimorfismos desse gene (Guy *et al.*, 2004; Morel *et al.*, 2002; Ning *et al.*, 2004; Tetlow *et al.*, 2001; Tetlow e Board, 2004).

Alguns dos polimorfismos presentes no gene *GSTA2* foram descritos em três populações etnicamente diferentes por Tetlow *et al.* (Tetlow *et al.*, 2001; Tetlow e Board, 2004). Alguns estudos funcionais desenvolvidos com o intuito de avaliar a

actividade proteica das variantes geradas pelos múltiplos polimorfismos, reportaram não encontrar nenhuma discrepância significativa na actividade proteica de todos os haplótipos gerados quando comparados com o haplótipo mais frequente [GSTA2 (P110; S112; E210)], excepto num [GSTA2 (S110; S112; E210)] (Ning *et al.*, 2004; Tetlow *et al.*, 2001; Tetlow e Board, 2004). Na verdade, a forma variante apresenta uma actividade mais reduzida relativamente à conjugação com determinados substratos, tais como o hidroperóxido de cumeno (Tetlow e Board, 2004), o que sugere que indivíduos com a variante S110 poderão ter um aumento de susceptibilidade para a doença relacionada com o stress oxidativo. Adicionalmente, a GSTA2 está também envolvida na isomerização de Δ^5 -androsteno-3,17-diona, um precursor de várias hormonas esteróides, como a progesterona. A forma variante de GSTA2 S110 mostra ainda uma baixa actividade perante este substrato, o que sugere que indivíduos portadores deste alelo variante podem apresentar baixos níveis de hormonas esteróides. Estes estudos que caracterizam estruturalmente os polimorfismos da classe *Alpha* da glutathione-S-transferase, mostram que os resíduos de aminoácidos de S112T e E210A não se encontram no centro activo. A alteração S112T é conservativa (ambos os resíduos são polares não-carregados) e não revelaram um efeito importante na estrutura da proteína, enquanto a alteração E210A é não-conservativa (A é um resíduo polar pequeno ao passo que E é um resíduo maior e ácido), contudo a posição destes resíduos poderá afectar apenas ligeiramente a função da proteína. Relativamente ao SNP P110S, os estudos efectuados mostraram que a substituição de um resíduo hidrofóbico, não polar (Pro), por um resíduo hidrofílico polar não-carregado (Ser), pode reduzir a rigidez do centro activo relacionado com a perda estrutural do resíduo, e aumentar o carácter hidrofílico (Tetlow e Board, 2004).

Neste trabalho o estudo efectuado visava a análise do possível papel de três polimorfismos comuns não-sinónimos do gene GSTA2 (P110S; S112T; E210A) na susceptibilidade genética para cancro da mama, o qual, no nosso conhecimento, foi o primeiro estudo a relacionar estes polimorfismos com o cancro da mama. No entanto, os resultados obtidos não revelaram qualquer associação entre os diferentes polimorfismos estudados e a susceptibilidade individual para neoplasia mamária (Tabela 3.3). Estes resultados poderão ser explicados pelo facto dos polimorfismos S112T e E210A não provocarem uma diminuição significativa da actividade catalítica, e

o SNP *GSTA2* P110S que poderia estar associado com a alteração desta actividade só se manifesta em heterozigotia e com uma frequência inferior a 10% em ambas as populações (doentes e controlos) (Tabela 3.3). Considerando os estudos epidemiológicos em que foram estudados alguns polimorfismos do gene *GSTA2*, não foram descritas associações entre o polimorfismo S112T e alteração de risco para cancro colorectal, mesotelioma pleural maligno (Landi *et al.*, 2007; Landi *et al.*, 2005) e cancro da próstata (Ning *et al.*, 2004). No entanto, o alelo variante deste polimorfismo parece estar associado com um aumento de risco para cancro do pulmão (Gemignani *et al.*, 2007). Adicionalmente, o nosso estudo mostra que todos os SNPs estudados estão em desequilíbrio de *linkage*. Através da combinação das diferentes variantes alélicas dos polimorfismos do *GSTA2*, foram obtidos seis haplótipos diferentes nas nossas populações (ver Tabela 3.5), o que é concordante com os resultados publicados em estudos prévios (Tetlow *et al.*, 2001; Tetlow e Board, 2004). Contudo, existem diferenças consistentes nas frequências dos haplótipos determinadas para outras populações (Tetlow *et al.*, 2001; Tetlow e Board, 2004), que poderão ser explicadas com base nas diferenças étnicas entre as populações.

Resultados recentes em estudos de cancerigénese têm relacionado expressões aberrantes de isozimas de GSTs com o desenvolvimento e expressão de resistência a uma grande diversidade de compostos químicos, especialmente dos utilizados na terapia anti-cancerígena. Estudos desenvolvidos em tecido tumoral e em linhas celulares resistentes a agentes quimioterapêuticos têm mostrado uma sobre-expressão dessas enzimas (Dialyna *et al.*, 2001; McIlwain *et al.*, 2006). Para além disso, estudos em tecido normal e tecido neoplásico de mama humano, revelaram que na maioria dos casos, o nível médio de GSTs está substancialmente elevado no tecido tumoral quando comparado com o tecido normal, fornecendo assim um mecanismo plausível para a resposta variável ao tratamento (Kelley *et al.*, 1994).

Mesmo assim os nossos resultados sugerem que os polimorfismos no gene *GSTA2* não estão associados com a susceptibilidade individual para cancro da mama, o que é concordante com um estudo recentemente publicado (Andonova *et al.*, 2009). No entanto, não pode ser excluído que outros SNPs em outros genes associados com a detoxificação, poderão estar envolvidos com a susceptibilidade para cancro da mama. Adicionalmente, não pode ser excluído que os polimorfismos avaliados neste trabalho

poderão estar associados com outros aspectos da doença oncológica como resistência a medicamentos/sobrevida dos pacientes, no entanto, serão necessários outros estudos que possam avaliar este efeito.

CAPÍTULO 4

POLIMORFISMOS ENVOLVIDOS NO METABOLISMO DE ESTROGÉNIOS (CYP17, HSD17 β 1, COMT E MNSOD) E O RISCO PARA CANCRO DA MAMA

4. POLIMORFISMOS ENVOLVIDOS NO METABOLISMO DE ESTROGÉNIOS (*CYP17*, *HSD17 β 1*, *COMT* E *MNSOD*) E O RISCO PARA CANCRO DA MAMA

Publicado em Silva *et al.*, 2006, *Oncology Reports*, 16, 781-788

4.1 INTRODUÇÃO

Diferentes polimorfismos genéticos têm sido identificados em genes envolvidos na biossíntese de estrogénios (*e.g.* *CYP17* e *CYP19*) e no metabolismo de estrogénios (*e.g.* *CYP1B1* e *COMT*), polimorfismos esses que podem influenciar as concentrações de estrogénios. Adicionalmente, os catecol-estrogénios produzem ROS o que sugere o potencial envolvimento das formas polimórficas de genes que codificam para enzimas com capacidade de destoxificação de ROS (*e.g.* *MnsOD*).

O gene *CYP17* codifica para uma proteína que está envolvida na conversão da 17-hidroxipregnenolona e 17-hidroxiprogesterona em dehidroepiandrosterona (DHEA) e androstenediona, respectivamente. A região 5'UTR da sequência que codifica para o *CYP17* possui um polimorfismo (T27C) que vai criar um local SP1 (CCACC box) adicional que poderá influenciar a expressão do gene, podendo resultar num aumento da expressão da enzima e conseqüentemente num aumento da síntese de estrogénios (Mitrunen e Hirvonen, 2003; Tworoger *et al.*, 2004). A função deste polimorfismo na susceptibilidade para cancro da mama tem sido intensamente estudada dando origem a resultados contraditórios (Bergman-Jungstrom *et al.*, 1999; Haiman *et al.*, 1999; Huang *et al.*, 1999).

O gene *HSD17 β 1* (17 β -hidroxiesteróide dehidrogenase tipo 1) codifica para uma enzima que catalisa o passo final da biossíntese do estradiol, ou seja a conversão de estrona a estradiol. Têm sido descritos vários SNPs existentes neste gene, no entanto, a função que possam desencadear permanece pouco evidente. Um dos polimorfismos mais estudados deste gene, em estudos de cancro da mama, localiza-se no exão 6 (Gly313Ser) e conduz a uma alteração de aminoácido de glicina para serina no codão 313, contudo os resultados publicados não são indicadores de uma

associação entre este polimorfismo e o cancro da mama (Setiawan *et al.*, 2004; Wu *et al.*, 2003).

O gene *COMT* (Catecol-*O*-Metiltransferase) catalisa a metilação de diversas substâncias endobióticas e xenobióticas prevenindo a formação de quinonas e de reacções redox, e desta forma pode exercer uma função de protecção do DNA das eventuais lesões oxidativas. Uma transição de G para A, dá origem a uma alteração de aminoácido de valina para metionina no codão 108 (Val108Met), sendo responsável por uma menor actividade da enzima COMT. A actividade enzimática conferida pelo genótipo Met/Met corresponde a cerca de 25% da obtida pelo genótipo mais comum (Val/Val), enquanto os indivíduos heterozigóticos apresentam um nível de actividade enzimática intermédio, no entanto, os resultados disponíveis que relacionem este SNP com o risco para cancro da mama são contraditórios (Cheng *et al.*, 2005; Kocabas *et al.*, 2002; Lavigne *et al.*, 1997; Yim *et al.*, 2001).

As enzimas com características antioxidantes, como as superóxido dismutase (SOD), têm como função proteger as células do stress oxidativo, e a formação de ROS tem sido relacionada com a etiologia de diversas doenças. A SOD catalisa a dismutação do radical superóxido em peróxido de hidrogénio (H₂O₂) e oxigénio molecular (O₂). Uma substituição de T para C caracteriza um SNP no gene *MnSOD* provocando uma alteração de valina para alanina no codão 16 (Val16Ala). Este SNP induz uma alteração na estrutura secundária da proteína capaz de afectar a eficiência de transporte da MnSOD para a mitocôndria (Sutton *et al.*, 2005), no entanto, e mais uma vez, os estudos que relacionam este SNP com a susceptibilidade individual para cancro da mama são contraditórios (Cai *et al.*, 2004; Egan *et al.*, 2003; Millikan *et al.*, 2004; Mitrunen e Hirvonen, 2003).

Os estudos de associação em cancro da mama publicados até à data que incluem os polimorfismos anteriormente descritos revelaram resultados contraditórios. Assim sendo, e podendo incluir neste trabalho uma população caracterizada por não apresentar casos de cancro da mama, os Índios Xavante de Sangradouro (Mato Grosso, Brasil), identificada pelo Dr. Guilherme Bezerra de Castro (Castro *et al.*, 2005), a frequência desses polimorfismos pôde ser avaliada também para esta população na tentativa de se obter uma melhor compreensão sobre o potencial papel destes polimorfismos na susceptibilidade para cancro da mama. Para

os polimorfismos que apresentaram uma diferença de frequência mais significativa (apenas para *MnSOD*), quando comparadas com os valores descritos para outras populações, foi efectuado um estudo caso-controlo para a população Portuguesa Caucasiana com o intuito de avaliar o efeito do polimorfismo do gene *MnSOD* (Val16Ala) na susceptibilidade para cancro da mama.

4.2 MATERIAIS E MÉTODOS

4.2.1 DESCRIÇÃO DA POPULAÇÃO

A descrição efectuada no Capítulo 3, ponto 3.2.1 aplica-se neste novo capítulo, à excepção do tamanho da amostra, já que para este estudo foram utilizados 241 doentes com cancro da mama e 457 indivíduos que constituem a população controlo. O diagnóstico histológico foi confirmado em todos os casos e a amostra inclui 213 carcinomas tipo ductal (88,4%), 11 carcinomas tipo lobular (4,6%) e 17 casos classificados como sendo de outro tipo de tumor da mama (7,1%).

Neste trabalho foi incluída uma segunda população controlo constituída por 179 indivíduos, cuja importante característica é não apresentar quaisquer casos de cancro da mama. Esta população foi identificada pelo Dr. Guilherme Bezerra de Castro numa reserva de Índios de Sangradouro. O grupo étnico predominante da reserva é a comunidade Xavante. Muito perto da reserva existe, desde 1906, uma missão Salesiana com uma estrutura de ambulatório médico com capacidade fidedigna para registar todos os eventos epidemiológicos. Esta unidade médica está à disposição da comunidade Xavante (cerca de 10.000 indivíduos) residente perto de Sangradouro. Os dados epidemiológicos disponíveis mostram que esta população se caracteriza por uma absoluta ausência de casos de neoplasia mamária, sendo de referir que a idade média de vida das índias desta reserva não se afasta da idade da população não-índia do Mato Grosso. Os Índios adoptam o modelo de cruzamento *Iroques* não ocorrendo qualquer cruzamento com outras tribos que vivam nas imediações. Um factor interessante, todas as etnias indígenas residentes no estado de Mato Grosso apresentam grupo sanguíneo O Rh⁺ (Castro *et al.*, 2005; Salzano *et al.*, 1997).

4.2.2 EXTRACÇÃO DE DNA

A metodologia utilizada encontra-se descrita no Capítulo 3, ponto 3.2.2.

4.2.3 GENOTIPAGEM DOS POLIMORFISMOS *CYP17* (5'-UTR, T27C), *HSD17β1* (GLY313SER), *COMT* (VAL158MET) E *MNSOD* (VAL16ALA)

A genotipagem dos polimorfismos estudados foi efectuada por PCR-RFLP, técnica detalhada no Capítulo 3, ponto 3.2.3. Os *primers* utilizados nas reacções de PCR bem como a temperatura de *annealing* utilizada para a amplificação de cada SNP, e os produtos obtidos especificamente após RFLP estão descritos na Tabela 4.1.

As condições de PCR utilizadas para amplificação do gene *CYP17* foram previamente descritas por Ambrosone *et al.* (Ambrosone *et al.*, 2003); para amplificação do gene *HSD17β1* foram descritas por Feigelson *et al.* (Feigelson *et al.*, 2001); Garner *et al.* (Garner *et al.*, 2002) descreveram as condições para amplificação do fragmento contendo o SNP do gene *COMT*, tendo sido efectuadas pequenas modificações. Considerando o polimorfismo presente no gene *MnSOD*, os *primers* utilizados foram os descritos por Egan *et al.* (Egan *et al.*, 2003). O PCR foi efectuado utilizando cerca de 50 ng de DNA num volume reaccional de 50 µl, contendo tampão de PCR 1x, 1,5 mM MgCl₂, 0,8 mM dNTP, 1,0 µM de cada *primer* e 1,25U de *AmpliTaq Gold* (Applied Biosystems). O processo de amplificação tem início com um passo de desnaturação a 95 °C durante 7 minutos, 35 ciclos de amplificação incluindo desnaturação a 94 °C durante 30 segundos, temperatura de emparelhamento de acordo com o SNP (ver Tabela 4.1) durante 30 segundos e extensão a 72 °C durante 30 segundos, e um período de extensão final de 10 minutos a 72 °C. Após amplificação 10 µl do produto de PCR resultante foi digerido com a enzima de restrição adequada (ver Tabela 4.1) e posteriormente aplicados em gel de agarose com brometo de etídio (1,0 µg/ml) para posterior visualização. Os produtos de restrição expectáveis para os SNP estudados estão apresentados na Tabela 4.1.

As experiências foram realizadas em duplicado em ensaios independentes, sendo os resultados inconclusivos reanalisados.

Tabela 4.1 – Condições de PCR-RFLP para os polimorfismos estudados.

SNPs	SEQUÊNCIA DOS PRIMERS	FRAGMENTO PCR (BP)	T _{ANNEALING} (°C)	EFEITO DA ENZIMA DE RESTRIÇÃO	PERFIS APÓS ANÁLISE DE RESTRIÇÃO
<i>CYP17</i> (5'-UTR, T27C)	<i>Forward:</i> 5`-CAT TCG CAC TCT GGA GTC-3` <i>Reverse:</i> 5`-AGG CTC TTG GGG TAC TTG-3`	414	56	T → C, local de reconhecimento adicional para <i>MspA1I</i>	TT: 414bp; TC: 414, 291, 123bp; CC: 291, 123bp
<i>HSD17β1</i> (Gly313Ser)	<i>Forward 1:</i> 5`-GGG AGC CGC TCT GGG GCG ATC T-3` <i>Reverse 1:</i> 5`-GGT GCC ACT GTG CTG ATT TTT AAA TTT TCT-3` <i>Forward 2:</i> 5`-AAG CCG ACC CTG CGC TAC TTC AC-3` <i>Reverse 2:</i> 5`-TCT ATC TTA ATT AGC CAC CCA CAG C-3`	349	60	G → A, perda de um local de reconhecimento para <i>BstUI</i>	GG: 192, 80, 77bp; GA: 269, 192, 80, 77bp; AA: 269, 80bp
<i>COMT</i> (Val158Met)	<i>Forward:</i> 5`-TACTGTGGCTACTCAGCTGTGC-3` <i>Reverse:</i> 5`-GTGAACGTGGTGTGAACACC-3`	236	60	G → A, local de reconhecimento adicional para <i>NlaIII</i>	GG: 114, 54, 40, 28bp; GA: 114, 96, 54, 40, 28, 18bp; AA: 96, 54, 40, 28, 18bp
<i>MnSOD</i> (Val16Ala)	<i>Forward:</i> 5`-TAG ACG GTC CCG CGG CGC TGA-3` <i>Reverse:</i> 5`-CCG TAG TCG TAG GGC AGG TCG GGG A-3`	134	66	T → C, perda de um local de reconhecimento para <i>BsaWI</i>	TT: 71, 63bp; TC: 134, 71, 63bp; CC: 134bp

4.2.4 ANÁLISE ESTATÍSTICA

A análise das frequências de Hardy-Weinberg dos alelos de todos os SNPs estudados, para a população de doentes, de controlos e para a população constituída pelos Índios Xavante, foi realizada através de testes de probabilidade exacta disponíveis no software Mendel (versão 5.7.2) (Lange *et al.*, 2001).

A abordagem estatística efectuada encontra-se descrita no Capítulo 3, ponto 3.2.4.

4.3 RESULTADOS

Na Tabela 4.2 estão apresentadas as frequências alélicas e genotípicas determinadas, para cada um dos polimorfismos estudados, na população constituída pelos Índios Xavante. A distribuição dos diferentes genótipos está em equilíbrio de Hardy-Weinberg para todos os polimorfismos excepto para o polimorfismo Val158Met do gene *COMT* ($P < 0,05$).

Tabela 4.2 - Determinação das frequências genotípicas e alélicas para a população constituída pelos Índios Xavante ($n = 179$), para os quatro polimorfismos em estudo.

GENE/POLIMORFISMO	FREQUÊNCIAS		FREQUÊNCIAS ALÉLICAS ± ERRO ASSOCIADO
	GENOTÍPICAS (N)	(%)	
<i>CYP17</i> (5'-UTR, T27C)	T/T = 58	(34,4%)	T = 0,5503 (± 0,0263)
	T/C = 81	(45,3%)	
	C/C = 40	(22,3%)	
<i>HSD17B1</i> Gly313Ser	Gly/Gly = 22	(12,3%)	Gly = 0,6955 (± 0,0243)
	Gly/Ser = 65	(36,3%)	
	Ser/Ser = 92	(51,4%)	
<i>COMT</i> Val158Met	Val/Val = 71	(39,7%)	Val = 0,6620 (± 0,0250)
	Val/Met = 95	(53,1%)	
	Met/Met = 13	(7,3%)	
<i>MnSOD</i> Val16Ala	Val/Val = 10	(5,6%)	Val = 0,2318 (± 0,0223)
	Val/Ala = 63	(35,2%)	
	Ala/Ala = 106	(59,2%)	

Após determinação das frequências alélicas e genotípicas do polimorfismo do gene *CYP17* na população de Índios Xavante, estas foram comparadas com os resultados publicados para outras populações (Caucasiana, Chinesa, Coreana, Japonesa, Afro-Americana) e observou-se que as frequências obtidas para este polimorfismo são similares aos valores reportados nas diferentes populações (Tabela 4.3). No entanto, em todos os resultados encontrados apenas um refere o impacto deste polimorfismo na susceptibilidade para cancro da mama e diz respeito a um pequeno estudo efectuado na população sueca (Bergman-Jungstrom *et al.*, 1999), não tendo sido encontrada qualquer relação em estudos desenvolvidos noutras populações Caucasiana (Haiman *et al.*, 1999), Chinesa (Huang *et al.*, 1999; Wu *et al.*, 2003), Coreana (Shin *et al.*, 2005), Japonesa (Miyoshi *et al.*, 2000) e Afro-Americana (Miyoshi *et al.*, 2000).

Relativamente ao polimorfismo no gene *HSD17β1* (Gly313Ser), os dados publicados até ao momento, em populações Caucasiana e Chinesa, não sugerem qualquer envolvimento deste polimorfismo no risco para neoplasia mamária (Setiawan *et al.*, 2004; Wu *et al.*, 2003). Contudo, as frequências alélicas e genotípicas observadas na população de Índios Xavante (Tabela 4.1) mostram que a frequência do alelo mais comum é maior do que a reportada nas populações Caucasiana e Chinesa (Setiawan *et al.*, 2004; Wu *et al.*, 2003) (ver Tabela 4.3).

Por sua vez, as frequências alélicas e genotípicas do polimorfismo *COMT* (Val108Met) determinadas para a população indígena encontram-se na faixa descrita para as populações Chinesa e Japonesa (Tabela 4.3), no entanto os estudos desenvolvidos em cancro da mama em populações Asiáticas são contraditórios (Cheng *et al.*, 2005; Kocabas *et al.*, 2002; Lavigne *et al.*, 1997; Yim *et al.*, 2001).

Em relação ao polimorfismo do gene *MnSOD* (Val16Ala), encontraram-se diferenças potencialmente relevantes entre as frequências alélicas e genotípicas determinadas para a população de Índios Xavante (Tabela 4.2) e as frequências publicadas em estudos efectuados noutras populações (Tabela 4.3). Tendo em consideração que os resultados disponíveis relativamente ao papel deste polimorfismo na susceptibilidade para cancro da mama, embora contraditórios, são significativamente diferentes dos obtidos para a nossa população, caracterizada pela inexistência de casos de cancro da mama, efectuámos um estudo caso-controlo na

população Portuguesa, com o objectivo de se avaliar o potencial efeito deste polimorfismo na susceptibilidade individual para cancro da mama.

Tabela 4.3 – Frequências alélicas e genotípicas dos polimorfismos estudados publicadas em estudos desenvolvidos em diferentes populações.

POPULAÇÃO	REFERÊNCIA	MAF* (%)	FREQUÊNCIA GENOTÍPICA (%)
CYP17 5'UTR, T27C			
Caucasiana	(Bergman-Jungestrom <i>et al.</i> , 1999)	C=31,2%	T/T=45,0%; T/C=47,0%
	(Haiman <i>et al.</i> , 1999)	C=40,0%	T/T=35,1%; T/C=49,7%
	(Miyoshi e Noguchi, 2003)	C=45,0%	T/T=34,0%; T/C=42,0%
	(Miyoshi <i>et al.</i> , 2000)	C=34,4%	T/T=49,0%; T/C=33,0%
	(Miyoshi <i>et al.</i> , 2000)	C=37,8%	T/T=39,0%; T/C=47,0%
	(Miyoshi <i>et al.</i> , 2000)	C=38,1%	T/T=37,0%; T/C=50,0%
Asiática	(Huang <i>et al.</i> , 1999)	T=47,2%	T/T=22,2%; T/C=50,0%
	(Shin <i>et al.</i> , 2005)	T=43,3%	T/T=20,8%; T/C=45,1%
	(Wu <i>et al.</i> , 2003)	T=41,1%	T/T=16,2%; T/C=49,6%
	(Miyoshi <i>et al.</i> , 2000)	C=48,2%	T/T=25,0%; T/C=54,0%
Outras	(Miyoshi <i>et al.</i> , 2000)	C=37,2%	T/T=39,0%; T/C=48,0%
	(Miyoshi e Noguchi, 2003)	C=35,0%	T/T=42,0%; T/C=46,0%
HSD17β1 Gly313Ser			
Caucasiana	(Setiawan <i>et al.</i> , 2004)	Gly=46,4%	Gly/Gly=20,7%; Gly/Ser=51,4%
Asiática	(Wu <i>et al.</i> , 2003)	Ser=43,4%	Gly/Gly=32,4%; Gly/Ser=48,3%
COMT Val158Met			
Caucasiana	(Wedren <i>et al.</i> , 2003)	Val=44,0%	Val/Val=18,3%; Val/Met=49,4%
	(Lavigne <i>et al.</i> , 1997)	Val=48,2%	Val/Val=23,7%; Val/Met=49,1%
	(Sazci <i>et al.</i> , 2004)	Met=39,7%	Val/Val=26,6%; Val/Met=62,7%
	(Kocabas <i>et al.</i> , 2002)	Met=39,3%	Val/Val=34,0%; Val/Met=53,4%
Asiática	(Yim <i>et al.</i> , 2001)	Met=23,9%	Val/Val=62,0%; Val/Met=28,2%
	(Cheng <i>et al.</i> , 2005)	Met=25,5%	Val/Val=62,0%; Val/Met=28,2%
	(Huang <i>et al.</i> , 1999)	Met=25,2%	Val/Val=52,8%; Val/Met=44,0%
	(Wu <i>et al.</i> , 2003)	Met=27,1%	Val/Val=53,3%; Val/Met=39,2%
	(Wu <i>et al.</i> , 2003)	Met=34,3%	Val/Val=43,7%; Val/Met=44,2%
	(Wu <i>et al.</i> , 2003)	Met=26,5%	Val/Val=54,2%; Val/Met=38,6%
MnSOD Val16Ala			
Caucasiana	(Egan <i>et al.</i> , 2003)	Ala=49,7%	Val/Val=26,2%; Val/Ala=48,3%
	(Mitrunen <i>et al.</i> , 2001)	Ala=44,3%	Val/Val=31,7%; Val/Ala=47,9%
	(Millikan <i>et al.</i> , 2004)	Val=49,3%	Val/Val=23,4%; Val/Ala=51,6%
	(Kocabas <i>et al.</i> , 2005)	Val=44,0%	Val/Val=27,0%; Val/Ala=39,0%
Asiática	(Cai <i>et al.</i> , 2004)	Ala=14,0%	Val/Val=73,9%; Val/Ala=24,2%
	(Cheng <i>et al.</i> , 2005)	Ala=14,6%	Val/Val=73,1%; Val/Ala=24,5%
Outras	(Millikan <i>et al.</i> , 2004)	Ala=44,7%	Val/Val=29,0%; Val/Ala=52,7%

* MAF – Alelo de Menor Frequência (do inglês, *Minor Allele Frequency*)

As características gerais das populações de doentes e controlos são apresentadas na Tabela 4.4. Estes resultados revelam não existir diferenças significativas entre os grupos etários, no entanto indivíduos com hábitos de consumo de álcool regular e indivíduos fumadores estão sob representados na população de doentes quando comparado com a população controlo (Tabela 4.4).

Tabela 4.4 – Características gerais das populações de doentes e controlos estudadas para o polimorfismo no gene *MnSOD*.

CARACTERÍSTICAS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a
Idade			
30 – 39	9 (3,7%)	20 (4,4%)	0,994
40 – 49	42 (17,4%)	78 (17,1%)	
50 – 59	56 (23,2%)	109 (23,9%)	
60 – 69	75 (31,1%)	141 (30,9%)	
≥ 70	59 (24,5%)	109 (23,9%)	
Hábitos Tabágicos			
Nunca e Ex-F.	208 (86,7%)	417 (91,6%) ^b	0,038
Fumadores	32 (13,3%)	38 (8,4%)	
N.D.	1	2	
Hábitos Alcoólicos			
Nunca	186 (77,2%)	376 (83,0%)	<0,001
Social	19 (7,9%)	52 (11,5%)	
Regular	36 (14,9%)	25 (5,5%)	
N.D.	4	0	
<i>MnSOD</i> Val16Ala			
CC	59 (24,5%)	99 (21,7%)	0,497
CG	146 (60,6%)	276 (60,4%)	
GG	36 (14,9%)	82 (17,9%)	

^a Ver Materiais e Métodos; ^b Incluí 6 Ex-Fumadores; N.D. – Não Determinado

Os resultados obtidos relativamente ao polimorfismo do gene *MnSOD*, demonstraram que as frequências alélicas dos alelos mais comuns são $0,5477 \pm 0,0277$ e $0,5186 \pm 0,0615$ para doentes e controlos, respectivamente, e que as frequências genotípicas (Tabela 4.4) observadas para a população controlo estão de acordo com os

resultados descritos previamente na literatura para outras populações Caucásicas (Egan *et al.*, 2003; Millikan *et al.*, 2004; Mitrunen *et al.*, 2001).

As frequências genótípicas obtidas para o SNP no gene *MnSOD* nas populações de doentes e controlos não se encontram em equilíbrio de Hardy-Weinberg ($P < 0,001$, teste de probabilidade exacta).

Os resultados obtidos para o polimorfismo presente no gene *MnSOD* não sustentam uma associação entre a presença de um determinado genótipo e a susceptibilidade para cancro da mama. Não foram registadas diferenças significativas entre as frequências genótípicas da população de doentes e da população controlo (Tabela 4.4), e mesmo após análise por regressão logística não se obtiveram valores de OR significativos relativamente ao efeito dos diferentes genótipos depois de ajustados para a idade, consumo de tabaco e consumo de álcool (Tabela 4.5). Contudo, após estratificação por status de amamentação observou-se que em mulheres que nunca tinham amamentado, a presença do alelo variante está associada com uma diminuição de risco para a neoplasia mamária, sendo que este efeito quase atinge a significância [$OR_{\text{Ajustado}} = 0,575 (0,327 - 1,011)$] ($P = 0,054$), ver Tabela 4.5.

Tabela 4.5 – Valores de OR (IC 95%) para cancro da mama determinado para os genótipos gerados pelo polimorfismo presente no gene *MnSOD* (Val16Ala).

GENÓTIPOS PARA O SNP VAL16 ALA; N (NÚMERO DE CASOS)	OR _{NÃO-AJUSTADO} (IC 95%)	OR _{AJUSTADO} (IC 95%)
TODOS OS DOENTES		
Val/Val; (n=59)	1 (Referência)	1 (Referência)
Val/Ala; (n=146)	0,888 (0,607 – 1,298)	0,860 (0,583 – 1,268)
Ala/Ala; (n=96)	0,737 (0,444 – 1,224)	0,733 (0,436 – 1,232)
Val/Ala + Ala/Ala; (n=182)	0,853 (0,590 – 1,533)	0,837 (0,574 – 1,219)
DOENTES QUE AMAMENTARAM		
Val/Val; (n=35)	1 (Referência)	1 (Referência)
Val/Ala; (n=106)	1,090 (0,698 – 1,703)	0,973 (0,612 – 1,547)
Ala/Ala; (n=27)	0,931 (0,521 – 1,665)	0,905 (0,495 – 1,653)
Val/Ala + Ala/Ala; (n=133)	1,054 (0,683 – 1,626)	0,958 (0,611 – 1,504)
DOENTES QUE NUNCA AMAMENTARAM		
Val/Val; (n=22)	1 (Referência)	1 (Referência)
Val/Ala; (n=39)	0,636 (0,359 – 1,125)	0,607 (0,339 – 1,088)
Ala/Ala; (n=9)	0,494 (0,216 – 1,131)	0,468 (0,201 – 1,088)
Val/Ala + Ala/Ala; (n=48)	0,566 (0,321 – 0,999)	0,575 (0,327 – 1,011)

OR's foram ajustados para: idade de diagnóstico (30-39, 40-49, 50-59, 60-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

4.4 DISCUSSÃO

Uma vez que os genes polimórficos *CYP17*, *HSD17β1*, *COMT* e *MnSOD*, estão envolvidos no metabolismo de estrogénios, foram seleccionados diferentes polimorfismos nesses genes, descritos previamente em diversos estudos caso-controlo para cancro da mama, com o objectivo de determinar a sua frequência na população indígena resistente a cancro da mama (Índios Xavante). Os resultados obtidos para o polimorfismo *CYP17* (T27C) mostraram que a sua frequência alélica na população indígena se encontra numa faixa idêntica à descrita para outras populações (Tabela 4.3). Uma vez que a frequência deste polimorfismo na população constituída pelos Índios Xavante é semelhante às frequências descritas para outras populações, e

apenas um (Bergman-Jungstrom *et al.*, 1999) em seis estudos ((Haiman *et al.*, 1999; Huang *et al.*, 1999; Miyoshi e Noguchi, 2003; Shin *et al.*, 2005; Wu *et al.*, 2003) reporta uma associação não elevada (OR = 2 na presença do alelo variante), é, portanto, razoável considerar-se que este polimorfismo possa não desenvolver um papel fundamental na susceptibilidade para cancro da mama.

Quando analisadas as frequências alélicas obtidas na população indígena para o polimorfismo no gene *HSD17β1*, observou-se que os alelos mais comuns estão sub-representados quando comparados com as populações Caucasiana e Chinesa (Tabelas 4.2 e 4.3). Existem ainda dois grandes estudos incidentes na susceptibilidade para cancro da mama que mostram consistentemente uma ausência de associação deste polimorfismo com esta neoplasia (Setiawan *et al.*, 2004; Wu *et al.*, 2003). Considerando os dados no seu conjunto, os resultados obtidos não sustentam um potencial envolvimento deste polimorfismo na susceptibilidade para cancro da mama.

Para o polimorfismo no gene *COMT*, a frequência observada na população constituída pelos Índios Xavante encontra-se no intervalo descrito para outras populações (Tabela 4.2) e encontram-se muito próximas das frequências descritas para as populações Asiáticas (Tabela 4.3). As frequências genotípicas observadas para a população indígena não se encontram em equilíbrio de Hardy-Weineberg, o que sugere que este polimorfismo está sujeito a pressão selectiva. Os resultados publicados até à data que incluem este SNP em estudos de risco para cancro da mama são contraditórios (Bergman-Jungstrom *et al.*, 1999; Haiman *et al.*, 1999; Huang *et al.*, 1999), sendo a maioria das associações positivas encontradas em populações Asiáticas (Tabela 4.3) e um estudo na população Turca que associa a presença do genótipo Met/Met com aumento de risco para cancro da mama (Sazci *et al.*, 2004). Os estudos com associação positiva para os alelos variantes do SNP no gene *COMT* e o risco de cancro da mama revelaram valores de OR entre 1,3 (n = 740) (Cheng *et al.*, 2005) e 3,6 (n = 125) (Huang *et al.*, 1999). De acordo com estes resultados, e uma vez que o número de casos avaliados nos diferentes estudos, em oito dos dez estudos o número de doentes estudados, é inferior a 250, não podemos excluir a potencial função do alelo variante deste polimorfismo na susceptibilidade para cancro da mama.

A frequência do alelo variante para o SNP do gene *MnSOD* na população de Índios Xavante é superior às frequências descritas para outras populações (Tabelas 4.2

e 4.3). Este gene codifica para uma proteína que catalisa a dismutação de dois radicais superóxido na mitocôndria, com formação de peróxido de oxigénio (H_2O_2) e oxigénio (O_2). Quando existe uma produção excessiva de radicais superóxido, consequência do metabolismo endógeno (*e.g.* do metabolismo do estradiol, no qual aniões superóxido são produzidos via ciclo redox de quinonas e semi-quinonas, e outros intermediários), da exposição a agentes tóxicos, ou devido a processos patológicos, e/ou quando não existem mecanismo de defesa *in vivo* suficientes, pode ocorrer stress oxidativo, o que conduz a lesão no DNA, lipoperoxidação, modificação da proteína, disrupção da membrana e lesão mitocondrial. A função do polimorfismo do gene *MnSOD* não está completamente compreendida. Estudos recentes mostraram que a forma variante deste SNP é transportada através da membrana da mitocôndria de forma mais eficiente (Sutton *et al.*, 2005), o que sugere que, indivíduos com pelo menos um alelo variante podem ter uma maior actividade de *MnSOD*.

A frequência do polimorfismo estudado do gene *MnSOD* na população Portuguesa é semelhante aos valores descritos para outras populações Caucásicas (Tabelas 4.3 e 4.4), pelo que o desvio ao equilíbrio de Hardy-Weineberg sugere que este polimorfismo poderá estar sob pressão selectiva. Os resultados obtidos relativamente ao papel deste SNP no cancro da mama e concretamente na população Portuguesa não revelaram o seu potencial envolvimento na susceptibilidade para esta neoplasia (ver Tabela 4.5), o que está de acordo com os resultados publicados por outros grupos (Cai *et al.*, 2004; Cheng *et al.*, 2005; Egan *et al.*, 2003; Kocabas *et al.*, 2005; Millikan *et al.*, 2004). Existe, contudo, um estudo publicado por Mitrunen *et al.*, cujos resultados revelaram que os portadores de pelo menos um alelo variante têm um aumento de risco para neoplasia mamária (Mitrunen *et al.*, 2001). A inconsistência de resultados pode ser justificada pelos diferentes padrões genéticos, ou diferentes estilos de vida (*e.g.* diferentes níveis de exposição a cancerígenos químicos, diferenças no consumo de antioxidantes). Diversos estudos descreveram uma interacção gene-ambiente entre este polimorfismo e os níveis de antioxidantes (Li *et al.*, 2005), no entanto esses níveis são altamente influenciados por hábitos dietéticos culturais, o que poderá justificar a inconsistência encontrada nos resultados publicados.

Após estratificação da população de acordo com o *status* de amamentação (mulheres que amamentaram *versus* mulheres que nunca amamentaram) observou-se

que, em mulheres que nunca tinham amamentado, a presença do alelo variante está associada com uma diminuição de risco para cancro da mama ($OR_{Ajustado} = 0,575$), sendo que este valor quase atinge a significância estatística ($P = 0,054$). Um estudo de revisão onde foram incluídos 47 estudos desenvolvidos em 30 países, envolvendo cerca de 50.000 mulheres com cancro da mama e 97.000 controlos, sugeriu que a amamentação pode ser responsável por 2/3 de redução de casos com cancro da mama. Quanto maior o período de amamentação menor o risco associado com o desenvolvimento de cancro da mama. Foi estimado que a incidência de cancro da mama em países desenvolvidos pode ser reduzida para menos de metade (de 6,3 a 2,7%) se o período de amamentação fosse mais longo (Rea, 2004). Sabe-se que o leite materno contém células ductais esfoliadas, sendo possível detectar nas células obtidas de amostras de leite materno diferentes tipos de lesão de DNA induzidas por compostos aromáticos (Thompson *et al.*, 2002). Deste modo, considerando estes resultados é razoável assumir-se que a esfoliação de células ductais como consequência da amamentação pode remover um significativo número de células com lesão genética, prevenindo a sua transformação em células neoplásicas. Neste estudo o potencial efeito protector atribuído à variante alélica (A1a) do SNP presente no gene *MnSOD*, em mulheres que nunca amamentaram, pode estar relacionado com um transporte para a mitocôndria mais eficiente resultando numa maior actividade como descrito por Sutton *et al.* (Sutton *et al.*, 2005), conduzindo a uma destoxificação mais eficiente da superóxido decorrente da exposição a genotóxicos ambientais ou endógenos (*e.g.* catecol-estrogénios). No entanto, e uma vez que existem outros polimorfismos neste mesmo gene⁴ e que existem outras enzimas polimórficas também envolvidas na destoxificação de ROS ou produtos provenientes de reacções celulares (*e.g.* GSTM1, catalase), não podemos excluir que outros polimorfismos neste ou noutros genes possam, sozinhos ou em associação, estar envolvidos na susceptibilidade para cancro da mama. De salientar que este mesmo trabalho foi incluído em dois estudos de meta-análise (Bag e Bag, 2008; Wang *et al.*, 2009).

Adicionalmente, através da comparação das frequências genotípicas e alélicas da população caracterizada pela ausência de casos de cancro mama (Índios Xavante),

⁴ <http://snp500cancer.nci.nih.gov/>

com as de outras populações com elevada incidência desta neoplasia, é possível obter-se informação sobre o papel dos diferentes polimorfismos genéticos no risco para cancro da mama, permitindo uma rápida identificação dos polimorfismos mais relevantes que possam estar relacionados com susceptibilidade para cancro da mama.

CAPÍTULO 5

A IDADE DA MENOPAUSA E POLIMORFISMOS NO
GENE *XRCC1*: O SEU PAPEL NO CANCRO DA MAMA

5. A IDADE DA MENOPAUSA E POLIMORFISMOS NO GENE XRCC1: O SEU PAPEL NO CANCRO DA MAMA

Publicado em parte em Silva *et al.*, 2007, *Cancer Detect Prev.*, 31, 303-309

5.1 INTRODUÇÃO

A utilização de metodologias funcionais, descritas em vários estudos, como medida de avaliação da actividade de reparação de DNA sugere a existência de variabilidade na capacidade de reparação entre diferentes populações, o que conduz à hipótese de que esta variação é consequência da múltipla combinação de alelos que revelam ligeiras variações na função biológica (Duell *et al.*, 2001). A possível função atribuída a deficiências na reparação de DNA no desenvolvimento do cancro, nomeadamente no cancro da mama, tem sido alvo de grande interesse uma vez que existem evidências de que doentes com cancro da mama apresentam uma menor capacidade na reparação de lesões de DNA induzidas por radiação (Duell *et al.*, 2001).

Diversos estudos têm sido desenvolvidos com o objectivo de identificar variantes polimórficas em genes de reparação potencialmente associados com o cancro da mama, sendo sugerido que variações nos genes *XRCC1*, *XRCC2*, *XRCC4*, *LIG4*, *RAD52*, *ERCC1* e *BRCA2* (Goode *et al.*, 2002; Kim *et al.*, 2002; Kuschel *et al.*, 2002; Lee *et al.*, 2005a; Rafii *et al.*, 2002) podem estar associadas com susceptibilidade para cancro da mama. De entre os vários genes envolvidos na via BER, foram seleccionados para este estudo os genes, *XRCC1* e *OGG1*.

As enzimas que iniciam e caracterizam particularmente a via BER são as DNA glicosilases, e a que codifica para o gene *OGG1* tem sido considerada como uma das DNA glicosilases mais importantes, visto catalisar a excisão da base modificada 8-hidroxidesoxiguanina (8-OhdG), a maior lesão produzida por ROS como subproduto do metabolismo endógeno ou exposição a agentes oxidantes ambientais (Xu *et al.*, 2002). A base 8-OhdG é altamente mutagénica e, se não for excisada, pode causar transversões G:C para T:A, que ocorrem com frequência em vários oncogenes e genes supressores de tumores (Choi *et al.*, 2003). Embora não existam evidências da sua

associação com o cancro da mama, o facto de ser uma das DNA glicosilases mais importantes na via BER levou a que fosse seleccionado para este estudo.

Considerando o papel dos polimorfismos do gene *XRCC1* no cancro da mama, os resultados publicados até à data não têm sido conclusivos (Deligezer e Dalay, 2004; Duell *et al.*, 2001; Figueiredo *et al.*, 2004; Kim *et al.*, 2002; Moullan *et al.*, 2003; Shen *et al.*, 2005; Shu *et al.*, 2003; Smith *et al.*, 2003b). O gene *XRCC1*, um dos vinte genes participantes na via de reparação por excisão de bases, desempenha variadas funções na reparação da base de DNA lesada e nas quebras de DNA em cadeia simples, relacionadas com uma grande variedade de lesões exógenas ou endógenas não-volumosas e quebras de cadeia simples (Chacko *et al.*, 2005; Duell *et al.*, 2001; Duell *et al.*, 2000; Han *et al.*, 2003). No entanto, a proteína que codifica para o gene *XRCC1* não tem actividade enzimática conhecida, é uma proteína constituída por três domínios distintos que apresentam locais de interacção com a DNA polimerase β , polimerase poli(ADP-ribose) e DNA ligase III, o que sugere que esta proteína possa actuar como factor nuclear na via BER favorecendo a ligação de diferentes componentes no local de acção para promover a eficiência da maquinaria de reparação. Têm sido identificados alguns SNPs no gene *XRCC1* (Duell *et al.*, 2001; Zhang *et al.*, 2006). Estes polimorfismos podem alterar a proficiência da via BER e, deste modo, conferir predisposição genética para o cancro da mama.

Já que os dados disponíveis correspondentes à função dos polimorfismos no gene *XRCC1* para o cancro da mama têm sido contraditórios (Deligezer e Dalay, 2004; Duell *et al.*, 2001; Figueiredo *et al.*, 2004; Kim *et al.*, 2002; Moullan *et al.*, 2003; Shen *et al.*, 2005; Shu *et al.*, 2003; Smith *et al.*, 2003b), para a elaboração deste trabalho foi efectuado um estudo caso-control numa população Caucasiana Portuguesa com o intuito de se avaliar o potencial papel desempenhado pelos polimorfismos R194W e R399Q presentes no gene *XRCC1* na susceptibilidade individual para cancro da mama. Foi igualmente avaliada a distribuição alélica nas populações em estudo do polimorfismo S326C do gene *OGG1*. Tal como descrito no capítulo anterior, secção 4.1, foi também inserida neste estudo a população indígena dos Índios Xavante caracterizada por não apresentar casos de cancro da mama.

5.2 MATERIAIS E MÉTODOS

5.2.1 DESCRIÇÃO DA POPULAÇÃO

A descrição efectuada no Capítulo 4, ponto 4.2.1 aplica-se neste novo capítulo.

5.2.2 EXTRACÇÃO DE DNA

A metodologia utilizada encontra-se descrita no Capítulo 3, ponto 3.2.2.

5.2.3 GENOTIPAGEM DOS POLIMORFISMOS *OGG1* (SER326CYS), *XRCC1* (ARG194TRP) E (ARG399GLN)

A genotipagem dos polimorfismos estudados foi efectuada por PCR-RFLP, técnica detalhada no Capítulo 3, ponto 3.2.3. Os *primers* utilizados nas reacções de PCR bem como a temperatura de *annealing* utilizada para a amplificação de cada SNP, e os produtos obtidos especificamente após RFLP estão descritos na Tabela 5.1. Com vista a validar a metodologia aplicada, PCR-RFLP, foram escolhidas e sequenciadas seis amostras com diferentes genótipos, tendo sido obtida uma concordância de 100% entre os resultados obtidos pelos dois processos. O processo de sequenciação foi desenvolvido pela StabVida.

O PCR foi efectuado utilizando cerca de 50 ng de DNA num volume reaccional de 50 µl. Para os polimorfismos *XRCC1* Arg399Gln e *OGG1* Ser326Cys a mistura reaccional continha: tampão PCR 1X, 0,2 mM de cada dNTP, 0,6 µM de cada *primer*, 2,5 mM de MgCl₂, 0,75 U de *Immolase* (Bioline). A mistura reaccional do polimorfismo *XRCC1* Arg194Trp é semelhante, tendo como diferença a concentração de MgCl₂ e de cada dNTP que foram optimizadas para 1,5 mM e 0,15mM, respectivamente.

As condições de amplificação para os três SNPs foram semelhantes. Para os polimorfismos *XRCC1* Arg194Trp e *XRCC1* Arg399Gln estas consistiram em 7 minutos de activação inicial a 95 °C, seguido de 32 ciclos de desnaturação a 94 °C durante 30 segundos, *annealing* a 62 °C e 64 °C, respectivamente durante 30 segundos, extensão a 72 °C durante 30 segundos e 10 minutos de extensão final a 72 °C. Para o polimorfismo *OGG1* Ser326Cys as condições foram iguais às utilizadas para amplificação do fragmento do SNP *XRCC1* Arg399Gln, tendo como única diferença o número de ciclos

que foi otimizado para 30. Após amplificação 10 µl do produto de PCR resultante foi digerido, segundo instruções do fabricante, com a enzima de restrição adequada (ver Tabela 5.1) e posteriormente aplicados em gel de agarose com brometo de etídio (1,0 µg/ml) para posterior visualização. Os produtos de restrição expectáveis para os SNP estudados estão apresentados na Tabela 5.1.

As experiências foram realizadas em duplicado em ensaios independentes, sendo os resultados inconclusivos reanalisados.

5.2.4 ANÁLISE ESTATÍSTICA

A análise das frequências de Hardy-Weinberg dos alelos de todos os SNPs estudados, para a população de doentes, de controlos e para a população constituída pelos Índios Xavante, foi realizada através de testes de probabilidade exacta disponíveis no software Mendel (versão 5.7.2) (Lange *et al.*, 2001).

A abordagem estatística efectuada encontra-se descrita no Capítulo 3, ponto 3.2.4.

5.3 RESULTADOS

Na Tabela 5.2 estão apresentadas as frequências alélicas e genóticas determinadas, para cada um dos polimorfismos estudados, na população constituídas pelos Índios Xavante. A distribuição dos diferentes genótipos está em equilíbrio de Hardy-Weinberg para todos os polimorfismos excepto para o polimorfismo Ser326Cys do gene *OGG1* ($P < 0,05$).

À data de realização deste trabalho existiam apenas dois estudos que incluíam o polimorfismo *OGG1* Ser326Cys (Choi *et al.*, 2003; Vogel *et al.*, 2003), de onde se verificou uma maior prevalência do alelo Cys na população Asiática quando comparado com a população Caucasiana (0.2408 <Cys< 0.5021). Assim, ao comparar-se a frequência alélica obtida para a população indígena constituída pelos Índios Xavante (Cys = 0.2765), observou-se que esta se encontrava dentro do intervalo registado para as outras populações. Em ambos os estudos efectuados não foi demonstrada nenhuma associação directa entre este polimorfismo e o cancro da mama.

Até à data outros estudos têm sido desenvolvidos que incluem o estudo deste polimorfismo embora não sejam muito conclusivos. Há estudos que não revelaram qualquer associação entre este SNP e o cancro da mama (Zhang *et al.*, 2006). O trabalho desenvolvido por Rossner *et al.* descreve um aumento de risco associado com a presença de pelo menos um alelo variante deste SNP e o consumo de álcool moderado em mulheres com elevado índice de massa corporal (BMI) (Rossner, Jr. *et al.*, 2006). Um outro estudo desenvolvido na população Tailandesa (Sangrajrang *et al.*, 2008) descreve um aumento de risco apenas perante o efeito conjunto de dois ou três alelos variantes para os genes *OGG1*, *XRCC1* e *APEX1*. Tudo indica que não há concordância entre a função atribuída a este polimorfismo e a sua influência na neoplasia mamária.

Dado que nenhum dos polimorfismos, na população indígena, apresente uma grande discrepância quanto às frequências genótípicas e alélicas observadas nas outras populações (Tabela 5.3), foram seleccionados os polimorfismos *XRCC1* Arg194Trp e Arg399Gln para um estudo caso-controlo na população Portuguesa, já que existem alguns estudos que os associam com o cancro da mama (Chacko *et al.*, 2005; Duell *et al.*, 2001; Kim *et al.*, 2002).

Tabela 5.1 - Condições de PCR-RFLP para os polimorfismos estudados.

SNPs	SEQUÊNCIA DOS <i>PRIMERS</i>	FRAGMENTO PCR (BP)	T _{ANNEALING} (°C)	EFEITO DA ENZIMA DE RESTRIÇÃO	PERFIS APÓS ANÁLISE DE RESTRIÇÃO
<i>XRCC1</i> (Arg194Trp)	<i>Forward:</i> 5'-CCG TGT GAA GGA GGA GGA TGA- 3' <i>Reverse:</i> 5'-CCT CCA GAC CTC TCA ACC CTC- 3'	275	62	C → T, perda de um local de reconhecimento para <i>MspI</i>	CC: 213, 42, 20bp; CT: 233, 213, 42, 20bp; TT: 233, 42bp
<i>XRCC1</i> (Arg399Gln)	<i>Forward:</i> 5'-GCA TCG TGC GTA AGG AGT- 3' <i>Reverse:</i> 5'-CCG CTC CTC TCA GTA GTC TGC- 3'	203	64	G → A, perda do local de reconhecimento para <i>MspI</i>	GG: 143, 60bp; GA: 203, 143, 60bp AA: 203bp
<i>OGG1</i> (Ser326Cys)	<i>Forward:</i> 5'-GGT GGC CCT AAA GGA CTC TCC-3' <i>Reverse:</i> 5'- CCA TCC TTA GCG CTG TCT CCC-3'	456	64	C → G, local de reconhecimento para <i>SatI</i>	CC: 456bp; CG: 456, 258, 198bp; GG: 258, 198bp

Tabela 5.2 - Determinação das frequências genotípicas e alélicas para a população constituída pelos Índios Xavante (n = 179), para os polimorfismos em estudo.

GENE/POLIMORFISMO	FREQUÊNCIAS GENOTÍPICAS (N)	(%)	FREQUÊNCIAS ALÉLICAS ± ERRO ASSOCIADO
<i>XRCC1</i> (Arg194Trp)	Arg/Arg = 100	(55,9%)	Arg = 0,7374 (± 0,0233)
	Arg/Trp = 64	(35,7%)	
	Trp/Trp = 15	(8,4%)	Trp = 0,2626 (± 0,0233)
<i>XRCC1</i> (Arg399Gln)	Arg/Arg = 124	(69,3%)	Arg = 0,8352 (± 0,0196)
	Arg/Gln = 51	(28,5%)	
	Gln/Gln = 4	(2,2%)	Gln = 0,1648 (± 0,0196)
<i>OGG1</i> (Ser326Cys)	Ser/Ser = 80	(44,7%)	Ser = 0,7235 (± 0,0236)
	Ser/Cys = 99	(55,3%)	
	Cys/Cys = 0	(0%)	Cys = 0,2765 (± 0,0236)

As características principais das populações envolvidas no estudo caso-controlo estão descritas na Tabela 5.4. Não existem diferenças significativas entre doentes e controlos quando se considera a idade e as frequências genotípicas calculadas para os polimorfismos presentes no gene *XRCC1*. Contudo, os grupos de fumadores e de consumidores de álcool são mais prevalentes na população de doentes do que na população controlo. Os resultados obtidos para as frequências alélicas para o alelo mais comum do SNP Arg194Trp foram 0,94 e 0,91 para a população controlo e para a população de doentes, respectivamente, e considerando agora o alelo mais comum do SNP Arg399Gln as frequências obtidas foram 0,65 e 0,68, respectivamente, população controlo e população de doentes.

As frequências genotípicas dos polimorfismos estudados do gene *XRCC1*, em ambas as populações, encontravam-se em equilíbrio de Hardy-Weineberg ($P < 0,001$, teste de probabilidade exacta), e as obtidas para a população controlo foram coincidentes com as frequências obtidas para outras populações Caucásicas (Duell *et al.*, 2001; Figueiredo *et al.*, 2004; Moullan *et al.*, 2003; Shen *et al.*, 2005; Smith *et al.*, 2003a). Em ambas as populações, controlo e doentes, os dois *loci* não se encontram em desequilíbrio de *linkage*.

Tabela 5.3 – Frequências genotípicas e alélicas dos polimorfismos estudados publicadas em estudos desenvolvidos em diferentes populações.

POPULAÇÃO	REFERÊNCIAS	FREQUÊNCIA GENOTÍPICA (%)	MAF (%)
<i>XRCC1</i> Arg194Trp			
Caucasiana	(Smith <i>et al.</i> , 2003b)	Arg/Arg = 91,0%; Arg/Trp = 8,7%	Trp = 4,7%
	(Moullan <i>et al.</i> , 2003)	Arg/Arg = 86,5%; Arg/Trp = 13,1%	Trp = 6,9%
	(Duell <i>et al.</i> , 2001)	Arg/Arg = 86,0%; Arg/Trp = 13,1%	Trp = 7,5%
	(Shen <i>et al.</i> , 2005)	Arg/Arg = 87,3% ; Arg/Trp ou Trp/Trp = 12,7%	
Asiática	(Kim <i>et al.</i> , 2002)	Arg/Arg = 44,9% ; Arg/Trp = 41,9%	Trp = 34,2%
Afro-Americana	(Duell <i>et al.</i> , 2001)	Arg/Arg = 87,5%; Arg/Trp = 12,5%	Trp = 13,5%
Indiana	(Chacko <i>et al.</i> , 2005)	Arg/Arg = 78,1 % ; Arg/Trp = 18,7%	Trp = 12,6%
Turca	(Deligezer e Dalay, 2004)	Arg/Arg = 89,5%; Arg/Trp = 10,5%	Trp = 5,3%
<i>XRCC1</i> Arg399Gln			
Caucasiana	(Moullan <i>et al.</i> , 2003)	Arg/Arg = 40,7%; Arg/Gln = 46,8%	Gln = 35,9%
	(Smith <i>et al.</i> , 2003b)	Arg/Arg = 43,1%; Arg/Gln = 46,1%	Gln = 33,9%
	(Duell <i>et al.</i> , 2001)	Arg/Arg = 43,0%; Arg/Gln = 41,0%	Gln = 36,2%
	(Shen <i>et al.</i> , 2005)	Arg/Arg = 40,0%; Arg/Gln = 48,3%	Gln = 35,9%
	(Figueiredo <i>et al.</i> , 2004)	Arg/Arg = 39,8%; Arg/Gln = 46,0%	Gln = 37,2%
Asiática	(Shu <i>et al.</i> , 2003)	Arg/Arg = 51,6% ; Arg/Gln = 42,1%	Gln = 27,3%
	(Kim <i>et al.</i> , 2002)	Arg/Arg = 43,9%; Arg/Gln = 49,3 %	Gln = 31,5%
Afro-Americana	(Duell <i>et al.</i> , 2001)	Arg/Arg = 74,0%; Arg/Gln = 24,0%	Gln = 13,5%
Indiana	(Chacko <i>et al.</i> , 2005)	Arg/Arg = 64,2% ; Arg/Gln = 28,5%	Gln = 21,5%
Turca	(Deligezer e Dalay, 2004)	Arg/Arg = 37,6%; Arg/Gln = 49,6%	Gln = 37,6%

Tabela 5.4 - Características gerais do grupo de doentes (n=241) e do grupo de controlos (n=457) estudados.

CARACTERÍSTICAS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a
Idade^{b,c}			
31 – 49	51 (21,2)	98 (21,4)	0,982
50 – 69	131 (54,4)	250 (54,7)	
≥ 70	59 (24,5)	109 (23,9)	
Hábitos Tabágicos			
Nunca e Ex – F.	208 (86,7)	417 (91,6)	0,038
Fumadores	32 (13,3)	38 (8,4)	
N.D.	1	2	
Hábitos Alcoólicos			
Nunca	186 (77,2)	376 (83,0)	<0,001
Social	19 (7,9)	52 (11,5)	
Regular	36 (14,9)	25 (5,5)	
N.D.	0	4	
XRCC1 (Arg194Trp)			
Arg/Arg	198 (82,5)	396 (87,4)	0,207
Arg/Trp	41 (17,1)	56 (12,4)	
Trp/Trp	1 (0,4)	1 (0,2)	
N.D.	1	4	
XRCC1 (Arg399Gln)			
Arg/Arg	112 (46,5)	191 (41,9)	0,504
Arg/Gln	104 (43,2)	212 (46,5)	
Gln/Gln	25(10,4)	53 (11,6)	
N.D.	0	1	

^a Ver Materiais e Métodos; ^b Idade de diagnóstico; ^c Idade do indivíduo controlo à data de diagnóstico do doente correspondente. N.D. – Não Determinado

Os resultados obtidos neste estudo não suportam uma associação entre a presença de qualquer genótipo gerado pelos polimorfismos presentes do gene *XRCC1* e a susceptibilidade individual para cancro da mama, uma vez que, e como já foi referido, as frequências genotípicas determinadas para ambas as populações não são significativamente diferentes (Tabela 5.4). Após análise por regressão logística, não foi

determinado um valor de OR significativo que afectasse os diferentes genótipos mesmo depois de ajustado aos possíveis factores confundentes, idade, hábitos tabágicos e alcoólicos (Tabela 5.5). No entanto, após estratificação de acordo com o *status* menopausico (pré-menopausa, idade de menopausa <44, 45 < idade de menopausa >54, e idade de menopausa ≥55), foi observado, para o polimorfismo Arg194Trp, que indivíduos com genótipo heterozigótico (Arg/Trp) com idade de menopausa compreendida entre os 45 e os 54 anos apresentam um aumento de risco para cancro da mama [OR_{Ajustado} = 1,964; IC a 95%, 1,174–3,288] (P = 0,01) e o mesmo efeito foi encontrado para os indivíduos portadores de pelo menos um alelo variante (Arg/Trp + Trp/Trp) [OR_{Ajustado} = 1,932; IC a 95%, 1,156–3,228] (P = 0,012) (ver Tabela 5.5). Para o polimorfismo Arg399Gln observou-se que mulheres com idade de menopausa superior a 55 anos e portadoras do genótipo homozigótico para a forma variante, apresentam igualmente um aumento de risco para esta neoplasia [OR_{Ajustado} = 4,074; IC a 95%, 1,562–10,626] (P = 0,004) (ver Tabela 5.6).

Tabela 5.5 – Valores de OR (IC 95%) para o polimorfismo *XRCC1* Arg194Trp e respectiva associação com cancro da mama.

	N	<i>XRCC1</i> (ARG194TRP)	OR_{NÃO-AJUSTADO} (IC 95%)	OR_{AJUSTADO} (IC 95%)^a
Todos os casos	240	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Trp	1,464 (0,945 – 2,268)	1,487 (0,953 – 2,321)
		Trp/Trp	2,000 (0,124 – 32,143)	1,657 (0,100 – 27,431)
		Arg/Trp + Trp/Trp	1,474 (0,955 – 2,273)	1,491 (0,959 – 2,317)
Status Menopausico	237			
Pré-Menopausa	19	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Trp	0,832 (0,187 – 3,697)	0,787 (0,168 – 3,681)
		Trp/Trp	0,047 (0,000 – 4,04E+17)	N.D.
		Arg/Trp + Trp/Trp	0,817 (0,184 – 3,631)	0,787 (0,168 – 3,681)
Pós-Menopausa Idade Menopausa < 44	218	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Trp	0,752 (0,287 – 1,972)	0,738 (0,278 – 1,959)
		Trp/Trp	8,425 (0,518 – 136,936)	5,171 (0,302 – 88,421)
		Arg/Trp + Trp/Trp	0,887 (0,363 – 2,168)	0,857 (0,346 – 2,121)
Entre 45 e 54	135	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Trp	1,953 (1,188 – 3,211)	1,964 (1,174 – 3,288)^c
		Trp/Trp	0,058 (0,000 – 5,56E+5)	0,025 (0,000 – 7,90E+9)
		Arg/Trp + Trp/Trp	1,919 (1,168 – 3,151)	1,932 (1,156 – 3,228)^d
≥ 55	30	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Trp	1,088 (0,366 – 3,233)	1,177 (0,388 – 3,571)
		Trp/Trp	0,031 (0,000 – 2,64E+17)	N.D.
		Arg/Trp + Trp/Trp	1,069 (0,360 – 3,175)	1,142 (0,377 – 3,459)

^a OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

^b Indivíduos com genótipo Arg/Arg considerados como grupo de referência.

^c P = 0,01

^d P = 0,012

Tabela 5.6 - Valores de OR (IC 95%) para o polimorfismo *XRCC1* Arg399Gln e respectiva associação com cancro da mama.

	N	<i>XRCC1</i> (ARG399GLN)	OR _{NÃO-AJUSTADO} (IC 95%)	OR _{AJUSTADO} (IC 95%) ^a
Todos os casos	240	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Gln	0,837 (0,601 – 1,165)	0,870 (0,620 – 1,222)
		Gln/Gln	0,805 (0,474 – 1,366)	0,873 (0,509 – 1,497)
		Arg/Gln + Gln/Gln	0,830 (0,606 – 1,137)	0,871 (0,631 – 1,201)
Status Menopausico Pré-Menopausa	237	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Gln	0,492 (0,178 – 1,355)	0,493 (0,171 – 1,423)
		Gln/Gln	0,655 (0,141 – 3,047)	0,520 (0,140 – 2,610)
		Arg/Gln + Gln/Gln	0,524 (0,207 – 1,328)	0,500 (0,188 – 1,324)
Pós-Menopausa Idade Menopausa < 44	218	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Gln	1,244 (0,686 – 2,255)	1,195 (0,654 – 2,186)
		Gln/Gln	0,515 (0,148 – 1,793)	0,492 (0,140 – 1,731)
		Arg/Gln + Gln/Gln	1,098 (0,614 – 1,964)	1,053 (0,584 – 1,898)
Entre 45 e 54	135	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Gln	0,780 (0,522 – 1,166)	0,813 (0,535 – 1,234)
		Gln/Gln	0,538 (0,259 – 1,117)	0,604 (0,285 – 1,278)
		Arg/Gln + Gln/Gln	0,732 (0,498 – 1,075)	0,722 (0,518 – 1,152)
≥ 55	30	Arg/Arg ^b	1 (Referência)	1 (Referência)
		Arg/Gln	0,737 (0,299 – 1,817)	0,803 (0,316 – 2,041)
		Gln/Gln	3,276 (1,320 – 8,129)	4,074 (1,562 – 10,626)^c
		Arg/Gln + Gln/Gln	1,245 (0,579 – 2,677)	1,394 (0,626 – 3,103)

^a OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

^b Indivíduos com genótipo Arg/Arg considerados como grupo de referência.

^c P = 0,004

Contudo, e uma vez que se observou com este estudo que a presença do alelo variante nos polimorfismos Arg194Trp e Arg399Gln parecia estar associada com um aumento de risco para cancro da mama em mulheres pós-menopausicas com idade de menopausa superior a 45 anos, posteriormente pretendeu avaliar-se o papel da associação dos dois polimorfismos na susceptibilidade para neoplasia mamária. Os resultados obtidos mostraram que a presença simultânea dos alelos variantes de ambos os SNPs não era frequente nas populações estudadas (Tabela 5.7), o que não permitiu este tipo de análise.

Tabela 5.7 – Combinação de genótipos dos polimorfismos do gene *XRCC1* estudados (Arg194Trp e Arg399Gln) na população Portuguesa.

<i>XRCC1</i> Arg194Trp	<i>XRCC1</i> Arg399Gln		
	Arg/Arg, n (%)	Arg/Gln, n (%)	Gln/Gln, n (%)
População Controlo			
Arg/Arg	156 (34,44)	187 (41,28)	53 (11,7)
Arg/Trp	33 (7,28)	23 (5,10)	0 (0)
Trp/Trp	1 (0,2)	0 (0)	0 (0)
População Doente			
Arg/Arg	83 (34,58)	90 (37,50)	25 (10,42)
Arg/Trp	27 (11,25)	14 (5,83)	0 (0)
Trp/Trp	1 (0,42)	0 (0)	0 (0)

5.4 DISCUSSÃO

Desde que se conhece o envolvimento do gene *XRCC1* na reparação de uma ampla variedade de lesões exógenas ou endógenas (aductos de baixo peso molecular) e quebras de cadeia simples, a função atribuída aos polimorfismos Arg194Trp e Arg399Gln tem sido muito estudada tendo em vista a avaliação do efeito das variantes alélicas nos diferentes níveis de lesão de DNA induzidos por cancerígenos/genotóxicos ambientais e também na susceptibilidade individual para diferentes tipos de cancro (para revisão consultar (Goode *et al.*, 2002; Hu *et al.*, 2005; Hung *et al.*, 2005)). Os resultados previamente publicados sobre o papel dos polimorfismos Arg194Trp e Arg399Gln do gene *XRCC1* na susceptibilidade individual para cancro da mama têm

sido contraditórios (Deligezer e Dalay, 2004; Duell *et al.*, 2001; Figueiredo *et al.*, 2004; Kim *et al.*, 2002; Moullan *et al.*, 2003; Shen *et al.*, 2005; Shu *et al.*, 2003; Smith *et al.*, 2003b). Os resultados apresentados no presente estudo não revelaram uma associação significativa entre os dois polimorfismos do gene *XRCC1* e a susceptibilidade para cancro da mama, o que está de acordo com diversos outros estudos desenvolvidos em diferentes populações: Americana (Bu *et al.*, 2006; Duell *et al.*, 2001; Han *et al.*, 2003; Shen *et al.*, 2005; Smith *et al.*, 2003a; Smith *et al.*, 2003b), Canadiana (Shen *et al.*, 2005), Francesa (Moullan *et al.*, 2003), Brasileira (Dufloth *et al.*, 2005), Turca (Deligezer e Dalay, 2004) e Chinesa (Shu *et al.*, 2003), e também em estudos de meta-análise (Zhang *et al.*, 2006). No entanto, outros estudos desenvolvidos na população Indiana (Chacko *et al.*, 2005) revelaram uma relevante associação entre a presença do alelo variante de ambos os polimorfismos e o risco para cancro da mama. Por outro lado, foi descrito para as populações Afro-Americana (Duell *et al.*, 2001) e Coreana (Kim *et al.*, 2002), que apenas a presença do alelo variante do SNP Arg399Gln estaria associado com o risco para a mesma neoplasia. As diferenças observadas nos diferentes estudos poderão estar relacionadas com as diferenças existentes entre as frequências alélicas determinadas nas diferentes populações, e/ou diferentes factores de risco característicos das diferentes populações.

Após estratificação dos nossos resultados de acordo com a idade de menopausa, observou-se que mulheres com idade de menopausa superior a 55 anos e homocigóticas para o alelo variante do SNP Arg399Gln apresentam um aumento de risco para cancro da mama, tendo sido descrito um efeito semelhante para mulheres com idade de menopausa entre os 45 e os 54 anos e homocigóticas para o alelo variante do SNP Arg194Trp. Estudos epidemiológicos revelaram que o número de ciclos menstruais aumenta o risco de cancro da mama (Geimba de Lima M. *et al.*, 2001). Tal evidência é igualmente suportada através de dados bioquímicos relacionados com o desempenho dos estrogénios no cancro da mama. De facto, os catecol-estrogénios podem actuar como moléculas de sinalização através da interacção com os receptores de estrogénios, mas também podem ser oxidados, originando quinonas que podem reagir com o DNA, conduzindo à formação de aductos, ou ser envolvidos em reacções redox induzindo a geração de espécies reactivas de oxigénio que podem provocar lesões oxidativas (Yager, 2000). Para além

disso, existem evidências crescentes que suportam um desempenho funcional dos polimorfismos Arg194Trp e Arg399Gln do gene *XRCC1*. Foi descrito que, em trabalhadores em radiologia, indivíduos portadores de pelo menos um alelo variante para o polimorfismo Arg399Gln apresentam um aumento significativo de micronúcleos (MN – do inglês *Micronuclei*) (Angelini *et al.*, 2005). Este polimorfismo está também associado com um aumento dos níveis de troca de cromátides irmãs (SCE – do inglês *Sister Chromatid Exchange*) em indivíduos fumadores (Lei *et al.*, 2002), com aumento dos níveis de lesão de DNA após exposição de linfócitos a compostos como a bleomicina e o benzo[*a*]pireno-diol-epóxido (Wang *et al.*, 2003), e ainda a um elevado número de aductos de DNA (Duell *et al.*, 2000; Matullo *et al.*, 2001b). Por outro lado, em linhas celulares EM9, linha celular utilizada como modelo celular *knockout* para *XRCC1*, observou-se que a expressão da forma mais comum, mas não para a forma variante, do polimorfismo Arg399Gln corrige a sensibilidade mutagénica observada nesta linha celular (Qu *et al.*, 2005). No entanto, este polimorfismo não parece estar associado com um aumento de expressão do gene *XRCC1* e com indução da apoptose após exposição a radiação ionizante ou ao benzo[*a*]pireno (Bu *et al.*, 2006). Os resultados disponíveis que consideram o polimorfismo Arg194Trp sugerem que, após exposição de linfócitos à bleomicina e ao benzo[*a*]pireno-diol-epóxido, os indivíduos homocigóticos para o alelo mais comum apresentam maiores índices de lesão de DNA do que a apresentada para os indivíduos homocigóticos para o alelo variante (Wang *et al.*, 2003).

Se se considerarem os resultados publicados por Qu *et al.* (Qu *et al.*, 2005), uma vez que os alelos variantes podem não apresentar capacidade de reparação adequada a lesões de DNA provocadas por erros espontâneos na proliferação das células epiteliais da mama, por aductos de DNA, e/ou radicais livres gerados por compostos de estrogénio activos, pode ser considerada uma explicação para a associação entre os polimorfismos do gene *XRCC1* e a idade tardia de menopausa.

Diversos estudos têm demonstrado que algumas proteínas polimórficas interagem com o gene *XRCC1* (Caldecott, 2003; Ronen e Glickman, 2001), gene este que apresenta vários outros polimorfismos (Chang-Claude *et al.*, 2005; Duell *et al.*, 2001; Moullan *et al.*, 2003; Smedby *et al.*, 2006; Zhang *et al.*, 2006). Assim, não pode ser excluído que outros polimorfismos presentes neste gene e/ou em outros genes

igualmente envolvidos na via BER podem, sozinhos ou em associação, estar relacionados com a susceptibilidade individual para cancro da mama.

Sumariamente, os resultados aqui apresentados sugerem que os polimorfismos presentes no gene *XRCC1* Arg194Trp e Arg399Gln, podem estar envolvidos na susceptibilidade individual para neoplasia mamária, contudo estudos adicionais são necessários, efectuados com populações com maior número de indivíduos, de forma a confirmarem-se estes resultados.

De salientar três estudos de meta-análise onde foram incluídos os resultados obtidos com este trabalho (Huang *et al.*, 2009; Li *et al.*, 2009; Saadat e Ansari-Lari, 2009).

CAPÍTULO 6

O PAPEL DE POLIMORFISMOS DO GENE *ERCC2* NO RISCO PARA CANCRO DA MAMA

6. O PAPEL DE POLIMORFISMOS DO GENE *ERCC2* NO RISCO PARA CÂNCRO DA MAMA

Publicado em parte em Silva *et al.*, 2006, *Cancer Genet Cytogenet.*, 170, 86-88

6.1 INTRODUÇÃO

Estudos recentes revelaram que algumas funções associadas à reparação de DNA são haploinsuficientes o que dá ênfase ao facto de que variações em genes envolvidos na reparação de DNA constituem parte de um espectro de variações que incluirão um fenótipo de reparação que pode contribuir para o risco de desenvolvimento de cancro. Esta variação na capacidade de reparação tem características expectáveis associada a genes de susceptibilidade para o cancro, proteínas codificadas por estes alelos podem apresentar uma função reduzida ou mesmo a ausência de função, o que será responsável pela doença. Encontram-se descritos alguns polimorfismos em genes que codificam para proteínas envolvidas na reparação de DNA, sendo que a grande maioria participa nas quatro maiores vias de reparação: reparação por excisão de bases (BER), reparação por excisão de nucleótidos (NER), reparação *mismatch* (MMR) e reparação de quebras e cadeia dupla/reparação por recombinação (DSBR) (Ronen e Glickman, 2001; Ruttan e Glickman, 2002). Alguns estudos têm apresentado como principal objectivo a identificação de variantes polimórficas em genes envolvidos na reparação de DNA e potencialmente associados com cancro da mama, sugerindo que variantes nos genes *XRCC1*, *XRCC2*, *XRCC4*, *LIG4*, *RAD52*, *ERCC1* e *BRCA2* (Goode *et al.*, 2002; Kim *et al.*, 2002; Kuschel *et al.*, 2002; Lee *et al.*, 2005b; Rafii *et al.*, 2002) podem estar associados com susceptibilidade individual para esta neoplasia.

A via NER é a maior via celular de reparação envolvida na reparação de uma grande variedade de lesões de DNA (fotoprodutos induzidos por UV; aductos químicos de elevado peso molecular e também lesão oxidativa) e está ainda associada a uma via de reparação relacionada com a transcrição (Friedberg, 2001). A via NER envolve a acção de cerca de 25 proteínas responsáveis pela localização da cadeia lesada, introdução de incisões que rodeiem a lesão, excisão de um oligonucleotido com 24-32 resíduos, e

preenchimento da região excisada por síntese e ligação das extremidades. O gene *ERCC2* (*XPD*) codifica para uma DNA helicase, subunidade do factor de transcrição basal TFIIH, necessário para a iniciação da transcrição pela RNA polimerase II, sendo também uma das principais proteínas envolvidas na via NER (Ronen e Glickman, 2001). Mutações presentes na linha germinal da região codificante do gene *ERCC2* estão correlacionadas com diferentes fenótipos clínicos tais como *Xeroderma pigmentosum*, tricotiodistrofia e síndrome de Cockayne (Lehmann, 2001).

À data da publicação deste trabalho tinham já sido identificados diversos polimorfismos não-sinónimos localizados em regiões codificantes. Os polimorfismos Ile199Met, His201Tyr, Asp312Asn e Lys751Gln do gene *ERCC2* dão origem a alteração de aminoácido. No entanto, não existem evidências bioquímicas directas sobre as suas consequências funcionais, para além de que os resultados publicados são divergentes quanto à função dos alelos variantes na reparação de DNA (Hemminki *et al.*, 2001) e na susceptibilidade para cancro da mama (Justenhoven *et al.*, 2004; Kuschel *et al.*, 2005; Lee *et al.*, 2005c; Zhang *et al.*, 2005).

Uma vez que os resultados sobre os polimorfismos do gene *ERCC2* têm sido contraditórios, e, por outro lado, a exposição a cancerígenos ambientais (*e.g.* PAH e aminas heterocíclicas) podem representar um factor de risco para cancro da mama (Gorlewska-Roberts *et al.*, 2002), e sendo as lesões normalmente provocadas por estes compostos reparadas por enzimas envolvidas na via NER, foi desenvolvido um estudo caso-controlo na população Portuguesa de forma a avaliar o potencial efeito dos polimorfismos neste gene na susceptibilidade individual para cancro da mama.

6.2 MATERIAIS E MÉTODOS

6.2.1 DESCRIÇÃO DA POPULAÇÃO

A descrição efectuada no Capítulo 4, ponto 4.2.1 aplica-se neste novo capítulo.

6.2.2 EXTRACÇÃO DE DNA

A metodologia utilizada para extracção de DNA encontra-se descrita no Capítulo 3, ponto 3.2.2.

6.2.3 GENOTIPAGEM DOS POLIMORFISMOS PRESENTES NO GENE *ERCC2* (ILE199MET), (HIS201TYR), (ASP312ASN) E (LYS751GLN)

A genotipagem dos polimorfismos estudados foi efectuada por PCR-RFLP, técnica detalhada no Capítulo 3, ponto 3.2.3. Os *primers* utilizados nas reacções de PCR bem como a temperatura de *annealing* utilizada para a amplificação de cada SNP, e os produtos obtidos especificamente após RFLP estão descritos na Tabela 6.1.

O PCR foi efectuado utilizando cerca de 50 ng de DNA num volume reaccional de 50 µl. Para os polimorfismos *ERCC2* Ile199Met e His201Tyr a mistura reaccional continha: tampão PCR 1X, 0,2 mM de cada dNTP, 1 µM de cada *primer*, 2,5 mM de MgCl₂ e 1,25 U de *AmpliTaq Gold* (Applied Biosystems). De notar que os SNPs Ile199Met e His201Tyr foram incluídos no mesmo amplicão, ou seja o fragmento de amplificação inclui ambos os polimorfismos.

As condições de amplificação foram iguais para ambos os SNP e consistiram em 7 minutos de activação inicial a 95 °C, seguido de 35 ciclos de desnaturação a 94 °C durante 30 segundos, *annealing* a 62 °C durante 30 segundos, extensão a 72 °C durante 30 segundos e 10 minutos de extensão final a 72 °C. Após amplificação 10 µl do produto de PCR resultante foi digerido, segundo instruções do fabricante, com a enzima de restrição adequada (ver Tabela 6.1) e posteriormente aplicados em gel de agarose com brometo de etídio (1,0 µg/ml) para posterior visualização. Os produtos de restrição expectáveis para os SNP estudados estão apresentados na Tabela 6.1.

As condições de PCR-RFLP utilizadas para os outros dois polimorfismos estudados (*ERCC2* Asp312Asn e Lys751Gln) foram efectuadas como descrito previamente (Silva *et al.*, 2005), estando sumarizados igualmente na Tabela 6.1.

As experiências foram realizadas em duplicado em ensaios independentes, sendo os resultados inconclusivos reanalisados.

6.2.4 ANÁLISE ESTATÍSTICA

A análise das frequências de Hardy-Weinberg dos alelos de todos os SNPs estudados, para a população de doentes, de controlos e para a população constituída pelos Índios Xavante, foi realizada através de testes de probabilidade exacta disponíveis no software Mendel (versão 5.7.2) (Lange *et al.*, 2001).

A abordagem estatística efectuada encontra-se descrita no Capítulo 3, ponto 3.2.4.

Tabela 6.1 - Condições de PCR-RFLP para os polimorfismos do gene *ERCC2* estudados.

SNPs	SEQUÊNCIA DOS <i>PRIMERS</i>	FRAGMENTO PCR (bp)	T _{ANNEALING} (°C)	EFEITO DA ENZIMA DE RESTRIÇÃO	PERFIS APÓS ANÁLISE DE RESTRIÇÃO
Ile199Met	<i>Forward:</i> 5'-TGT CTC TAT CCA TCT GCT CAT ACT TCT GCC- 3'	195	62	C → G, perde o local de reconhecimento para <i>Bsp143I</i>	CC: 144, 51bp; CG: 195, 144, 51bp; GG: 195bp.
His201Tyr	<i>Reverse:</i> 5'-GGG GTC CAG GAG GTA GTG GTA G- 3'			C → T, perde o local de reconhecimento para <i>PaeI</i>	CC: 155, 40bp; CT: 195, 155, 40bp; TT: 195bp.
Asp312Asn	<i>Forward:</i> 5'-GCA TCG TGC GTA AGG AGT- 3' <i>Reverse:</i> 5'-CCG CTC CTC TCA GTA GTC TGC- 3'	751	60	G → A, ganha dois locais de reconhecimento para <i>Eco130I</i>	GG: 506, 245bp; GA: 32, 245, 474, 506bp; AA: 32, 245, 474bp.
Lys751Gln	<i>Forward:</i> 5'-CCC TCT CCC TTT CCT CTG TTC TCT GC- 3' <i>Reverse:</i> 5'-GGA ACA GTG CAG GAG GGA TGG G - 3'	359	62	A → C, ganha dois locais de reconhecimento para <i>PstI</i>	AA: 331, 28bp; AC: 331, 268, 63, 28bp; CC: 268, 63, 28bp.

6.3 RESULTADOS

As frequências dos polimorfismos Ile199Met, His201Tyr, Asp312Asn e Lys751Gln presentes no gene *ERCC2* foram determinadas para cada uma das populações descritas no ponto 6.2.1. Na Tabela 6.2 estão descritas as principais características das populações de doentes e controlo. Os resultados obtidos não revelaram diferenças significativas entre as idades e cada um dos polimorfismos estudados. No entanto, após estratificação por hábitos alcoólicos e tabágicos verificou-se que consumidores regulares de álcool e fumadores se encontram sob representados na população de doentes quando comparado com a população controlo (ver Tabela 6.2).

Para os polimorfismos Ile199Met e His201Tyr, os resultados obtidos revelaram que a frequência do alelo mais comum em ambos os SNPs na população Caucasiana Portuguesa (doentes e controlos) e também na população Xavante é de 100%, o que sugere que a prevalência destes polimorfismos nestas populações é muito baixa. Os resultados obtidos para os outros dois polimorfismos em estudo, Asp312Asn e Lys751Gln, revelaram que a frequência alélica para o alelo mais comum para cada um dos SNPs é $0,6847 \pm 0,0155$ e $0,6400 \pm 0,0159$, respectivamente. As frequências genotípicas descritas na Tabela 6.2 para a população controlo estão de acordo com resultados previamente publicados em outras populações Caucasianas (Benhamou e Sarasin, 2005; Silva *et al.*, 2005).

As frequências genotípicas determinadas para os SNPs Asp312Asn e Lys751Gln nas populações de doentes e controlos respeitam o equilíbrio de Hardy-Weinberg. O teste para o desequilíbrio de *linkage* genético efectuado entre os dois *loci* foi altamente significativo ($P < 0,0001$) o que indica que os dois SNPs estão fortemente correlacionados, os alelos Asp312 e Lys751 e os alelos Asn312 e Gln751, na nossa população controlo ($D' = 0,802444$; $P < 0,0001$) e na população de doentes ($D' = 0,69279$; $P < 0,0001$), como descrito previamente para outras populações (Butkiewicz *et al.*, 2001; Liang *et al.*, 2003).

Tabela 6.2 - Características gerais do grupo de doentes (n=241) e do grupo de controlos (n=452) estudados.

CARACTERÍSTICAS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a
Idade			
30 – 39	9 (3,7)	20 (4,4)	0,994
40 – 49	42 (17,4)	78 (17,1)	
50 – 59	56 (23,2)	109 (23,9)	
60 – 69	75 (31,1)	141 (30,9)	
≥ 70	59 (24,5)	109 (23,9)	
Hábitos Tabágicos			
Nunca e Ex – F.	208 (86,7)	417 (91,6)	0,038
Fumadores	32 (13,3)	38 (8,4)	
N.D.	1	2	
Hábitos Alcoólicos			
Nunca	186 (77,2)	376 (83,0)	<0,001
Social	19 (7,9)	52 (11,5)	
Regular	9 (7,0)	25 (5,5)	
N.D.	4	0	
ERCC2 (Ile199Met)			
Ile/Ile	241 (100)	456 (100)	-
Ile/Val	0	0	
Val/Val	0	0	
ERCC2 (His201Tyr)			
His/His	241 (100)	456 (100)	
His/Tyr	0	0	
Tyr/Tyr	0	0	
ERCC2 (Asp312Asn)			
Asp/Asp	121 (50,2)	218 (48,2)	0,465
Asp/Asn	100 (41,5)	183 (40,5)	
Asn/Asn	20 (8,3)	51 (11,3)	
ERCC2 (Lys751Gln)			
Lys/Lys	113 (46,9)	195 (42,7)	0,467
Lys/Gln	99 (41,1)	195 (42,7)	
Gln/Gln	29 (12,0)	67 (14,6)	

^a Ver Materiais e Métodos. N.D. – Não Determinado

Os resultados obtidos para a frequência dos polimorfismos em estudo na população constituída pelos Índios Xavante, caracterizada pela total inexistência de casos de cancro da mama (ver ponto 6.2.1), revelaram que as frequências destes polimorfismos são similares às descritas para populações de Chineses e Coreanos mas diferentes das descritas na população Caucasiana ($0,8182 \pm 0,0206$ e $0,8212 \pm 0,0203$ para os polimorfismos Asp312Asn e Lys751Gln, respectivamente) (Benhamou e Sarasin, 2005). Na população de Índios Xavante os polimorfismos Asp312Asn e Lys751Gln encontram-se em equilíbrio de Hardy-Weinberg ($P > 0,1$) e em desequilíbrio de *linkage* absoluto ($D' = 1$; $P < 0,0001$).

Os resultados obtidos para os polimorfismos Asp312Asn e Lys751Gln não revelaram uma associação entre a presença de um genótipo destes SNPs e a susceptibilidade individual para cancro da mama, uma vez que as frequências genotípicas tanto na população de doentes como na população controlo não são significativamente diferentes (Tabela 6.2). Mesmo após regressão logística não se observaram valores de OR significativos que justifiquem o efeito dos diferentes genótipos, após ajuste a factores como a idade, hábitos tabágicos e alcoólicos (Tabela 6.3). Contudo, uma vez que se observou que a presença de alelos variantes em ambos os SNPs (Asp312Asn e Lys751Gln) conduzia a uma diminuição nos valores de OR, fomos avaliar o papel associado com a presença de alelos variantes na susceptibilidade para cancro da mama. Os resultados obtidos mostraram que a presença de alelos variantes, em ambos os polimorfismos, não está associada com susceptibilidade individual para cancro da mama (Tabela 6.3).

Uma vez que ambos os polimorfismos estão em desequilíbrio de *linkage* determinou-se a frequência derivada da combinação dos diferentes genótipos em ambas as populações (ver Tabela 6.4).

Tabela 6.3 - Valores de OR (IC 95%) para os polimorfismos *ERCC2* Asp312Asn e Lys751Gln e respectiva associação com cancro da mama.

GENÓTIPOS	OR _{NÃO-AJUSTADO} (IC 95%)	OR _{AJUSTADO} (IC 95%) ^a
Asp312Asn (G/A)		
Asp/Asp	1 (Referência)	1 (Referência)
Asp/Asn	0,985 (0,708 – 1,369)	0,991 (0,706 – 1,391)
Asn/Asn	0,707 (0,402 – 1,240)	0,685 (0,385 – 1,218)
Asp/Asn + Asn/Asn	0,924 (0,676 – 1,263)	0,923 (0,669 – 1,272)
Lys751Gln (A/C)		
Lys/Lys	1 (Referência)	1 (Referência)
Lys/Gln	0,876 (0,627 – 1,225)	0,825 (0,584 – 1,164)
Gln/Gln	0,747 (0,456 – 1,223)	0,778 (0,440 – 1,268)
Lys/Gln + Gln/Gln	0,843 (0,616 – 1,154)	0,813 (0,589 – 1,122)

^a OR's foram ajustados para: idade de diagnóstico (30-39, 40-49, 50-59, 60-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

Tabela 6.4 – Combinação dos genótipos obtidos para os polimorfismos Asp312Asn e Lys751Gln do gene *ERCC2* nas populações estudadas.

Lys751Gln \ Asp312Asn	Asp/Asp	Asp/Asn	Asn/Asn
	População Controlo (n=452)		
Lys/Lys	172 (38,1%)	19 (4,2%)	9 (2,0%)
Lys/Gln	37 (8,2%)	142 (31,4%)	13 (2,9%)
Gln/Gln	9 (2,0%)	22 (4,9%)	36 (8,0%)
População Doente (n=241)			
Lys/Lys	91 (37,7%)	19 (7,9%)	3 (1,2%)
Lys/Gln	26 (10,8%)	69 (28,6%)	4 (1,7%)
Gln/Gln	4 (1,7%)	12 (5,0%)	13 (5,4%)
População de Índios Xavante (n=176)			
Lys/Lys	120 (68,2%)	0	0
Lys/Gln	0	48 (27,3%)	0
Gln/Gln	0	0	7 (3,9%)

A análise por regressão logística da combinação dos diferentes genótipos mostrou um significativo aumento de risco apenas em portadores dos genótipos Asp/Asn + Lys/Lys [OR_{Ajustado} = 2,136 (1,045 – 4,369)], para todas as outras combinações não foram encontradas associações (Tabela 6.5).

Tabela 6.5 - Valores de OR (IC 95%) obtidos por regressão logística para as combinações de genótipos dos polimorfismos *ERCC2* Asp312Asn e Lys751Gln e associação com cancro da mama.

GENÓTIPOS	OR _{NÃO-AJUSTADO} (IC 95%)	OR _{AJUSTADO} (IC 95%) ^a
Asp/Asp + Lys/Lys	1 (Referência)	1 (Referência)
Asp/Asp + Lys/Gln	1,328 (0,757 – 2,330)	1,244 (0,697 – 2,222)
Asp/Asp + Gln/Gln	0,840 (0,252 – 2,803)	0,910 (0,271 – 3,057)
Asp/Asn + Lys/Lys	1,890 (0,953 – 3,749)	2,136 (1,045 – 4,369)^p
Asp/Asn + Lys/Gln	0,918 (0,626 – 1,348)	0,880 (0,593 – 1,305)
Asp/Asn + Gln/Gln	1,031 (0,488 – 2,178)	1,151 (0,536 – 2,471)
Asn/Asn + Lys/Lys	2,835 (0,465 – 17,275)	2,267 (0,359 – 14,302)
Asn/Asn + Lys/Gln	0,582 (0,184 – 1,835)	0,534 (0,165 – 1,728)
Asn/Asn + Gln/Gln	0,683 (0,345 – 1,352)	0,679 (0,339 – 1,360)

^a OR's foram ajustados para: idade de diagnóstico (30-39, 40 -49, 50-59, 60-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

^p P=0,038

6.4 DISCUSSÃO

O gene *ERCC2* encontra-se envolvido na reparação de aductos de elevado peso molecular e de lesões induzidas por radiação UV, deste modo o estudo do papel desempenhado pelos polimorfismos presentes neste gene (entre os quais se salientam Ile199Met, His201Tyr, Asp312Asn e Lys751Gln) tem sido desenvolvido com o intuito de se avaliar o efeito das variantes alélicas nos diferentes tipos de cancro. Apesar dos resultados contraditórios observados nos estudos desenvolvidos em cancro do pulmão, um estudo de meta-análise sugere que a presença dos alelos variantes em homozigotia dos polimorfismos Asp312Asn e Lys751Gln do gene *ERCC2* está associada com o aumento de susceptibilidade para cancro do pulmão (Hu *et al.*, 2004) e também

para cancro da tiróide (Silva *et al.*, 2005). No entanto, não se observou o mesmo efeito em outros tipos de tumor normalmente associado com o consumo de tabaco como por exemplo o cancro da bexiga (Matullo *et al.*, 2001a; Schabath *et al.*, 2005; Shen *et al.*, 2003; Stern *et al.*, 2002), e o cancro da cabeça e pescoço (Sturgis *et al.*, 2000). O efeito destes polimorfismos do gene *ERCC2* foi também descrito em estudos de cancro da mama embora com resultados contraditórios. Os estudos desenvolvidos nas populações Coreana (Lee *et al.*, 2005c) e Caucasiana (Kuschel *et al.*, 2005; Shi *et al.*, 2004; Tang *et al.*, 2002) não revelaram qualquer efeito relacionado com ambos os polimorfismos.

Contudo, resultados obtidos para uma população Chinesa (Zhang *et al.*, 2005) sugerem que a presença do genótipo homozigótico para a variante do SNP Asp312Asn (Asn/Asn) pode estar associado com uma diminuição de risco para cancro da mama [OR = 0,51 (0,27 – 0,94); P<0,05]. Outro estudo desenvolvido numa população Caucasiana Alemã revelou que indivíduos homozigóticos para o alelo mais comum no mesmo polimorfismo (Asp/Asp) apresentam um significativo aumento de risco para a mesma patologia [OR = 2,06 (1,39 – 3,07)] sendo este risco ainda maior em mulheres igualmente portadoras do genótipo Gln751Gln (Justenhoven *et al.*, 2004), o que está parcialmente de acordo com o descrito por Terry *et al.* (Terry *et al.*, 2004), que mostrou que os portadores de pelo menos um alelo variante do polimorfismo Lys751Gln apresentam maior risco para cancro da mama.

Os resultados obtidos neste trabalho mostram que os polimorfismos Ile199Met e His201Tyr do gene *ERCC2*, localizados no exão 8 do gene, são muito raros na população Portuguesa. Estes resultados são concordantes com as frequências descritas numa população Americana branca não-hispânica (Sturgis *et al.*, 2002b) de 99 e 100% para o alelo mais comum de ambos os SNPs, respectivamente. Os resultados disponíveis na base de dados *Cancer Genome Anatomy Project SNP500 Cancer Database*⁵ são igualmente concordantes. O efeito destes polimorfismos tendo em conta o risco individual para patologia neoplásica foi também estudado para o carcinoma espinocelular da cabeça e pescoço (Sturgis *et al.*, 2002a) e gliomas (Caggana *et al.*, 2001). Igualmente nestes estudos a baixa frequência dos alelos variantes não

⁵ (<http://snp500cancer.nci.nih.gov/>)

permitiu encontrar qualquer associação entre uma das formas alélicas e o risco para cancro. Uma vez que a frequência observada para estes polimorfismos é a mesma em ambas as populações (Doentes e Controlos), e também não foram detectados na população constituída pelos Índios Xavante, pode ser excluída qualquer associação entre estes dois SNPs (Ile199Met e His201Tyr) e a susceptibilidade para cancro da mama.

Relativamente aos outros dois polimorfismos estudados (Asp312Asn e Lys751Gln), os resultados obtidos não suportam qualquer associação entre os SNPs e a susceptibilidade individual para cancro da mama, o que está de acordo com os resultados publicados por outros autores, Lee *et al.* (Lee *et al.*, 2005c), Kuschel *et al.* (Kuschel *et al.*, 2005), Shi *et al.* (Shi *et al.*, 2004) e Tang *et al.* (Tang *et al.*, 2002), nas populações Coreana e Caucasiana. Contudo, considerando o efeito conjunto destes dois polimorfismos observou-se que indivíduos com os genótipos Asp/Asn + Lys/Lys pareciam aumentar o risco para esta neoplasia originando um $OR_{Ajustado} = 2,136$ (1,045 – 4,369) ($P=0,038$) (ver Tabela 6.5). Estes resultados podem, contudo, estar associados com um erro estatístico tipo I, uma vez que o número de indivíduos presentes em ambos os grupos (população de doentes e controlos) é baixo ($n = 19$ para cada grupo).

Justenhoven *et al.* (Justenhoven *et al.*, 2004) identificaram o genótipo Asp312Asp como sendo de alto risco, risco este que é ainda maior em mulheres que sejam igualmente portadoras do genótipo Gln751Gln. Outro estudo identificou a presença de pelo menos um alelo variante do polimorfismo Lys751Gln como estando associado com aumento de risco para cancro da mama (Terry *et al.*, 2004), contudo num estudo desenvolvido numa população Chinesa o genótipo Asp312Asp foi identificado como associado com a diminuição de risco (Zhang *et al.*, 2005).

Há fortes evidências de que estes SNPs se encontram em Desequilíbrio de *Linkage* em diferentes populações (Butkiewicz *et al.*, 2001; Justenhoven *et al.*, 2004; Liang *et al.*, 2003; Silva *et al.*, 2005) o que foi igualmente observado nas populações envolvidas neste estudo, resultando num baixo número de indivíduos com as combinações de genótipo de alto risco identificadas por Justenhoven *et al.* (Asp/Asp + Gln/Gln): 2,0 e 1,7% para as populações de doentes e controlos envolvidas neste trabalho e 6,0 e 2,0% segundo dados publicados por Justenhoven *et al.* (Justenhoven *et al.*, 2004). Um facto interessante é que esta combinação de genótipos de alto risco

está totalmente ausente na população constituída pelos índios Xavante, já que se encontram em desequilíbrio de *linkage* absoluto ($D'=1$), e a que a frequência de pelo menos um alelo variante do SNP Asp312Asp é muito mais baixa nos Índios Xavante do que a reportada na população americana (Terry *et al.*, 2004) e associada com aumento de risco (32% *versus* 59%). Tendo em conta estes resultados, a inconsistência de resultados observada nos diferentes estudos que envolveram o estudo simultâneo de ambos os polimorfismos pode ser explicada pelas diferenças nas frequências alélicas e/ou com a extensão do desequilíbrio de *linkage* da populações.

Para além dos factores relacionados com a história reprodutiva e/ou *status* hormonal, a dieta e exposição ocupacional são igualmente vistos como importantes factores etiológicos para o cancro da mama. O espectro de mutações do gene *TP53* encontrado em tumores da mama fornece possíveis pistas sobre a etiologia da doença sugerindo que outros cancerígenos exógenos poderão estar envolvidos na doença (Phillips *et al.*, 2002). De facto, em amostras de DNA isoladas de células epiteliais ductais exfoliadas de leite materno e de tecido mamário normal foram detectados aductos de compostos cancerígenos, tais como aminas aromáticas heterocíclicas, benzo[*a*]pireno e hidrocarbonetos aromáticos policíclicos. Estes dados sugerem que as mulheres estão expostas a diversas classes de cancerígenos dietéticos e ambientais e que esses compostos podem reagir com o DNA em células ductais epiteliais da mama (Gorlewska-Roberts *et al.*, 2002). Considerando os polimorfismos Asp312Asn e Lys751Gln, a maioria dos dados reportados, com algumas excepções, revelam níveis mais elevados de aductos de DNA em indivíduos portadores dos genótipos Asn e Gln do que em portadores Asp e Lys, o que é interpretado como uma eficiência de reparação reduzida para os portadores dos alelos Asn e Gln (Benhamou e Sarasin, 2005). Um efeito semelhante foi descrito considerando a capacidade de reparar lesões induzidas por radiação UV (Qiao *et al.*, 2002). Os dados disponíveis que consideram o papel dos polimorfismos no gene *ERCC2* e os níveis de aductos derivados de hidrocarbonetos aromáticos policíclicos no tecido mamário, sugerem que portadores dos alelos variantes apresentam níveis mais elevados de aductos de DNA quando comparados com os indivíduos portadores dos alelos mais frequentes (Tang *et al.*, 2002).

Considerando os resultados obtidos neste trabalho, seria razoável assumir o envolvimento dos polimorfismos Asp312Asn e Lys751Gln do gene *ERCC2* na susceptibilidade para cancro da mama, no entanto os resultados não suportam tal hipótese pelo número reduzido resultante da estratificação de fenótipos associados com OR significativo (n=19). Para além dos potenciais factores genéticos envolvidos e discutidos anteriormente, diferentes tipos de exposição a cancerígenos ambientais (e.g fumo de tabaco, cancerígenos de origem alimentar), podem ser diferentes entre as populações conduzindo a resultados inconsistentes sobre o papel destes polimorfismos na neoplasia mamária.

Tem sido demonstrado que diversas proteínas, tal como o p53, podem modular a via NER através de interações proteína-proteína com *XPB* e *XPD*, e também a existência de outros polimorfismos no gene *ERCC2*. Assim, não se pode excluir que outros polimorfismos neste gene e/ou em outros genes associados com a regulação da via NER possam, sozinhos ou em associação, estar relacionados com a susceptibilidade para cancro da mama.

CAPÍTULO 7

ASSOCIAÇÃO DE VARIANTES EM GENES ENVOLVIDOS
NA VIA DE REPARAÇÃO *MISMATCH* E
SUSCEPTIBILIDADE PARA CANCRO DA MAMA:
ESTUDO MULTIGÉNICO

7. ASSOCIAÇÃO DE VARIANTES EM GENES ENVOLVIDOS DA VIA DE REPARAÇÃO *MISMATCH* E SUSCEPTIBILIDADE PARA CANCRO DA MAMA: ESTUDO MULTIGÉNICO

Publicado em Conde *et al.*, 2009, BMC Cancer, 9:344

7.1 INTRODUÇÃO

Diversos factores genéticos e ambientais têm sido identificados como estando associados com a susceptibilidade individual para cancro da mama. Assim, o desafio de se identificarem indivíduos de risco para os casos mais prevalentes, ou seja, os casos esporádicos, tem vindo a aumentar.

A ocorrência frequente de polimorfismos nos genes de reparação de DNA tem sido relacionada com um aumento de risco para cancro devido à sua função crítica na manutenção da integridade do genoma (Ford *et al.*, 2000).

Múltiplos dados indicam que a grande maioria dos diferentes tipos de cancro mostram instabilidade em sequências específicas constituídas por repetições dinucleotídicas. Este fenótipo denominado de instabilidade de microsatélites (MSI) é normalmente observado em defeitos na via de reparação de *mismatch* (MMR) (Hoeijmakers, 2001). De facto, a instabilidade de microsatélites e a perda de heterozigotia (LOH) foram as alterações detectadas em cerca de 83% de amostras de pele provenientes de doentes com carcinoma ductal invasivo da mama, o que sugere o potencial envolvimento da reparação de *mismatch* na susceptibilidade para cancro da mama (Moinfar *et al.*, 2008).

A reparação de *mismatch* pós-replicativa (MMR), conservada dos procariotas a todos os eucariotas, incluindo o Homem, actua na substituição de bases *mismatches* e em *loops* de inserção/delecção que ocorrem como resultado de erros de replicação que tenham escapado à função correctiva da DNA polimerase (Kunkel e Erie, 2005; Li, 2008). A via MMR contribui grandemente para fidelidade de replicação. Deste modo, uma actividade da MMR diminuída confere um fenótipo mutante pelo qual a taxa de mutação espontânea está muito aumentada. Uma característica das células deficientes

em MMR é a instabilidade em regiões microsatélites caracterizadas por regiões de repetições de mono- e dinucleotidos. A marca mais comum da perda de actividade MMR em células tumorais é a existência de regiões MSI (Couch *et al.*, 2008).

A perda das funções MMR conduz a elevadas taxas de mutação, MSI, LOH, diminuição dos processos de apoptose e aumento da sobrevivência celular, bem como predisposição para a cancerigénese (Schofield e Hsieh, 2003; Schroering *et al.*, 2007). Esta via de reparação tem ainda a capacidade de suprimir a recombinação homóloga e desempenhar importantes funções na sinalização da lesão de DNA (Smith *et al.*, 2008).

O gene *MSH2* desempenha uma função central no reconhecimento da lesão *mismatch*, existindo diversos estudos que reportam mutações (Murata *et al.*, 2002) e polimorfismos em diversas variantes de *MSH2* (Poplawski *et al.*, 2005; Wong *et al.*, 2008). No entanto, e devido à falta de dados sobre o envolvimento de polimorfismos noutros genes da via MMR na susceptibilidade para cancro da mama, desenvolvemos um estudo caso-controlo numa população Caucasiana Portuguesa de forma a avaliar o potencial papel de polimorfismos nos genes *MSH3*, *MSH4*, *MSH6*, *MLH1*, *MLH3*, *PMS1* e *MUTYH* na susceptibilidade individual para cancro da mama.

7.2 MATERIAIS E MÉTODOS

7.2.1 DESCRIÇÃO DA POPULAÇÃO

A descrição efectuada no Capítulo 3, ponto 3.2.1 aplica-se neste novo capítulo, à excepção do tamanho da amostra, já que para este estudo foram utilizados 287 doentes com cancro da mama e 547 indivíduos que constituem a população controlo. O diagnóstico histológico foi confirmado em todos os casos e a amostra inclui 251 carcinomas tipo ductal (87,4%), 14 carcinomas tipo lobular (4,9%) e 22 casos classificados como sendo de outro tipo de tumor da mama (7,7%).

7.2.2 EXTRACÇÃO DE DNA

A metodologia utilizada encontra-se descrita no Capítulo 3, ponto 3.2.2.

7.2.3 GENOTIPAGEM DOS POLIMORFISMOS

Foram seleccionados os seguintes polimorfismos: *MSH3* (A1045T, A>G, rs26279; R940Q, G>A, rs184967), *MSH4* (N914S, G>A, rs5745549; A97T, A>G, rs5745325), *MSH6* (G39E, C>T, rs1042821), *MLH1* (I219V, A>G, rs1799977), *MLH3* (L844P, G>A, rs175080), and *MUTYH* (H335Q, G>C, rs3219489). Todos os SNP seleccionados são não-sinónimos à excepção de um SNP presente no gene *PMS1* (G>C, rs5742933), uma transição de G-para-C na região 5' UTR. Todos os polimorfismos apresentavam um MAF superior a 5%.

Todos os polimorfismos [*MSH3* (rs26279; rs184967), *MSH4* (rs5745549; rs5745325), *MSH6* (rs1042821), *MLH1* (rs1799977), *MLH3* (rs175080), *PMS1* 5'UTR (rs5742933) e *MUTYH* (rs3219489)] foram genotipados por PCR em tempo real (AB7300) utilizando os ensaios da Applied Biosystems *TaqMan*[®] *SNP Genotyping Assays* (as referências dos ensaios são C_800002_1_, C_907914_10, C_1184803_10, C_3286081_10, C_8760558_10, C_1219076_20, C_1082805_10, C_29329633_10 e C_27504565_10, respectivamente para cada SNP) segundo instruções fornecidas pelo fabricante.

7.2.4 ANÁLISE ESTATÍSTICA

A abordagem estatística efectuada encontra-se descrita no Capítulo 3, ponto 3.2.4.

Neste trabalho foi efectuada a análise de interacção SNP-SNP. Considerando esta análise, os génotipos de risco para as interacções SNP-SNP foram determinados como valores de risco ORs com IC a 95% e valores de P. Os efeitos combinados entre dois génotipos foram estudados através da criação de diferentes variáveis, cada uma delas representando uma combinação de dois génotipos, tendo sido considerado como categoria de referência a combinação de génotipos homocigóticos supostos como sendo de “baixo-risco”. Todas as interacções de génotipos com frequência superior a 5% foram consideradas comuns, enquanto as interacções raras foram agrupadas. De notar que todas estas análises estatísticas foram determinadas utilizando o programa SPSS (versão 15.0, SPSS Inc Chicago, IL).

7.3 RESULTADOS

Este estudo envolveu 287 doentes com cancro da mama e 547 controlos saudáveis. Os testes de Kolmogorov-Smirnov e Levene mostraram que a população segue uma distribuição normal e homogeneidade de variáveis contínuas (idade). As características principais (idade, hábitos tabágicos e hábitos alcoólicos) das populações caso-controlo estão apresentadas na Tabela 7.1. Não existem diferenças significativas entre a população de doentes e controlos quando se consideram os factores idade e

Tabela 7.1 – Características gerais das populações de doentes (n=287) e controlos (n=547) envolvidas no estudo.

CARACTERÍSTICAS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P*
Idade^{b,c}			
≤ 30	1 (0,3%)	2 (0,4%)	0,97 ^c
31 – 49	66 (23,0%)	119 (21,8%)	
50 – 69	158 (55,1%)	301 (55,0%)	
≥ 70	62 (21,6%)	125 (22,9%)	
N.D.	0	0	
Hábitos Tabágicos			
Nunca e Ex – F.	250 (87,4%)	490 (91,6%)	0,06 ^c
Fumadores	36 (12,6%)	45 (8,4%)	
N.D.	1	12	
Hábitos Alcoólicos			
Nunca	219 (76,3%)	441 (82,6%)	<0,001 ^c
Social	25 (8,7%)	59 (11,0%)	
Regular	43 (15,0%)	34 (6,4%)	
N.D.	0	13	
Diagnóstico Histológico			
Carcinoma Ductal	251 (87,4%)	-	-
Carcinoma Lobular	14 (4,9%)	-	
Outros tipos de tumor	22 (7,7%)	-	

^a Idade de diagnóstico; ^b Idade do indivíduo controlo à data de diagnóstico do doente correspondente;

^c Grupo doentes *versus* Controlos; * Valor de P χ^2 (Ver Materiais e Métodos).

hábitos alcoólicos. No entanto, os consumidores de álcool são mais prevalentes na população constituída pelos doentes do que na população controlo ($P < 0,001$).

Foram calculadas as frequências de Hardy-Weinberg dos alelos de todos os SNPs estudados, tendo-se verificado, que a distribuição dos genótipos para todos os polimorfismos estão em acordo com o equilíbrio de Hardy-Weinberg ($P > 0,05$), excepto dois dos SNPs (*MUTYH* – rs3219489 e *MLH1* – rs1799977) que desviaram do HWE na população controlo ($P = 0,02$ e $P = 0,04$, respectivamente). Estes resultados não se devem a erros de genotipagem, uma vez que nenhum SNP desvia deste equilíbrio em ambas as populações estudadas e, para além disso, a discriminação alélica dos genótipos dos 9 SNPs incluídos no estudo foi consistente.

Na Tabela 7.2 estão apresentados as frequências genotípicas, a frequência do alelo menos frequente (MAF), obtidas para as populações de doentes e controlos estudadas, bem como os valores de OR com IC a 95%, para os 9 SNPs estudados. As frequências genotípicas determinadas para todos os SNPs não diferem entre as populações estudadas.

Os testes de associação de SNPs obtidos por regressão logística, não revelaram diferenças entre doentes e controlos para 8 dos 9 SNPs estudados (ver Tabela 7.2). Contudo, um dos polimorfismos (*MLH3* Leu844Pro – rs175080) revelou estar associado com uma diminuição de risco para cancro da mama em indivíduos portadores do genótipo homocigótico para a variante (Pro/Pro) [$OR_{Ajustado} = 0,62$ (0,41 – 0,94)] ($P = 0,03$) e o mesmo efeito em indivíduos heterocigóticos (Leu/Pro) [$OR_{Ajustado} = 0,65$ (0,45 – 0,95)] ($P = 0,03$).

Em dois dos genes (*MSH3* e *MSH4*) incluídos neste estudo foram estudados, em cada um, dois polimorfismos. Esta abordagem permitiu avaliar o efeito da combinação de genótipos na susceptibilidade para cancro da mama. Deste modo, determinou-se a frequência da combinação de genótipos e os valores de OR associados em cada interacção, não tendo sido encontradas diferenças significativas entre as interacções estabelecidas (ver Tabela 7.3).

Foram igualmente consideradas interacções SNP-SNP, para todos os polimorfismos dos diferentes genes, com plausibilidade biológica de forma a avaliar o risco individual para neoplasia mamária (ver Tabela 7.4). Os resultados obtidos, após regressão logística, indicaram que duas interacções (*MSH3* Ala1045Thr/*MSH6*

Gly39Glu e *MSH4* Ala97Thr/*MLH3* Leu844Pro) estão associadas com risco individual para a doença (ver Tabela 7.5). A interação *MSH3/MSH6* AA/TC revelou estar associada com uma diminuição de risco para neoplasia mamária [$OR_{Ajustado} = 0,43$ (0,21 – 0,83)] ($P=0,01$). Adicionalmente, três interações entre os polimorfismos *MSH4* Ala97Thr /*MLH3* Leu844Pro foram encontrada como estando associadas com aumento de risco para esta neoplasia (interação AG/AA: $OR_{Ajustado} = 2,35$ (1,23 – 4,49)] ($P=0,01$); interação GG/AA: $OR_{Ajustado} = 2,11$ (1,12 – 3,98)] ($P=0,02$); e a interação GG/AG $OR_{Ajustado} = 1,88$ (1,23 – 3,15)] ($P=0,02$)).

Tabela 7.2 – Frequências genotípicas e frequências do menor alelo (MAF) na população controlo (n=547) e na população de doentes (n=287) e análise de associação dos SNPs estudados individualmente com risco para cancro da mama.

SNPs	GENÓTIPOS	CONTROLOS N (%)	DOENTES N (%)	MAF		TODOS OS CASOS	
				Controlos	Doentes	OR Não Ajustado (IC a 95%)	OR Ajustado (IC a 95%) ^A
<i>MSH3</i> Ala1045Thr A>G	Ala/Ala	246 (45,2%)	121 (42,3%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	Ala/Thr	240 (44,1%)	129 (45,1%)	G: 0,33 (±0,01)	G: 0,35 (±0,02)	1,093 (0,81-1,48)	1,13 (0,83-1,55)
	Thr/Thr	58 (10,7%)	36 (12,6%)			1,26 (0,79-2,02)	1,294 (0,80-2,09)
<i>MSH3</i> Arg940Gln G>A	Arg/Arg	371 (68,1%)	182 (63,6%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	Arg/Gln	158 (29,0%)	96 (33,6%)	A: 0,17 (±0,01)	A: 0,20 (±0,02)	1,24 (0,91-1,69)	1,23 (0,90-1,68)
	Gln/Gln	16 (2,9%)	8 (2,8%)			1,02 (0,43-2,43)	1,03 (0,43-2,47)
<i>MSH4</i> Asn914Ser G>A	Asn/Asn	496 (91,0%)	263 (91,6%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	Asn/Ser	49 (9,0%)	23 (8,0%)	A: 0,05 (±0,01)	A: 0,04 (±0,01)	0,89 (0,53-1,49)	0,83 (0,49-1,41)
	Ser/Ser	0 (0%)	1 (0,3%)			–	–
<i>MSH4</i> Ala97Thr A>G	Ala/Ala	260 (47,7%)	145 (50,7%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	Ala/Thr	239 (43,9%)	117 (40,9%)	G: 0,30 (±0,01)	G: 0,29 (±0,02)	0,88 (0,65-1,19)	0,85 (0,63-1,16)
	Thr/Thr	46 (8,4%)	24 (8,4%)			0,94 (0,55-1,60)	1,03 (0,60-1,79)
<i>MSH6</i> Gly39Glu C>T	Gly/Gly	354 (65,1%)	195 (68,2%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	Gly/Glu	174 (32%)	79 (27,6%)	T: 0,19 (±0,01)	T: 0,18 (±0,02)	0,82 (0,60-1,13)	0,83 (0,60-1,15)
	Glu/Glu	16 (2,9%)	12 (4,2%)			1,36 (0,63-2,94)	1,38 (0,62-3,03)
<i>MLH1</i> Ile219Val A>G	Ile/Ile	255 (46,7%)	129 (44,9%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	Ile/Val	251 (46,0%)	129 (44,9%)	G: 0,30 (±0,01)	G: 0,33 (±0,02)	1,02 (0,75-1,37)	1,01 (0,74-1,37)
	Val/Val	40 (7,3%)	29 (10,1%)			1,43 (0,85-2,42)	1,35 (0,79-2,31)
<i>MLH3</i> Leu844Pro G>A	Pro/Pro	166 (30,5%)	76 (26,6%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	Leu/Pro	283 (52%)	141 (49,3%)	A: 0,43 (±0,02)	A: 0,49 (±0,02)	0,69 (0,47-0,99)[†]	0,65 (0,45-0,95)[†]
	Leu/Leu	95 (17,5%)	69 (24,1%)			0,63 (0,42-0,95)[†]	0,618 (0,41-0,94)[†]

Tabela 7.2 (Continuação) - Frequências genotípicas e frequências do menor alelo (MAF) na população controlo (n=547) e na população de doentes (n=287) e análise de associação dos SNPs estudados individualmente com risco para cancro da mama.

SNPs	GENÓTIPOS	CONTROLOS N (%)	DOENTES N (%)	MAF		TODOS OS CASOS	
				Controlos	Doentes	OR Não Ajustado (IC a 95%)	OR Ajustado (IC a 95%) ^A
PMS1 5'UTR Ex1-4 G>C	GG	352 (64,7%)	181 (63,5%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	CG	179 (32,9%)	90 (31,6%)	C: 0,19 (±0,01)	C: 0,21 (±0,02)	0,98 (0,72-1,33)	1,01 (0,74-1,39)
	CC	13 (2,4%)	14 (4,9%)			2,09 (0,96-4,55)	1,88 (0,85-4,15)
MUTYH His335Gln G>C	His/His	283 (51,7%)	162 (56,4%)			1 (REFERÊNCIA)	1 (REFERÊNCIA)
	His/Gln	235 (43,0%)	107 (37,3%)	C: 0,27 (±0,01)	C: 0,25 (±0,02)	0,80 (0,59-1,07)	0,81 (0,60-1,10)
	Gln/Gln	29 (5,3%)	18 (6,3%)			1,08 (0,58-2,01)	1,05 (0,56-1,98)

^a OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

[†] RESULTADOS A NEGRITO REPRESENTAM OS ESTATISTICAMENTE SIGNIFICATIVOS (P<0,05). OS VALORES DE P ESTÃO AJUSTADOS PARA REGRESSÃO LOGÍSTICA.

Tabela 7.3 – Efeito da combinação de genótipos entre dois polimorfismos de um mesmo gene na susceptibilidade para cancro da mama.

INTERACÇÕES		GENÓTIPOS	CONTROLOS N (%)	DOENTES N (%)	TODOS OS CASOS			
<i>MSH3</i> Ala1045Thr (A>G) / <i>MSH3</i> Arg940Gln (G>A)					OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P*	OR _{AJUSTADO} (IC A 95%) ^a	VALOR DE P*
	AA/GG		246 (45,2%)	120 (42,0%)	1 (REFERÊNCIA)	0,56	1 (REFERÊNCIA)	0,63
	AA/AG		24 (4,4%)	17 (5,9%)	1,45 (0,75-2,81)	0,27	1,47 (0,75-2,87)	0,26
	AG/AA							
	GG/GG		116 (21,3%)	54 (18,9%)	0,95 (0,65-1,41)	0,81	1,03 (0,69-1,53)	0,88
	GG/AA							
	AG/GG		123 (22,6%)	75 (26,2%)	1,25 (0,87-1,79)	0,23	1,26 (0,87-1,82)	0,23
	GG/AG		35 (6,4%)	20 (7,0%)	1,17 (0,65-2,12)	0,60	1,22 (0,67-2,24)	0,52
<i>MSH4</i> Asn914Ser (G>A) / <i>MSH4</i> Ala97Thr (A>G)			CONTROLS N (%)	CASES N (%)	OR _{NÃO AJUSTADO} (IC a 95%)	Valor de P	OR _{Ajustado} (IC a 95%) ^A	Valor de P
	GG/GG		230 (42,2%)	133 (46,5%)	1 (REFERÊNCIA)	0,56	1 (REFERÊNCIA)	0,46
	AA/GG		19 (3,5%)	13 (4,5%)	1,18 (0,57-2,47)	0,66	1,120 (0,528-2,378)	0,77
	AG/AG							
	AG/GG		30 (5,5%)	11 (3,8%)	0,63 (0,31-1,31)	0,22	0,598 (0,287-1,246)	0,17
	GG/AA		46 (8,4%)	24 (8,4%)	0,90 (0,53-1,55)	0,71	0,991 (0,570-1,723)	0,98
	GG/AG		220 (40,4%)	105 (36,7%)	0,83 (0,60-1,13)	0,23	0,802 (0,581-1,107)	0,18

^a OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

* Os valores de P estão ajustados para regressão logística.

Tabela 7.4 – Efeitos das interações SNP-SNP no cancro da mama.

INTERACÇÕES		CONTROLOS N (%)	DOENTES N (%)	TODOS OS CASOS			
<i>MLH1</i> Ile219Val (A>G) / <i>PMS1</i> 5'UTR Ex1-4(G>C)	GENÓTIPOS			OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
	GG/GG	24 (4,4%)	18 (6,3%)	1 (REFERÊNCIA)	0,27	1 (REFERÊNCIA)	0,26
	AACC AG/CC GG/CC GG/CG	28 (5,1%)	23 (8,1%)	1,10 (0,48-2,50)	0,83	1,23 (0,53-2,86)	0,63
	AA/CG	87 (16,0%)	48 (16,8%)	0,74 (0,36-1,49)	0,39	0,89 (0,43-1,83)	0,74
	AA/GG	162 (29,8%)	72 (25,3%)	0,59 (0,30-1,16)	0,13	0,66 (0,33-1,31)	0,23
	AG/CG	77 (14,2%)	33 (11,6%)	0,57 (0,27-1,19)	0,14	0,61 (0,29-1,30)	0,20
	AG/GG	166 (30,5%)	91 (31,9%)	0,73 (0,38-1,42)	0,35	0,84 (0,43-1,66)	0,62
<i>MLH1</i> Ile219Val (A>G) / <i>MLH3</i> Leu844Pro (G>A)		CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P
	AA/AA	48 (8,8%)	29 (10,1%)	1 (REFERÊNCIA)	0,18	1 (REFERÊNCIA)	0,19
	AA/AG	128 (23,5%)	69 (24,1%)	0,89 (0,52-1,54)	0,68	0,83 (0,47-1,44)	0,50
	AA/GG	78 (14,3%)	30 (10,5%)	0,68 (0,34-1,19)	0,16	0,60 (0,32-1,14)	0,12
	AG/AA	40 (7,4%)	32 (11,2%)	1,32 (0,69-2,55)	0,40	1,27 (0,65-2,48)	0,48
	AG/AG	131 (24,1%)	57 (19,9%)	0,72 (0,41-1,26)	0,25	0,66 (0,38-1,17)	0,15
	AG/GG	79 (14,5%)	40 (14,0%)	0,84 (0,46-1,52)	0,56	0,80 (0,43-1,47)	0,47
	GG/AA GG/AG GG/GG	40 (7,4%)	29 (10,1%)	1,20 (0,62-2,33)	0,59	1,07 (0,54-2,101)	0,85

Tabela 7.4 (Continuação) - Efeitos das interações SNP-SNP no cancro da mama.

<i>MSH3</i> Ala1045Thr (A>G) / <i>MSH6</i> Gly39Glu (C>T)	CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
GG/CC	29 (5,3%)	26 (9,1%)	1 (REFERÊNCIA)	0,36	1 (REFERÊNCIA)	0,26
AA/TT						
AG/TT	43 (7,9%)	22 (7,7%)	0,57 (0,27-1,19)	0,14	0,53 (0,25-1,13)	0,10
GG/TT						
GG/TC						
AA/TC	81 (14,9%)	33 (11,5%)	0,45 (0,23-0,89)	0,02	0,43 (0,21-0,83)	0,01
AA/CC	158 (29,1%)	83 (29,0%)	0,59 (0,32-1,06)	0,07	0,55 (0,30-1,00)	0,05
AG/TC	66 (12,2%)	36 (12,6%)	0,61 (0,31-1,19)	0,14	0,61 (0,31-1,21)	0,16
AG/CC	166 (30,6%)	86 (30,1%)	0,58 (0,32-1,04)	0,07	0,55 (0,30-1,01)	0,05
<i>MSH3</i> Arg940Gln (G>A) / <i>MSH6</i> Gly39Glu (C>T)	CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
AG/CC	98 (18,0%)	60 (21,0%)	1 (REFERÊNCIA)	0,59	1 (REFERÊNCIA)	0,62
GG/TT						
AG/TT	32 (5,9%)	20 (7,0%)	1,02 (0,54-1,95)	0,95	1,02 (0,54-2,00)	0,92
AA/TC						
AA/CC						
GG/TC	113 (20,8%)	48 (16,8%)	0,69 (0,44-1,11)	0,12	0,70 (0,44-1,13)	0,14
GG/CC	246 (45,2%)	128 (44,8%)	0,85 (0,58-1,25)	0,41	0,86 (0,58-1,27)	0,45
AG/TC	55 (10,1%)	30 (10,5%)	0,89 (0,52-1,54)	0,68	0,92 (0,52-1,60)	0,75

Tabela 7.4 (Continuação) - Efeitos das interações SNP-SNP no cancro da mama.

<i>MUTYH</i> His335Gln (G>C) / <i>MSH6</i> Gly39Glu (C>T)	CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
CG/CC	163 (30,0%)	69 (24,1%)	1 (REFERÊNCIA)	0,15	1 (REFERÊNCIA)	0,26
GG/TT						
CG/TT	45 (8,3%)	30 (10,5%)	1,58 (0,92-2,71)	0,10	1,50 (0,86-2,60)	0,15
CC/TC						
CC/CC						
GG/TC	92 (16,9%)	44 (15,4%)	1,13 (0,72-1,78)	0,60	1,11 (0,69-1,76)	0,67
GG/CC	175 (32,2%)	112 (39,2%)	1,51 (1,05-2,19)	0,03	1,44 (0,99-2,10)	0,06
CG/TC	69 (12,7%)	31 (10,8%)	1,06 (0,64-1,77)	0,82	1,02 (0,61-1,71)	0,94
<i>MSH4</i> Asn914Ser (G>A) / <i>MLH3</i> Leu844Pro (G>A)	CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
GG/GG	152 (27,9%)	73 (25,5%)	1 (REFERÊNCIA)	0,28	1 (REFERÊNCIA)	0,17
AA/AG						
AG/AA	22 (4,0%)	16 (5,6%)	1,51 (0,75-3,06)	0,25	1,50 (0,74-3,06)	0,26
AG/GG						
AG/AG	27 (5,0%)	8 (2,8%)	0,62 (0,27-1,43)	0,26	0,54 (0,23-1,28)	0,16
GG/AA	87 (16,0%)	57 (19,9%)	1,36 (0,88-2,11)	0,16	1,41 (0,90-2,20)	0,13
GG/AG	256 (47,1%)	132 (46,2%)	1,07 (0,76-1,52)	0,69	1,05 (0,73-1,49)	0,81

Tabela 7.4 (Continuação) - Efeitos das interações SNP-SNP no cancro da mama.

<i>MSH4</i> Ala97Thr (A>G) / <i>MLH3</i> Leu844Pro (G>A)	CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
GG/GG	85 (15,6%)	28 (9,8%)	1 (REFERÊNCIA)	0,02	1 (REFERÊNCIA)	0,01
AA/AA						
AA/AG	46 (8,5%)	24 (8,4%)	1,58 (0,83-3,04)	0,17	1,70 (0,87-3,32)	0,12
AA/GG						
AG/AA	41 (7,5%)	31 (10,8%)	2,30(1,22-4,32)	0,01	2,35 (1,23-4,49)	0,01
AG/AG	127 (23,3%)	47 (16,4%)	1,12 (0,65-1,93)	0,67	1,03 (0,59-1,78)	0,93
AG/GG	71 (13,1%)	39 (13,6%)	1,67 (0,94-2,98)	0,08	1,54 (0,85-2,78)	0,15
GG/AA	44 (8,1%)	32 (11,2%)	2,21 (1,18-4,12)	0,01	2,11 (1,12-3,98)	0,02
GG/AG	130 (23,9%)	85 (29,7%)	1,99 (1,20-3,30)	0,01	1,88 (1,12-3,15)	0,02
<i>MLH1</i> Ile219Val (A>G) / <i>MSH4</i> Asn914Ser (G>A)	CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
GG/GG	34 (6,2%)	27 (9,4%)	1 (REFERÊNCIA)	0,36	1 (REFERÊNCIA)	0,49
AA/AA						
AA/AG	49 (9,0%)	24 (8,4%)	0,62 (0,31-1,25)	0,18	0,63 (0,30-1,28)	0,20
AG/AG						
GG/AG						
AA/GG	236 (43,3%)	114 (39,7%)	0,61 (0,35-1,06)	0,08	0,66 (0,37-1,16)	0,15
AG/GG	226 (41,5%)	122 (42,5%)	0,68 (0,39-1,18)	0,17	0,73 (0,41-1,29)	0,27

Tabela 7.4 (Continuação) - Efeitos das interações SNP-SNP no cancro da mama.

<i>MLH1</i> Ile219Val (A>G) / <i>MSH4</i> Ala97Thr (A>G)	CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
AG/GG	121 (22,2%)	59 (20,6%)	1 (REFERÊNCIA)	0,29	1 (REFERÊNCIA)	0,19
AA/AA						
AG/AA						
GG/AA	81 (14,9%)	48 (16,8%)	1,22 (0,76-1,95)	0,42	1,29 (0,79-2,10)	0,31
GG/AG						
GG/GG						
AA/AG	116 (21,3%)	45 (15,7%)	0,80 (0,50-1,27)	0,33	0,79 (0,49-1,27)	0,32
AA/GG	118 (21,7%)	73 (25,5%)	1,27 (0,83-1,94)	0,27	1,32 (0,85-2,04)	0,21
AG/AG	109 (20,0%)	61 (21,3%)	1,15 (0,74-1,79)	0,54	1,17(0,74-1,83)	0,50

^a OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

[†] Resultados a negrito representam os estatisticamente significativos (P<0,05). Os valores de P estão ajustados para regressão logística.

Tabela 7.5 – Interações entre SNPs de diferentes genes com plausibilidade biológica e risco para cancro da mama.

INTERACÇÕES		GENÓTIPOS	CONTROLOS N (%)	DOENTES N (%)	TODOS OS CASOS			
<i>MSH3</i> Ala1045Thr (A>G) / <i>MSH6</i> Gly39Glu (C>T)					OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]
	GG/CC	29 (5,3%)	26 (9,1%)	1 (REFERÊNCIA)	0,36	1 (REFERÊNCIA)	0,26	
	AA/TT AG/TT GG/TT GG/TC	43 (7,9%)	22 (7,7%)	0,57 (0,27-1,19)	0,14	0,53 (0,25-1,13)	0,10	
	AA/TC	81 (14,9%)	33 (11,5%)	0,45 (0,23-0,89)	0,02	0,43 (0,21-0,83)	0,01	
	AA/CC	158 (29,1%)	83 (29,0%)	0,59 (0,32-1,06)	0,08	0,55 (0,30-1,00)	0,05	
	AG/TC	66 (12,2%)	36 (12,6%)	0,61 (0,31-1,19)	0,14	0,61 (0,310-1,21)	0,16	
	AG/CC	166 (30,6%)	86 (30,1%)	0,58 (0,32-1,04)	0,07	0,55 (0,30-1,01)	0,05	
<i>MSH4</i> Ala97Thr (A>G) / <i>MLH3</i> Leu844Pro (G>A)		CONTROLOS N (%)	DOENTES N (%)	OR _{NÃO AJUSTADO} (IC A 95%)	VALOR DE P [†]	OR _{AJUSTADO} (IC A 95%) ^A	VALOR DE P [†]	
	GG/GG	85 (15,6%)	28 (9,8%)	1 (REFERÊNCIA)	0,02	1 (REFERÊNCIA)	0,01	
	AA/AA AA/AG AA/GG	46 (8,5%)	24 (8,4%)	1,58 (0,83-3,04)	0,17	1,70 (0,87-3,32)	0,12	
	AG/AA	41 (7,5%)	31 (10,8%)	2,30 (1,22-4,32)	0,01	2,35 (1,23-4,49)	0,01	
	AG/AG	127 (23,3%)	47 (16,4%)	1,12 (0,65-1,93)	0,67	1,03 (0,59-1,78)	0,93	
	AG/GG	71 (13,1%)	39 (13,6%)	1,67 (0,94-2,98)	0,08	1,54 (0,85-2,78)	0,15	
	GG/AA	44 (8,1%)	32 (11,2%)	2,21 (1,18-4,12)	0,01	2,11 (1,12-3,98)	0,02	
	GG/AG	130 (23,9%)	85 (29,7%)	1,99 (1,20-3,30)	0,01	1,88 (1,12-3,15)	0,02	

^a OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

[†] Resultados a negrito representam os estatisticamente significativos (P<0,05). Os valores de P estão ajustados para regressão logística.

7.4 Discussão

Que seja do nosso conhecimento, este é o primeiro estudo realizado com o intuito de analisar o possível papel de 9 polimorfismos em 7 genes da via MMR e a sua influência na susceptibilidade genética para cancro da mama. O objectivo principal foi compreender a possível contribuição atribuída à eventual relevância funcional dos polimorfismos e das interacções SNP-SNP nos genes da via MMR no risco para cancro da mama. Os resultados apontam para alguns genes candidatos e interacções SNP-SNP associados com susceptibilidade individual para cancro da mama.

A primeira associação plausível encontrada foi no polimorfismo *MLH3* (Leu844Pro), estando os genótipos homozigótico para a variante (Pro/Pro) e heterozigótico (Leu/Pro) relacionados com uma diminuição de risco na susceptibilidade para cancro da mama. Este gene (*MLH3*) tem sido maioritariamente relacionado com alguns casos de cancro colorectal (HNPCC, do inglês *Hereditary nonpolyposis colorectal cancer*), onde está descrito um considerável número de mutações consistentes com o seu potencial envolvimento no desenvolvimento de HNPCC (Lipkin *et al.*, 2001). Uma vez que este gene desempenha funções na reparação de IDLs (*Loops de Inserção/Delecção*) e o seu envolvimento na MSI, alterações na sua estrutura e função podem desencadear eventos oncogénicos (Lipkin *et al.*, 2000), inclusive no cancro da mama.

Um estudo desenvolvido por Michielis *et al.* (Michiels *et al.*, 2007) numa população com neoplasia de pulmão (constituída por 151 casos de cancro do pulmão e 172 controlos), mostrou que o polimorfismo *MLH3* (Leu844Pro) estaria associado com um aumento de risco para cancro do pulmão [OR: 1,97 (1,06 – 3,65); P=0,04]. No entanto, contrastando com o cancro da mama, a maioria dos casos de cancro do pulmão são normalmente provocados por hábitos tabágicos, o que mostra que a etiologia das duas doenças é diferente. Todavia, o facto de existirem polimorfismos em genes envolvidos na reparação de DNA pode modificar o risco para qualquer tipo de cancro (Sakiyama *et al.*, 2005).

Relativamente ao polimorfismo no gene *MLH1* (Ile219Val), neste estudo não foi encontrada qualquer associação entre este SNP e o cancro da mama, o que é consistente com um estudo anterior desenvolvido por Lee *et al.* (Lee *et al.*, 2005b) que obteve os mesmos resultados numa população Coreana (872 doentes e 671 controlos).

Outro estudo publicado por Smith *et al.* (Smith *et al.*, 2008), descreveu uma diminuição de risco [OR: 0,49 (0,29 – 0,85); P <0,05] na susceptibilidade para cancro da mama numa população Caucasiana (336 doentes e 416 controlos). Estas diferenças entre estudos podem estar relacionadas com diferentes factores de risco e/ou diferenças no perfil genético inerentes às diferentes populações. A via de reparação MMR envolve a acção de diversas proteínas heterodiméricas. Deste modo, é importante avaliar neste estudo as interacções gene-gene. A análise dos genótipos de risco específico revela diferenças significativas para duas associações entre diferentes genes MMR. A interacção *MSH3* Ala1045Thr/*MSH6* Gly39Glu (AA/TC) está associada com uma diminuição de risco [OR_{Ajustado} = 0,43 (0,21 – 0,83); P=0,01]. De salientar que os genes *MSH3* e *MSH6* actuam como sensores da via MMR, detectando erros que ocorrem durante a replicação de DNA, estando envolvidos na correcção de *mismatches* pós-replicativos (Li, 2008).

Encontrou-se outra interacção (*MSH4* Ala97Thr/*MLH3* Leu844Pro), associada com um aumento de risco para cancro da mama. Como descrito por Santucci-Darmanin *et al.* (Santucci-Darmanin *et al.*, 2002) a proteína MLH3 está associada com a proteína MSH4 específica da meiose em células meióticas de mamíferos, o que suporta fortemente a possibilidade do gene *MLH3* desempenhar funções na recombinação meiótica em mamíferos (Santucci-Darmanin *et al.*, 2002).

Foi reconhecido durante alguns anos que a via de reparação MMR afecta a eficiência da recombinação meiótica (Jiricny, 2006). Todavia, as proteínas envolvidas na via MMR estão igualmente envolvidas nos processos de recombinação mitótica e desempenham um papel crítico na manutenção da estabilidade mitótica dos genomas eucariotas (Harfe e Jinks-Robertson, 2000). Durante a recombinação mitótica as proteínas MMR previnem a troca entre sequências não-idênticas. De facto, tem sido demonstrado que determinadas sequências homólogas apenas recombinaem quando a via MMR está inactiva (Jiricny, 2006). A inactivação da via MMR conduz a um aumento na frequência da recombinação mitótica e uma recombinação ineficiente pode aumentar o risco da susceptibilidade para o desenvolvimento de cancro. Consequentemente, a função anti-recombinação da via MMR não suprime apenas a recombinação homóloga mas também actua de forma a prevenir rearranjos cromossómicos envolvendo translocações e deleções (Schofield e Hsieh, 2003).

Como resultado, é possível que modificações estruturais ou funcionais relacionadas com a interacção *MSH4* Ala97Thr/*MLH3* Leu844Pro possam estar associadas com aumento de risco para o cancro da mama, modificando a progressão da via MMR e aumentando assim as taxas de recombinação mitótica nas células da glândula mamária. No entanto, são necessários estudos com maior número de amostras que permitam verificar o papel dos genes *MSH4* e *MLH3* na recombinação mitótica.

Concluindo, os resultados obtidos com este trabalho indicam que a variante do gene *MLH3* (844Pro) está associada com uma diminuição de risco para cancro da mama. Por outro lado, foi ainda observado uma diminuição de risco entre os genótipos homozigótico e heterozigótico para a interacção *MSH3* Ala1045Thr/*MSH6* Gly39Glu e um risco aumentado entre diversos genótipos combinados da interacção *MSH4* Ala97Thr/*MLH3* Leu844Pro o que aponta para um modelo de interacção gene-gene multiplicativo.

A associação entre o risco para cancro e os diversos genótipos obtidos reforça a hipótese da função da via MMR na susceptibilidade para cancro da mama. Na realidade, a via MMR desempenha uma função muito importante na estabilidade genómica, contribuindo para a supressão de tumor através da redução de mutações e promovendo a apoptose na resposta a lesões de DNA durante os processos de replicação e recombinação mitótica (Buermeyer *et al.*, 1999; Li, 2008). Diferentes actividades e funções destes genes, bem como variações polimórficas podem alterar o nível de reparação, conduzindo assim a elevadas taxas de mutações e, por conseguinte, um aumento do risco para cancro da mama ou, inversamente, desempenhar um papel protector na cancerigénese mamária. Contudo, estudos independentes efectuados com populações maiores são essenciais para suportar estes resultados.

CAPÍTULO 8

RISCO PARA CANCRO DA MAMA E SNPs EM GENES
DA VIA DE REPARAÇÃO POR RECOMBINAÇÃO
HOMÓLOGA *XRCC2*, *XRCC3*, *NBS1* E *RAD51*

8. RISCO PARA CANCRO DA MAMA E SNPs EM GENES DA VIA DE REPARAÇÃO POR RECOMBINAÇÃO HOMÓLOGA *XRCC2*, *XRCC3*, *NBS1* E *RAD51*

Publicado em Silva *et al.*, 2010, *Cancer Epidemiology*, **34**: 85 - 92

8.1 INTRODUÇÃO

Como tem vindo a ser focado ao longo desta dissertação, o cancro da mama é a patologia neoplásica mais comum entre as mulheres, sendo provocado por uma complexa combinação de factores genéticos e ambientais. Na realidade, os genes e o ambiente partilham a progressão para a maioria, senão de todos, os tipos de cancro não familiares.

Agentes cancerígenos como a radiação ionizante (IR), agente cancerígeno associado com o cancro da mama (Hu *et al.*, 2002), e radicais livres induzem quebras de DNA em cadeia dupla (DSBs) (Lieber *et al.*, 2003). As DSBs não reparadas ou em que tenha ocorrido uma falha de reparação podem resultar em morte celular, rearranjos cromossomais e instabilidade genómica, fenómenos envolvidos no processo de cancerigénese. Uma deficiente reparação de DSBs pode, deste modo, contribuir para a susceptibilidade para cancro da mama (Han *et al.*, 2004a). A via de reparação de DSBs pode ser dividida em duas vias diferentes, via de reparação por recombinação homóloga (HRR) e via de reparação não homóloga (NHEJ) (descritas anteriormente na Capítulo 1 secção 1.7.5). Um dos pontos relevantes do estudo de DSBs no risco para neoplasia mamária assenta no facto dos dois mais importantes genes associados com cancro da mama (*BRCA1* e *BRCA2*) estarem envolvidos na via HRR. Enquanto a via NHEJ é uma via propensa a erros, a via HRR repara as quebras de DNA com alta fidelidade utilizando para tal a sequência homóloga como modelo para sintetizar uma nova cadeia livre de erros (Thompson e Schild, 2001). Diversos polimorfismos genéticos, especialmente polimorfismos de uma única base (SNPs), têm sido identificados nos genes envolvidos na via HRR (*e.g.* *XRCC2*, *XRCC3*, *NBS1* e *Rad51*) podendo influenciar a

capacidade de reparação em doentes com cancro da mama e, por sua vez, conferir predisposição genética para a doença.

O gene *NBS1* codifica para a proteína NBS1 que está envolvida no complexo proteico MRE11-NBS1-RAD51, complexo este responsável pelo reconhecimento da DSB e pelos passos iniciais do processo de reparação (Dudas e Chovanec, 2004; Ronen e Glickman, 2001). A contribuição de regiões distintas da proteína NBS1 na resposta a lesões induzidas por irradiação, e em particular na reparação de DSB, tem sido extensamente avaliada fornecendo uma visão mais aprofundada da interacção entre a NBS1 e outras proteínas (i.e. ATM) envolvidas no processo (Berkovich *et al.*, 2007; Cariveau *et al.*, 2007; Difilippantonio *et al.*, 2007). O polimorfismo no gene *NBS1* (Ex5-32C>G, E185Q, rs1805794) tem sido estudado em estudos caso-controlo com o intuito de avaliar a susceptibilidade individual para cancro da mama. (Forsti *et al.*, 2004; Kuschel *et al.*, 2002; Millikan *et al.*, 2005; Zhang *et al.*, 2005). Os resultados obtidos para este SNP não revelam qualquer associação em diferentes populações: Caucasiana (Forsti *et al.*, 2004; Kuschel *et al.*, 2002; Millikan *et al.*, 2005), Afro-Americana (Millikan *et al.*, 2005) e Chinesa (Zhang *et al.*, 2005).

Por sua vez, a proteína RAD51 desempenha um papel central na via de reparação HRR, na qual é responsável pela invasão das extremidades quebradas das DSBs em cromátides irmãs intactas (Rodrigue *et al.*, 2006; Thompson e Schild, 2001). O polimorfismo existente no gene *RAD51* na região 5'UTR (Ex1-59G>T, rs1801321) não tem sido muito explorado em estudos de susceptibilidade para cancro da mama (Kuschel *et al.*, 2002; Lee *et al.*, 2005b). Do nosso conhecimento, os resultados disponíveis para este SNP não revelaram qualquer associação no papel deste polimorfismo no risco para neoplasia mamária nem na população Caucasiana (Kuschel *et al.*, 2002) nem na população Coreana (Lee *et al.*, 2005b).

A proteína XRCC2 é uma proteína relacionada com a RAD51, essencial na eficiência da via HRR, e consequentemente na manutenção da estabilidade cromossómica, fazendo parte de um filamento de nucleoproteína que actua como um co-factor para as actividades de invasão e troca de cadeias pela RAD51. Contudo, existem outras indicações sobre o seu envolvimento nas etapas finais da via HRR, nomeadamente na migração da cadeia e resolução da junção de *Holliday* (Dudas e Chovanec, 2004; Hoeijmakers, 2001; Rodrigue *et al.*, 2006; Wood *et al.*, 2005). O

polimorfismo no gene *XRCC2* (Ex3+442G>A, R188H, rs3218536) tem sido estudado num amplo conjunto de estudos de associação na susceptibilidade para cancro da mama (Breast Cancer Association Consortium, 2006; Han *et al.*, 2004a; Kuschel *et al.*, 2002; Lee *et al.*, 2005b; Loizidou *et al.*, 2008; Millikan *et al.*, 2005; Pooley *et al.*, 2008; Webb *et al.*, 2005). Na realidade, este SNP é um dos mais estudados da via HRR, contudo os resultados obtidos têm sido contraditórios. Diversos estudos descrevem um efeito marginalmente protector em portadores do alelo variante (188His) (Loizidou *et al.*, 2008; Pooley *et al.*, 2008). Han *et al.* (Han *et al.*, 2004a) descreve um efeito protector em mulheres com elevados níveis plasmáticos de α -caroteno. Outros estudos reportaram resultados nulos em populações de Afro-Americanas (Millikan *et al.*, 2005), Caucasianas (Millikan *et al.*, 2005; Tranah *et al.*, 2004; Webb *et al.*, 2005) e Coreanas (Lee *et al.*, 2005b). Um outro estudo desenvolvido por Kuschel *et al.* (Kuschel *et al.*, 2002) mostrou um risco marginalmente aumentado em mulheres portadoras do alelo variante numa população Caucásiana.

A proteína *XRCC3* está também envolvida na via de reparação por HRR e é também membro da família das proteínas relacionadas com *RAD51*. O polimorfismo *XRCC3* (Ex8-5C>T, T241M, rs861539) tem sido estudado em diversos estudos caso-controlo com vista a avaliar o seu envolvimento na susceptibilidade individual para o cancro (Manuguerra *et al.*, 2006).

O resultado dos estudos de associação dos polimorfismos acima referidos e susceptibilidade para cancro da mama têm sido contraditórios e justificam, deste modo, a extensão/continuidade destes estudos noutras populações. Dada a contraditoriedade de resultados publicados até à data, desenvolvemos um estudo caso-controlo na população Caucásiana Portuguesa (uma população constituída por 289 doentes com cancro da mama e outra por 548 controlos saudáveis) de forma a avaliar o potencial envolvimento das variações genéticas polimórficas em quatro genes da via HRR (*RAD51*, *NBS1*, *XRCC3* e *XRCC2*) na susceptibilidade para cancro da mama, utilizando 45 tagSNPs característicos destes genes. A abordagem por tagSNPs foi utilizada para aumentar a variabilidade em cada gene estudado.

8.2 MATERIAIS E MÉTODOS

8.2.1 SELECÇÃO DA POPULAÇÃO

A descrição efectuada no Capítulo 3, ponto 3.2.1 aplica-se neste novo capítulo, à excepção do tamanho da amostra, já que para este estudo foram utilizados 289 doentes com cancro da mama e 548 indivíduos que constituem a população controlo. O diagnóstico histológico foi confirmado em todos os casos e a amostra inclui 256 carcinomas tipo ductal (88,3%), 14 carcinomas tipo lobular (4,8%) e 22 casos classificados como sendo de outro tipo de tumor da mama (6,9%).

8.2.2 EXTRACÇÃO DE DNA

A metodologia utilizada encontra-se descrita no Capítulo 3, ponto 3.2.2.

8.2.3 SELECÇÃO DOS TAGSNPs

A abordagem utilizando tagSNPs permite identificar um conjunto de SNPs que se encontrem em desequilíbrio de *linkage* de forma a incluir eficientemente todos os SNPs conhecidos, e alguns mesmo que desconhecidos, no mesmo gene. Os tagSNPs foram identificados recorrendo à base de dados *GVS: Genome Variation Server*⁶. Na utilização desta ferramenta foi seleccionado um conjunto de tagSNPs de forma que para cada gene, todos os SNPs conhecidos e mais comuns (MAF > 0,05) apresentassem um valor de $r^2 > 0,8$ (factor de correlação, corresponde ao valor mínimo que se impõe para as variações que pertencem a um mesmo gene). Foram incluídos neste estudo quatro SNPs em quatro genes que actuam na via HRR, *XRCC2* (Ex3+442G>A, R188H, rs3218536), *XRCC3* (Ex8-5C>T, T241M, rs861539), *NBS1* (Ex5-32C>G, E185Q, rs1805794) e *RAD51* 5'UTR (Ex1-59G>T, rs1801321), o que corresponde à inclusão de um total de 45 SNPs presentes nestes genes, com um r^2 na ordem de 1,0 (Tabela 8.1), o que indica um desequilíbrio de *linkage* elevado e baixa recombinação meiótica. Os tagSNPs seleccionados estão maioritariamente localizados nas regiões intrónicas dos genes estudados.

⁶ <http://gvs-p.gs.washington.edu/GVS/>

Tabela 8.1 – Relação dos tagSNPs seleccionados para os genes em estudo (fonte: Genome Variation Server).

GENES	TAGSNPs ESTUDADOS	SNPs TAGGED	MAF	r^2	FUNÇÃO
<i>XRCC2</i>	rs3218536*	rs3218446	9%	1.000	Intrão
		rs3218455		1.000	Intrão
		rs3218480		1.000	Intrão
<i>XRCC3</i>	rs861539*	rs861531	41%	0.966	Intrão
		rs861534		0.961	Intrão
<i>RAD51</i>	rs1801321*	rs957603	46%	1.000	Intrão
		rs4924496		1.000	Intrão
		rs7177265		1.000	Intrão
		rs7180135		1.000	3'-UTR
		rs12148915		1.000	Intrão
		rs12592524		1.000	Intrão
<i>NBS1</i>	rs1805794*	rs741777	28%	1.000	Intrão
		rs741778		0.919	Intrão
		rs6985793		0.878	Intrão
		rs9649958		0.919	Intrão
		rs1061302		1.000	Sinónimo-codificante
		rs1063045		1.000	Sinónimo-codificante
		rs1805790		1.000	Intrão
		rs1805795		1.000	Intrão
		rs1805797		1.000	Intrão
		rs1805799		1.000	Intrão
		rs1805818		1.000	Intrão
		rs1805844		1.000	Intrão
		rs2073635		1.000	Intrão
		rs2234744		0.959	Intrão
		rs2280780		1.000	Intrão
		rs2308962		1.000	Intrão
		rs2339025		1.000	Intrão
		rs3736639		1.000	Intrão
		rs6470523		1.000	Intrão
		rs6470524		1.000	Intrão
		rs6987873		1.000	Intrão
		rs6990969		1.000	Intrão
		rs7006322		1.000	Intrão
rs7829246	1.000	Intrão			
rs7840099	1.000	Intrão			
rs9792335	1.000	Intrão			
rs10956375	1.000	Intrão			
rs13275276	1.000	Intrão			
rs13278453	1.000	Intrão			
rs13312935	1.000	Intrão			

* SNPs não-sinónimos – codificantes; r^2 – factor de correlacção associado ao valor estimado de desequilíbrio de linkage (D'), que varia entre 0 e 1.

8.2.4 GENOTIPAGEM DOS POLIMORFISMOS

A genotipagem dos polimorfismos presentes nos genes *XRCC2* e *XRCC3* foi efectuada por PCR-RFLP. Os *primers* utilizados nas reacções de PCR bem como a temperatura de *annealing* utilizada para a amplificação de cada SNP, e os produtos obtidos especificamente após RFLP estão descritos na Tabela 8.2.

A reacção de amplificação dos fragmentos pretendidos para cada polimorfismo realizou-se num volume final de 50 µl contendo 100-150 ng de DNA genómico, 0,6 µM de cada *primer*, 0,75 U de *Immolase* (Bioline), 2,5 mM de MgCl₂ e 0,8 mM de dNTPs, em tampão específico da polimerase à concentração final de 1,3× e 1,0× para os polimorfismos nos genes *XRCC2* e *XRCC3*, respectivamente. As condições de PCR aplicadas consistiram num período inicial de activação de 7 minutos a 95 °C, 35 (*XRCC3*) ou 32 (*XRCC2*) ciclos de amplificação incluindo desnaturação a 94 °C durante 30 segundos, temperatura de *annealing* de acordo com o SNP (ver Tabela 8.2) durante 30 segundos e extensão a 72 °C durante 30 segundos, e um período de extensão final de 10 minutos a 72 °C. De forma a confirmar a amplificação dos respectivos segmentos, os produtos de PCR foram submetidos a electroforese em gel de agarose.

Estes polimorfismos originam ganho ou perda de locais de restrição, os quais permitirão discriminar por RFLP, após digestão com enzima de restrição, os alelos comum e variante que caracterizam este SNP. Os perfis de restrição característicos estão apresentados na Tabela 8.2. As experiências foram realizadas em duplicado em ensaios independentes, sendo os resultados inconclusivos reanalisados.

Os polimorfismos nos genes *NBS1* e *RAD51* 5'UTR foram genotipados por PCR em tempo real (AB7300) utilizando os ensaios da Applied Biosystems *TaqMan*[®] *SNP Genotyping Assays* (as referências dos ensaios são C__26470398_10 e C__7482700_10, respectivamente para cada SNP) segundo instruções fornecidas pelo fabricante. Para uma eficiente discriminação alélica as amostras de DNA foram quantificadas utilizando o Reagente de Quantificação Quant-iT[™] PicoGreen dsDNA Assay (Molecular Probes, Invitrogen) de acordo com as instruções fornecidas pelo fabricante.

Tabela 8.2 - Condições de PCR-RFLP para os polimorfismos *XRCC2* R188H e *XRCC3* T241M.

SEQUÊNCIA DOS <i>PRIMERS</i>	T _{ANNEALING} (°C)	FRAGMENTO PCR (BP)	EFEITO DA ENZIMA DE RESTRICÇÃO	PERFIS APÓS ANÁLISE DE RESTRICÇÃO
<i>XRCC2</i> R188H				
<i>Forward</i> : 5'-GGT GTA CTG CAG TAG TAG CAC CCA CTT AC-3'	62 °C	307	Arg188His G→A, cria um local de	G/G : 307 bp; G/A : 307, 214, 93 bp;
<i>Reverse</i> : 5'-CAC ATC ACA CAG TCG TCG AGA GGC-3'			reconhecimento para <i>SexAI</i>	A/A : 214, 93 bp
<i>XRCC3</i> T241M				
<i>Forward</i> : 5'-GTA CTG CTG TCT CGG GGC ATG- 3'	64 °C	315	Thr241Met C→T cria um local de	C/C : 22, 293 bp C/T : 22, 105, 188, 293 bp
<i>Reverse</i> : 5' -CGA TGG TTA GGC ACA GGC TGC- 3'			reconhecimento para <i>NlaIII</i>	T/T : 22, 105, 188 bp.

8.2.5 ANÁLISE ESTATÍSTICA

A abordagem estatística efectuada encontra-se descrita no Capítulo 3, ponto 3.2.4.

De notar que todas estas análises estatísticas foram determinadas utilizando o programa SPSS (versão 15.0, SPSS Inc Chicago, IL). No entanto, para garantir a exactidão dos resultados, a análise estatística foi também realizada utilizando o software SNPStats (Sole *et al.*, 2006).

8.3 RESULTADOS

Este estudo foi constituído por 289 doentes com cancro da mama diagnosticado e confirmado histologicamente e 548 controlos saudáveis emparelhados para a idade sem doença maligna actual ou prévia. Na Tabela 8.3 estão descritas as características principais das populações envolvidas no estudo. Estes resultados não revelaram diferenças significativas nas distribuições de idade e hábitos tabágicos entre

doentes e controlos. No entanto, o consumo de álcool regular está sob-representado na população de doentes quando comparado com a população controlo (Tabela 8.3).

Tabela 8.3 – Características gerais para as populações de doentes com cancro da mama (n=289) e controlo (n=548).

CARACTERÍSTICAS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a
Idade			
30-49	69 (23,9)	122 (22,3)	0,79
50-69	159 (55,0)	301 (54,9)	
> 70	61 (21,1)	125 (22,8)	
Idade Média	59,72	60,00	
Hábitos Tabágicos			
Nunca e Ex – F.	252 (87,8)	491 (91,6)	0,08
Fumadores	35 (12,2)	45 (8,4)	
N.D.	2	12	
Hábitos Alcoólicos			
Nunca	219 (75,8)	442 (82,6)	P ≤ 0,01
Social	26 (9,0)	59 (11,0)	
Regular	44 (15,2)	34 (6,4)	
N.D.	-	13	

^a Valor de P χ^2 (Ver Materiais e Métodos). N.D. – Não Determinado

As frequências genotípicas determinadas na população Portuguesa, para os polimorfismos estudados (*XRCC2* R188H, *XRCC3* T241M, *NBS1* E185Q e *RAD51* 5'UTR), são similares às descritas para outras populações Caucásicas (Kiuru *et al.*, 2008; Kuschel *et al.*, 2002; Lu *et al.*, 2006; Millikan *et al.*, 2005). O desvio ao equilíbrio de Hardy-Weinberg foi apenas observado para o polimorfismo *XRCC2* R188H na população controlo (P = 0,0095).

Os resultados obtidos após análise por regressão logística não revelaram qualquer associação entre os polimorfismos estudados e a susceptibilidade para cancro da mama, uma vez que não existem diferenças significativas na distribuição de genótipos dos diferentes polimorfismos entre as duas populações (Tabela 8.4).

Tabela 8.4 – Distribuição de genótipos e risco para cancro da mama associado aos polimorfismos *XRCC2* R188H, *NBS1* E185Q, *XRCC3* T241M e *RAD51* 5'UTR nas populações de doentes (n = 289) e controlos (n = 548).

GENÓTIPOS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a	OR _{NÃO AJUSTADO} (IC A 95%)	OR _{AJUSTADO} (95% CI) ^b
<i>XRCC2</i> R188H (Ex3+442G>A)					
Arg/Arg	243 (84,1)	445 (81,2)		1,00 (Referência)	1,00 (Referência)
Arg/His	46 (15,9)	103 (18,8)	0,30	0,82 (0,56-1,20)	0,83 (0,56-1,23)
His/His	0 (0)	0 (0)		-	-
<i>NBS1</i> E185Q (Ex5-32C>G)					
Glu/Glu	152 (52,6)	277 (50,5)		1,00 (Referência)	1,00 (Referência)
Glu/Gln	112 (38,8)	219 (40,0)	0,83	0,93 (0,69-1,26)	0,92 (0,67-1,25)
Gln/Gln	25 (8,7)	52 (9,5)		0,88 (0,52-1,47)	0,89 (0,53-1,51)
Glu/Gln + Gln/Gln	137 (47,4)	271 (49,5)		0,92 (0,69-1,22)	0,91 (0,68-1,22)
<i>RAD51</i> 5'UTR (Ex1-59G>T)					
GG	94 (32,6)	168 (30,7)		1,00 (Referência)	1,00 (Referência)
GT	139 (48,3)	275 (50,2)	0,83	0,90 (0,65-1,25)	0,89 (0,64-1,24)
TT	55 (19,1)	105 (19,2)		0,94 (0,62-1,41)	0,93 (0,61-1,42)
GT+TT	194 (67,4)	380 (69,3)		0,91 (0,67-1,24)	0,90 (0,66-1,23)
<i>XRCC3</i> T241M (Ex8-53C>T)					
Thr/Thr	109 (37,7)	178 (32,5)		1,00 (Referência)	1,00 (Referência)
Thr/Met	138 (47,8)	276 (50,4)	0,28	0,82 (0,60-1,12)	0,83 (0,60-1,14)
Met/Met	42 (14,5)	94 (17,2)		0,73 (0,47-1,13)	0,740 (0,46-1,15)
Thr/Met + Met/Met	180 (62,3)	370 (67,5)		0,79 (0,59-1,07)	0,80 (0,59-1,09)

^a Valor de $P \chi^2$ (Ver Materiais e Métodos). ^b OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥ 70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

Contudo, após estratificação da população de acordo com o *status* de amamentação, observou-se que indivíduos que nunca tinham amamentado e que apresentavam um genótipo heterozigótico para o polimorfismo *XRCC2* R188H estavam sujeitos a uma diminuição de risco para cancro da mama [$OR_{Ajustado}=0,45$ (0,22-0,92); ($P=0,03$)] (ver Tabela 8.5). Quando a estratificação da população é feita de acordo com o *status* menopausico, os resultados sugerem que mulheres pós-menopausicas e, simultaneamente, portadoras de pelo menos um alelo variante para o polimorfismo *XRCC3* T241M, apresentam um risco diminuído para cancro da mama [$OR_{Ajustado}=0,67$ (0,47-0,94); ($P=0,04$)] (ver Tabela 8.6).

Tabela 8.5 – Associação dos diferentes polimorfismos *XRCC2* R188H, *NBS1* E185Q, *RAD51* 5'UTR e *XRCC3* T241M e risco para cancro da mama na população de doentes (n=289), de acordo com o *status* de amamentação.

GENÓTIPOS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a	OR _{NÃO AJUSTADO} (IC A 95%)	OR _{AJUSTADO} (IC A 95%) ^b
<i>XRCC2</i> R188H					
Doentes que nunca amamentaram					
Arg/Arg	87 (90,6)	445 (81,2)		1,00 (Referência)	1,00 (Referência)
Arg/His	9 (9,4)	103 (18,8)	0,02	0,45 (0,23-0,92)*	0,45 (0,22-0,92)**
His/His	0 (0)	0 (0)		-	-
Doentes que amamentaram					
Arg/Arg	153 (80,5)	445 (81,2)		1,00 (Referência)	1,00 (Referência)
Arg/His	37 (19,5)	103 (18,8)	0,84	1,04 (0,69-1,59)	1,08 (0,70-1,65)
His/His	0 (0)	0 (0)		-	-
<i>NBS1</i> E185Q					
Doentes que nunca amamentaram					
Glu/Glu	53 (55,2)	277 (50,5)		1,00 (Referência)	1,00 (Referência)
Glu/Gln	37 (38,5)	219 (40,0)	0,51	0,88 (0,56-1,39)	0,82 (0,51-1,30)
Gln/Gln	6 (6,3)	52 (9,5)		0,60 (0,25-1,48)	0,60 (0,24-1,47)
Glu/Gln + Gln/Gln	43 (44,8)	271 (49,5)		0,83 (0,54-1,28)	0,78 (0,50-1,21)
Doentes que amamentaram					
Glu/Glu	98 (51,6)	277 (50,5)		1,00 (Referência)	1,00 (Referência)
Glu/Gln	73 (38,4)	219 (40,0)	0,93	0,94 (0,66-1,34)	0,94 (0,65-1,34)
Gln/Gln	19 (10,0)	52 (9,5)		1,03 (0,58-1,83)	1,06 (0,59-1,89)
Glu/Gln + Gln/Gln	92 (48,4)	271 (49,5)		0,96 (0,69-1,34)	0,96 (0,69-1,34)
<i>RAD51</i> 5'UTR					
Doentes que nunca amamentaram					
GG	31 (32,3)	168 (30,7)		1,00 (Referência)	1,00 (Referência)
GT	46 (47,9)	275 (50,2)	0,92	0,91 (0,55-1,49)	0,91 (0,55-1,51)
TT	19 (19,8)	105 (19,1)		0,98 (0,53-1,82)	1,01 (0,54-1,90)
GT+TT	65 (67,7)	380 (69,3)		0,93 (0,58-1,48)	0,94 (0,58-1,51)
Doentes que amamentaram					
GG	62 (32,8)	168 (30,7)		1,00 (Referência)	1,00 (Referência)
GT	92 (48,7)	275 (50,2)	0,86	0,91 (0,62-1,32)	0,89 (0,61-1,30)
TT	35 (18,5)	105 (19,1)		0,90 (0,56-1,46)	0,91 (0,56-1,49)
GT+TT	127 (67,2)	380 (69,3)		0,91 (0,64-1,30)	0,89 (0,62-1,28)

Tabela 8.5 (Continuação) - Associação dos diferentes polimorfismos XRCC2 R188H, NBS1 E185Q, RAD51 5'UTR e XRCC3 T241M e risco para cancro da mama na população de doentes (n=289), de acordo com o *status* de amamentação.

GENÓTIPOS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a	OR _{NÃO AJUSTADO} (IC A 95%)	OR _{AJUSTADO} (IC A 95%) ^b
XRCC3 T241M					
Doentes que nunca amamentaram					
Thr/Thr	32 (33,3)	178 (32,5)		1,00 (Referência)	1,00 (Referência)
Thr/Met	50 (52,1)	276 (50,4)	0,82	1,01 (0,62-1,63)	1,02 (0,61-1,66)
Met/Met	14 (14,6)	94 (17,1)		0,83 (0,42-1,63)	0,84 (0,42-1,66)
Thr/Met + Met/Met	64 (66,7)	370 (67,5)		0,96 (0,61-1,52)	0,97 (0,61-1,55)
Doentes que amamentaram					
Thr/Thr	75 (39,5)	178 (32,5)		1,00 (Referência)	1,00 (Referência)
Thr/Met	87 (45,8)	276 (50,4)	0,21	0,75 (0,52-1,07)	0,75 (0,52-1,09)
Met/Met	28 (14,7)	94 (17,1)		0,71 (0,43-1,17)	0,71 (0,43-1,18)
Thr/Met + Met/Met	115 (65,7)	370 (67,5)		0,74 (0,52-1,04)	0,74 (0,52-1,05)

^a Valor de P χ^2 (Ver Materiais e Métodos). ^b OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥ 70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

*P_{Não-Ajustado}=0,03 e **P_{Ajustado}=0,03 (Valores de P ajustado segundo modelo de regressão logística).

Os valores de OR associados aos diferentes polimorfismos estudados após estratificação para os outros factores que poderiam ser relevantes para o desenvolvimento de cancro da mama (i.e. idade de menarca, paridade, idade da primeira gravidez) foram determinados sem, no entanto, revelarem qualquer associação com a doença (dados não apresentados).

Tabela 8.6 - Associação dos diferentes polimorfismos *XRCC2* R188H, *NBS1* E185Q, *RAD51* 5'UTR e *XRCC3* T241M e risco para cancro da mama na população de doentes (n=289), de acordo com o *status* menopausico.

GENÓTIPOS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a	OR _{NÃO-AJUSTADO} (IC A 95%)	OR _{AJUSTADO} (IC A 95%) ^b
<i>XRCC2</i> R188H					
Doentes Pré-menopausicas					
Arg/Arg	66 (88,0)	445 (81,2)		1,00 (Referência)	1,00 (Referência)
Arg/His	9 (12,0)	103 (18,8)	0,15	0,59 (0,28-1,22)	0,61 (0,28-1,30)
His/His	0 (0,0)	0 (0)		-	-
Doentes Pós-menopausicas					
Arg/Arg	176 (83,0)	445 (81,2)		1,00 (Referência)	1,00 (Referência)
Arg/His	36 (17,0)	103 (18,8)	0,56	0,88 (0,58-1,34)	0,93 (0,60-1,43)
His/His	0 (0)	0 (0)		-	-
<i>NBS1</i> E185Q					
Doentes Pré-menopausicas					
Glu/Glu	38 (50,7)	277 (50,5)		1,00 (Referência)	1,00 (Referência)
Glu/Gln	29 (38,7)	219 (40,0)	0,94	0,96 (0,58-1,62)	0,95 (0,55-1,64)
Gln/Gln	8 (10,7)	52 (9,5)		1,12 (0,50-2,54)	1,06 (0,44-2,53)
Glu/Gln + Gln/Gln	37 (49,3)	271 (49,5)		1,00 (0,61-1,61)	0,98 (0,59-1,62)
Doentes Pós-menopausicas					
Glu/Glu	114 (53,8)	277 (50,5)		1,00 (Referência)	1,00 (Referência)
Glu/Gln	81 (38,2)	219 (40,0)	0,67	0,90 (0,64-1,26)	0,87 (0,61-1,24)
Gln/Gln	17 (8,0)	52 (9,5)		0,79 (0,44-1,43)	0,80 (0,44-1,47)
Glu/Gln + Gln/Gln	96 (45,7)	271 (49,5)		0,88 (0,64-1,21)	0,86 (0,62-1,19)
<i>RAD51</i> 5'UTR					
Doentes Pré-menopausicas					
GG	24 (32,4)	168 (30,7)		1,00 (Referência)	1,00 (Referência)
GT	35 (47,3)	275 (50,2)	0,90	0,89 (0,51-1,55)	0,87 (0,49-1,57)
TT	15 (20,3)	105 (19,1)		1,00 (0,50-1,99)	0,97 (0,47-2,02)
GT+TT	50 (67,6)	380 (69,3)		0,92 (0,55-1,55)	0,90 (0,52-1,56)
Doentes Pós-menopausicas					
GG	69 (32,5)	168 (30,7)		1,00 (Referência)	1,00 (Referência)
GT	103 (48,6)	275 (50,2)	0,88	0,91 (0,64-1,31)	0,92 (0,63-1,34)
TT	40 (18,9)	105 (19,1)		0,94 (0,59-1,47)	0,94 (0,58-1,51)
GT+TT	143 (67,5)	380 (69,3)		0,92 (0,65-1,29)	0,92 (0,65-1,32)

Tabela 8.6 (Continuação) - Associação dos diferentes polimorfismos *XRCC2* R188H, *NBS1* E185Q, *RAD51* 5'UTR e *XRCC3* T241M e risco para cancro da mama na população de doentes (n=289), de acordo com o *status* menopausico.

GENÓTIPOS	DOENTES N (%)	CONTROLOS N (%)	VALOR DE P ^a	OR _{NÃO-AJUSTADO} (IC A 95%)	OR _{AJUSTADO} (IC A 95%) ^b
XRCC3 T241M					
Doentes Pré-menopausicas					
Thr/Thr	21 (28,0)	178 (32,5)		1,00 (Referência)	1,00 (Referência)
Thr/Met	44 (58,7)	276 (50,4)	0,39	1,35 (0,78-2,35)	1,45 (0,81-2,59)
Met/Met	10 (13,3)	94 (17,1)		0,90 (0,41-1,99)	0,99 (0,43-2,28)
Thr/Met + Met/Met	54 (72)	370 (67,5)		1,24 (0,73-2,11)	1,34 (0,76-2,34)
Doentes Pós-menopausicas					
Thr/Thr	87 (41,0)	178 (32,5)		1,00 (Referência)	1,00 (Referência)
Thr/Met	94 (44,3)	276 (50,4)	0,08	0,70 (0,49-0,99)*	0,68 (0,47-0,97)**
Met/Met	31 (14,6)	94 (17,1)		0,68 (0,42-1,09)	0,64 (0,39-1,04)
Thr/Met + Met/Met	125 (59,0)	370 (67,5)		0,69 (0,50-0,96)†	0,67 (0,47-0,94)††

^a Valor de P χ^2 (Ver Materiais e Métodos). ^b OR's foram ajustados para: idade de diagnóstico (30-49, 50-69 e ≥ 70), sendo o grupo de referência o de menor idade; Hábitos Alcoólicos (consume: nunca, social and regular), os não consumidores são considerados grupo de referência; e hábitos tabágicos (fumadores/não fumadores), não fumadores grupo de referência.

*P_{Não-Ajustado}=0,04 e **P_{Ajustado}=0,03 (Valores de P ajustado segundo modelo de regressão logística).

†P_{Não-Ajustado}=0,03 e ††P_{Ajustado}=0,04 (Valores de P ajustado segundo modelo de regressão logística).

8.4 DISCUSSÃO

Os polimorfismos *XRCC2* (Ex3+442G>A, R188H, rs3218536), *XRCC3* (Ex8-5C>T, T241M, rs861539), *NBS1* (Ex5-32C>G, E185Q, rs1805794) e *RAD51* 5'UTR (Ex1-59G>T, rs1801321) têm sido estudados em diversos estudos caso-controlo tendo em vista a avaliação do seu papel na susceptibilidade individual para o cancro. Os estudos em cancro da mama efectuados e publicados até à data que incluam estes genes têm sido contraditórios (Brooks *et al.*, 2008; Costa *et al.*, 2007; Garcia-Closas *et al.*, 2006; Han *et al.*, 2004a; Hsu *et al.*, 2007; Krupa *et al.*, 2009; Kuschel *et al.*, 2002; Pooley *et al.*, 2008; Popanda *et al.*, 2006; Rafii *et al.*, 2002; Sangrajrang *et al.*, 2007).

A frequência dos diferentes polimorfismos na população controlo Portuguesa é semelhante à reportada para outras populações Caucásicas. O desvio ao equilíbrio de Hardy-Weinberg foi apenas observado para o polimorfismo no gene *XRCC2* (Ex3+442G>A, R188H, rs3218536) para a população controlo. O equilíbrio de Hardy Weinberg depende de uma série de condições sobre a população estudada, incluindo,

por exemplo, o tamanho da população, reprodução aleatória (panmixia), ausência de migração, ausência de deriva genética e ausência de ocorrência de qualquer selecção (Costa *et al.*, 2007). Considerando que não foi encontrado qualquer desvio para todos os outros genótipos, o desvio encontrado no SNP do gene *XRCC2* pode estar relacionado com o acaso ou com a violação de algum destes pressupostos, em vez de estar associado com erros de genotipagem.

Um trabalho prévio mostrou que o polimorfismo do gene *XRCC2* não estava associado com o cancro da mama considerando a doença benigna da mama como um dos factores de risco (Jorgensen *et al.*, 2009). Outros estudos não revelaram qualquer associação deste polimorfismo com o cancro da mama (Breast Cancer Association Consortium, 2006; Brooks *et al.*, 2008; Jorgensen *et al.*, 2009; Kuschel *et al.*, 2002; Millikan *et al.*, 2005; Webb *et al.*, 2005), nem com efeitos colaterais agudos da radioterapia em doentes com cancro da mama (Popanda *et al.*, 2006). No entanto, em diversos outros estudos, a presença do alelo variante parece estar associada com uma diminuição de risco para cancro da mama (Loizidou *et al.*, 2008; Pooley *et al.*, 2008). Os resultados obtidos com este trabalho não revelaram qualquer associação entre o polimorfismo R188H do gene *XRCC2 per se* e o risco para cancro da mama. Contudo, após estratificação da população de acordo com o *status* de amamentação, observou-se que indivíduos que nunca amamentaram e que eram, simultaneamente, portadores de pelo menos um alelo variante deste polimorfismo (188His) apresentam uma diminuição de risco para neoplasia mamária.

Um trabalho de revisão que incluiu 47 estudos desenvolvidos em 30 países, com o envolvimento de cerca de 50.000 mulheres com cancro da mama e 97.000 controlos, sugeriu que a amamentação pode ser responsável por 2/3 da redução estimada para o cancro da mama. Quanto maior a duração do período de amamentação, menor é o potencial risco para cancro da mama. Tem sido estimado que a incidência da neoplasia mamária em países desenvolvidos poderia ser reduzida para menos de metade (de 6,3 a 2,7%) se o período de amamentação fosse superior (Rea, 2004). Sabe-se que o leite materno contém células ductais esfoliadas, e que nestas mesmas células obtidas a partir de amostras de leite materno é possível detectar diversos tipos de aductos de DNA (Thompson *et al.*, 2002). Pela análise destes resultados é razoável assumir que a esfoliação das células ductais como consequência

da amamentação pode remover um significativo número de células com lesão genética, prevenindo assim a sua transformação em células neoplásicas. Neste sentido, o potencial efeito protector da variante alélica do SNP do gene *XRCC2*, em mulheres que nunca amamentaram, pode estar relacionado com uma reparação mais eficiente da lesão de DNA. Na realidade, quando células deficientes em *XRCC2* foram complementadas com a forma comum ou variante do polimorfismo deste gene, observou-se que a presença da forma variante estava associada com um aumento na resistência à cisplatina, mas não à lesão de DNA induzida pela mitomicina C, o que sugere que a forma variante apresenta uma maior capacidade na reparação de determinados tipos de lesão de DNA (Danoy *et al.*, 2007), o que pode explicar os resultados obtidos no presente trabalho quanto ao efeito protector associado à presença da variante polimórfica em mulheres que nunca amamentaram. Embora não existam resultados semelhantes em estudos de associação que incluam este polimorfismo e o risco para cancro da mama, existe um estudo prévio (Capítulo 4) que reporta este mesmo efeito em mulheres que nunca amamentaram mas associado com a presença da variante alélica de um polimorfismo presente numa enzima de detoxificação de espécies reactivas de oxigénio, a MnSOD (Silva *et al.*, 2006b).

Relativamente ao polimorfismo no gene *NBS1* (Ex5-32C>G, E185Q, rs1805794), os resultados obtidos não apresentaram qualquer associação entre o SNP e o risco para cancro da mama, o que está de acordo com os resultados apresentados em estudos anteriores e desenvolvidos nas populações Caucasiana (Forsti *et al.*, 2004; Kuschel *et al.*, 2002) e Chinesa (Zhang *et al.*, 2005), onde não foram encontradas diferenças significativas entre as frequências genótípicas das populações de doentes e controlos. Adicionalmente, não foram estabelecidas quaisquer associações entre a sobrevivência de cancro da mama (Goode *et al.*, 2002) e nem com os efeitos colaterais agudos da radioterapia em doentes com cancro da mama (Popanda *et al.*, 2006). No entanto, Smith *et al.* (Smith *et al.*, 2008) mostraram que existia uma tendência significativa no risco para cancro da mama com o aumento do número de genótipos de risco de *NBS1* 185EQ/QQ numa população Afro-Americana. No estudo publicado por Lu *et al.* (Lu *et al.*, 2006) também foi descrita uma associação entre a forma variante deste polimorfismo no gene *NBS1* e o risco para neoplasia mamária em mulheres brancas não-hispânicas com idade inferior ou igual a 55 anos. Estas discrepâncias entre

os diferentes estudos podem estar associadas com diferentes perfis genéticos, diferentes factores de risco nas diferentes populações e, até mesmo, o tamanho da amostra populacional estudada em cada caso. No entanto, estes resultados não excluem o potencial envolvimento de outros polimorfismos presentes no gene *NBS1* na susceptibilidade individual para esta patologia, como foi descrito por Hsu *et al.* (Hsu *et al.*, 2007). Este último grupo estuda o papel dos genes MRN no risco para cancro, tendo descrito uma ligeira associação no desenvolvimento do cancro da mama apenas para um polimorfismo do gene *NBS1* (rs1805790) em mulheres de Taiwan.

A análise das frequências genótípicas do polimorfismo *RAD51* 5'UTR (Ex1-59G>T, rs1801321) permite concluir, como num outro estudo publicado por Kuschel *et al.* (Kuschel *et al.*, 2002), que não existe qualquer associação entre a presença do alelo variante deste SNP e o risco para cancro da mama. Este polimorfismo foi igualmente estudado num outro trabalho não tendo sido descrita qualquer associação entre o SNP e a sobrevivência ao cancro da mama (Goode *et al.*, 2002), num estudo conduzido com vista a avaliar o efeito de variações da linha germinal em diversos polimorfismos na sobrevivência entre mulheres com cancro da mama.

Relativamente ao polimorfismo no gene *XRCC3* (Ex8-5C>T, T241M, rs861539), os resultados publicados têm sido contraditórios (Costa *et al.*, 2007; Forsti *et al.*, 2004; Garcia-Closas *et al.*, 2006; Lee *et al.*, 2007; Sangrajrang *et al.*, 2007). Num estudo desenvolvido por Krupa *et al.* (Krupa *et al.*, 2009), foi descrito um ligeiro aumento de risco de metástases locais em doentes com cancro da mama e também uma correlação entre este polimorfismo e um outro presente no gene *RAD51* (rs1801320 – 135G>C), o que pode ser avaliado como uma possível indicação de causa poligénica na ocorrência e progressão do cancro da mama. Um outro estudo desenvolvido numa população Portuguesa (Costa *et al.*, 2007) revelou um aumento de risco para cancro da mama não-familiar em mulheres portadoras do alelo variante. Contudo, estes resultados não estão de acordo com os obtidos no presente trabalho. O polimorfismo presente no gene *XRCC3*, como todos os outros SNPs estudados, individualmente não suporta uma associação de risco com cancro da mama esporádico. Não obstante, após estratificação para o *status* menopausico, os resultados obtidos sugerem uma diminuição de risco para neoplasia mamária em mulheres pós-menopausicas portadoras de pelo menos um alelo variante. Alguns estudos referem-se ao alelo

241Met como sendo um alelo de risco para cancro da mama, para além disso este SNP tem sido descrito como responsável por elevados níveis de aductos de grande peso molecular em estudos funcionais (Matullo *et al.*, 2001b), ou como responsável por deficiências mitóticas e diminuição da capacidade de reparação de lesões induzidas por raios-X (Lindh *et al.*, 2006). No entanto, os resultados sugerem que este alelo variante pode estar envolvido numa reparação da lesão de DNA mais eficiente em mulheres pós-menopausicas.

A discrepância de resultados entre os diferentes estudos de associação caso-controlo que consideram os polimorfismos nos genes *XRCC2*, *XRCC3*, *NBS1* e *RAD51* podem ser considerados como variação no perfil genético e nos diferentes tipos de exposição a cancerígenos a que as diferentes populações estudadas estão sujeitas. Outros factores que podem também contribuir para os diferentes resultados baseiam-se em amostras populacionais demasiado pequenas ou derivados de um inadequado controlo a determinados factores confundentes.

Outra linha de pesquisa em estudos de associação que envolvam alelos de baixa penetrância relaciona-se com a avaliação do suposto papel desempenhado pelos microRNAs. As variantes genéticas em genes que codificam para microRNAs podem alterar potencialmente a regulação de genes chave para o cancro da mama afectando assim os níveis de expressão de genes supressores de tumor ou oncogenes e, deste modo, o risco para cancro. Do mesmo modo, SNPs nas regiões 3'UTR de genes alvo para os microRNAs podem apresentar um efeito semelhante. Se todas estas variantes conferem susceptibilidade genética para o cancro da mama ou podem modificar a expressão dos alelos de baixa penetrância permanece por esclarecer. Sendo este o caso, a ocorrência de resultados não-reprodutíveis em diferentes estudos pode, pelo menos em parte, ser explicado pelo papel ainda discreto dos microRNAs (Shen *et al.*, 2009; Tchatchou *et al.*, 2009).

Sumariamente, os resultados obtidos com este trabalho sugerem que os polimorfismos nos genes *XRCC2* (Ex3+442G>A, R188H, rs3218536) e *XRCC3* (Ex8-5C>T, T241M, rs861539) poderão estar envolvidos na susceptibilidade genética para cancro da mama, no entanto, serão necessários mais estudos com populações mais numerosas de forma a confirmar estes resultados.

CAPÍTULO 9

APLICABILIDADE DA TÉCNICA DE *POOLS* DNA

9. APLICABILIDADE DA TÉCNICA DE *POOLS* DE DNA

9.1 INTRODUÇÃO

Os polimorfismos de base única (SNPs) ocorrem frequentemente ao longo do genoma humano e, também por isso, são os tipos de marcadores mais comuns utilizados em análise genética (Mattarucchi *et al.*, 2005; Yang *et al.*, 2006), nomeadamente para estudos epidemiológicos de forma a avaliar qual o papel atribuído às variantes genéticas na etiologia de doenças humanas (Little *et al.*, 2002). A identificação de polimorfismos é assim fundamental para aumentar a compreensão sobre a susceptibilidade genética no sentido de melhorar a análise individual de risco. Consequentemente, serão necessários muito mais estudos de associação que permitam identificar, em genes candidatos, alelos que possam conferir pequenos incrementos na susceptibilidade individual para a doença (Norton *et al.*, 2002).

Os estudos caso-controlado desenvolvidos até à data têm mostrado que o estudo de SNPs apenas aumenta o risco de doença associado a um factor, normalmente, inferior a dois. Por outro lado, sabe-se que a predictibilidade de risco individual pode ser aumentada quando se avalia o efeito atribuído à associação de múltiplos polimorfismos (Tabor *et al.*, 2002), permitindo assim o desenvolvimento de estratégias preventivas a nível populacional. Deste modo, a identificação de variantes alélicas dos genes humanos, a descrição das frequências dessas mesmas variantes em diferentes populações ou grupos étnicos (devido à heterogeneidade do perfil genético), a identificação de doenças influenciadas por essas variantes, e a avaliação da magnitude associada à doença, serão a base da análise genética em doenças humanas (Little *et al.*, 2002; Yang *et al.*, 2006).

De facto, o papel emergente dos polimorfismos genéticos em estudos de associação clínica tem criado a necessidade de metodologias de genotipagem mais eficientes que permitam uma rápida identificação dos polimorfismos mais relevantes no estudo do risco individual para cancro, para aplicabilidade em estudos de associação de larga escala, ou seja que envolva populações numerosas. Uma forma de resolver os problemas de custo, tempo e mão-de-obra envolvidos na genotipagem em larga escala é desenvolver técnicas de análise, não em amostras individuais mas sim,

em *pools* de amostras construídas a partir do DNA de muitos indivíduos (Sham *et al.*, 2002), com a mais valia de serem tratadas como uma amostra única (Chowdari *et al.*, 2007; Shaw *et al.*, 1998; Yang *et al.*, 2006; Zuo *et al.*, 2006). Deste modo, a técnica de *DNA pooling* quando aplicada em estudos caso-controlo pode ser utilizada na pré-selecção de variantes genéticas, potencialmente associadas a uma patologia, que permitam subsequente análise individual útil na abordagem caso-controlo (Rollinson *et al.*, 2004; Zuo *et al.*, 2006).

Na realidade, a ideia de se utilizarem as amostras de DNA em *pool* com o intuito de diminuir a genotipagem não é nova e foi sugerido pela primeira vez, no contexto de ser aplicado a estudos caso-controlo, por Arnheim em 1985 aplicado a um estudo em doentes com diabetes *mellitus* dependente de insulina (Arnheim *et al.*, 1985). Desde então, um determinado número de métodos de genotipagem tem sido utilizado para análise de amostras de DNA em *pool*, incluindo análise de fragmentos de restrição em gel (RFLP), PCR em tempo real (TaqMan), espectroscopia de massa, cromatografia líquida desnaturante de alta resolução (dHPLC), *SnaPshot* e pirosequenciação, sendo que todas estas técnicas apresentam uma precisão de resultados comparável (Chowdari *et al.*, 2007; Le Hellard *et al.*, 2002; Rollinson *et al.*, 2004).

Desta forma, com este trabalho pretendeu testar-se a precisão da análise de amostras em *pool* utilizando a metodologia de RFLP como método de determinação das frequências alélicas. Assim, compararam-se as frequências alélicas de alguns polimorfismos em diferentes genes, previamente determinadas em estudos anteriores (*ERCC2* Lys751Gln, *OGG1* Ser326Cys, *XRCC1* Arg194Trp e Arg399Gln) por genotipagem individual (Silva *et al.*, 2005; Silva *et al.*, 2006a; Silva *et al.*, 2007), com amostras de diferentes *pools* criadas de forma a validar a reprodutibilidade da metodologia.

9.2 MATERIAIS E MÉTODOS

9.2.1 QUANTIFICAÇÃO DAS AMOSTRAS DE DNA E CONSTRUÇÃO DOS DIFERENTES POOLS

As amostras de DNA utilizadas foram as normalmente usadas no laboratório nos estudos de associação de cancro da mama, as quais foram extraídas a partir de 250 µl de sangue total através da utilização de um kit disponível comercialmente e segundo instruções do fabricante (QIAamp DNA mini kit; Qiagen, Hilden, Germany).

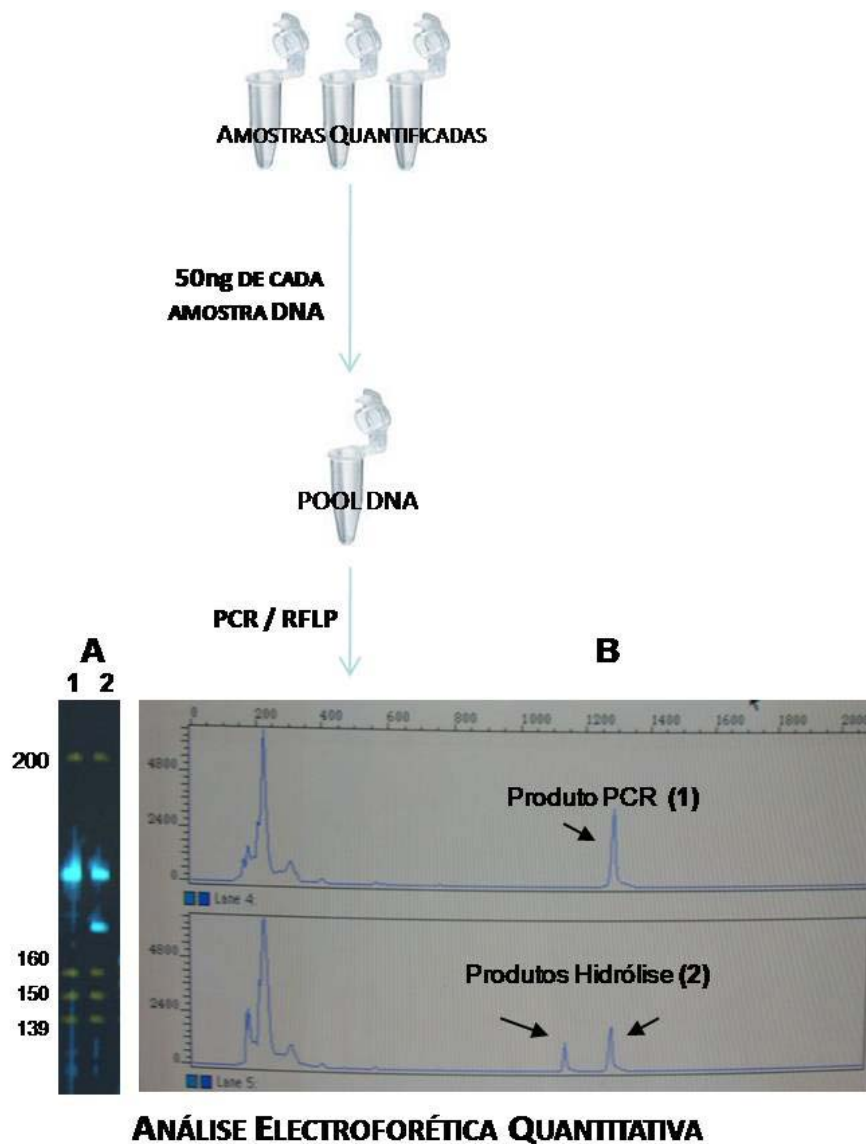
A determinação da concentração das amostras de DNA utilizadas na construção dos *pools* de DNA foi efectuada através da utilização de *PicoGreen dsDNA Quantification Reagent* (Molecular Probes, Eugene, Ore., USA) tendo a leitura sido efectuada num leitor de microplacas *Anthos Zenyth 3100* (Anthos-Labtec Instruments GmbH, Austria). A linearidade do método foi verificada utilizando a curva padrão a partir de uma solução 2 µg/ml de DNA de fago γ stock a 100 µg/ml fornecida pelo kit de quantificação, de acordo com as recomendações do fabricante, previamente à determinação da concentração de DNA. Todos os resultados de quantificação foram considerados desde que apresentassem erros padrão (*Standard Errors*) inferiores a 10%.

Após quantificação, todas as amostras foram diluídas à mesma concentração (ver Tabela 9.1), e subsequentemente misturadas utilizando para tal o mesmo volume de cada amostra individual (Ver Figura 9.1). Para testar a linearidade deste método foram utilizados quatro *pools* (Pool_A até Pool_D, ver Tabela 9.1) construídos com DNA de 20 indivíduos distintos e com diferentes percentagens de cada alelo (90, 80, 70 e 50% do alelo mais frequente, respectivamente) para o polimorfismo presente no gene *ERCC2* Lys751Gln. Seguidamente utilizaram-se os mesmos *pools* e verificou-se se as frequências obtidas através desta metodologia eram concordantes com as frequências obtidas previamente por genotipagem individual do polimorfismo no gene *OGG1* Ser326Cys. Após a optimização inicial, o número de indivíduos que constituíram os diferentes *pools* criados foi sendo aumentado para 50 e 100, a partir de amostras de DNA escolhidas aleatoriamente (Pool_E e Pool_F, respectivamente), e finalmente para dois *pools* correspondentes às populações de controlos e doentes, normalmente

utilizadas nos estudos de associação de cancro da mama, constituídos por 200 indivíduos cada (Pool_H e Pool_I, respectivamente). De salientar ainda um último *pool* (Pool_G) do qual fazem parte todos os indivíduos que constituem a população constituída pelo Índios Xavante (descrição desta população no Capítulo 4, secção 4.2, ponto 4.2.1). Todos os *pools* foram analisados por PCR-RFLP para os polimorfismos estudados em diferentes genes, previamente genotipados em estudos já descritos e cujas frequências alélicas eram já conhecidas, de forma a determinar a precisão deste método.

Tabela 9.1 – Descrição da composição de todos os *pools* construídos.

ID POOLS	AMOSTRAS DNA NO POOL (N)	CONCENTRAÇÃO DAS AMOSTRAS DE DNA EM POOL	POPULAÇÃO
POOL_A	20	2,5 ng/ μ l	Controlos
POOL_B	20		Controlos
POOL_C	20		Controlos
POOL_D	20		Controlos
POOL_E	50	3,5 ng/ μ l	Índios Xavante
POOL_F	100		Índios Xavante
POOL_G	178		Índios Xavante
POOL_H	200	2,5 ng/ μ l	Controlos
POOL_I	200		Doentes Cancro da mama



ANÁLISE ELECTROFORÉTICA QUANTITATIVA

Figura 9.1 – Estratégia utilizada para análise e preparação de *pools* de DNA. Após quantificação, o *pool* é preparado com igual quantidade de cada amostra que o constitui, sendo posteriormente utilizado como única amostra nas reacções de PCR/RFLP. Os produtos de reacção são separados em gel de poliacrilamida e quantificados.

(A) – Produto de PCR (Linha 1) e Produtos de Restrição (Linha 2) – Bandas azuis. As bandas amarelas representam o marcador de pesos moleculares.

(B) – Análise através de electroforetograma, permite a quantificação da frequência de cada alelo por determinação da altura de cada pico que é fornecida pelo software *GENESCAN* (Applied Biosystems, CA, EUA).

9.2.2 GENOTIPAGEM INDIVIDUAL

A genotipagem individual dos polimorfismos utilizados na optimização desta abordagem foi descrita ao longo dos diferentes capítulos que constituem a presente dissertação, pelo que nada será acrescentado neste ponto.

9.2.3 FREQUÊNCIAS ALÉLICAS DETERMINADAS APÓS ANÁLISE DOS POOLS

Como já foi referido, a abordagem escolhida no desenvolvimento da técnica de *pools* assenta na metodologia de RFLP. Este método gera fragmentos específicos e correspondentes aos alelos de cada SNP, que podem ser detectados por electroforese de géis convencionais, o que faz desta abordagem a mais simples deste tipo.

As amostras foram analisadas em gel de acrilamida num sequenciador *ABI 377 automated DNA sequencer* e o software *GENESCAN* (Applied Biosystems, CA, EUA) para quantificação alélica a partir da amplificação dos polimorfismos seleccionados (Rollinson *et al.*, 2004). O processo de amplificação por PCR dos fragmentos contendo os SNP pretendidos é igual ao já descrito com excepção dos *primers reverse* que são marcados na extremidade 5' com um fluorocromo (6-FAM). O gel de acrilamida (4%) foi preparado por adição de 5 ml de acrilamida líquida, 18 g de ureia (Merck, Darmstadt, Germany), 0,5 g de Amberlite, 25 ml de água desionizada, seguindo-se a homogeneização desta mistura. A referida solução é posteriormente filtrada com filtros *Millipore* 0,22 µm e desgaseificada durante 10 min., recorrendo para tal a uma bomba de vácuo. Após filtragem adicionam-se 5,0 ml de tampão TBE 10× e perfaz-se o volume a 50 ml. Posteriormente, são adicionados à mistura os agentes polimerizadores da acrilamida APS (Persulfato de Amónio) (Sigma, St. Louis, USA) e TEMED (N,N,N,N – tetrametiletlenodiamina) (Sigma, St. Louis, USA), 250 µl e 35 µl, respectivamente. A solução foi posteriormente adicionada entre os vidros de sequenciação e após polimerização colocados no *Genescanner*. Antes de se iniciar a corrida electroforética, foi efectuado um *plate check*, com vista a identificar possíveis interferências com a leitura laser. As condições de corrida foram as seguintes 1200V, 40W, PMT: 800V, durante aproximadamente 4 horas.

As frequências alélicas foram depois determinadas a partir dos valores relativos obtidos através da altura dos picos de fluorescência gerados, e extraídos automaticamente do software, e que correspondem a cada alelo detectado na amostra *pool*, como descrito por Shaw et al. (Shaw *et al.*, 1998).

9.3 RESULTADOS

Com o intuito de se avaliar a exactidão da abordagem metodológica conferida pela utilização de amostras em *pool*, utilizaram-se diversos SNPs (*ERCC2* Lys751Gln, *OGG1* Ser326Cys, *XRCC1* Arg194Trp e Arg399Gln) com frequências alélicas conhecidas. A construção de *pools* de frequência alélica conhecida permitiu verificar a linearidade da metodologia.

9.3.1 REPRODUCTIBILIDADE DA TÉCNICA

Este estudo foi realizado de forma a testar a viabilidade da metodologia de *pool* como técnica de projecção na determinação de frequências alélicas de diversos SNPs em diferentes amostras *pool* construídas, comparando-as com as frequências alélicas determinadas a partir da genotipagem individual.

A primeira etapa da optimização desta técnica consistiu na construção de 4 *pools* (Pool_A até Pool_D) baseados nas frequências alélicas previamente obtidas para os polimorfismos do gene *ERCC2* Lys751Gln e do gene *OGG1* Ser326Cys. Para avaliação da reprodutibilidade da técnica estabeleceu-se um *cut-off* de $\pm 5\%$, ou seja, as diferenças obtidas entre os dois métodos de genotipagem não deveriam exceder 5%. Este valor permite suportar alguns dos erros inerentes à preparação da amostra *pool*, bem como algum erro decorrente do processo de restrição pela endonuclease. Deste modo, e seguindo esta padronização, os resultados obtidos entre a genotipagem individual e as amostras *pool* revelaram ser muito heterogéneos (Tabela 9.2). Quando se comparam as frequências alélicas obtidas por genotipagem por PCR/RFLP com a obtida através dos *pools*, para o SNP no gene *ERCC2* apenas um dos *pools* apresenta um valor de *cut-off* superior ao estipulado (Pool_A com $\Delta = 0,086$). Já os valores obtidos para o SNP no gene *OGG1* variaram muito mais, ultrapassando na maioria dos

casos o valor de *cut-off* estabelecido ($\Delta > 0,050$), o que pode estar associado com uma fraca digestão da reacção de restrição, que corresponde à existência de produto de PCR que não sofreu acção da endonuclease de restrição, originando diferenças significativas na discriminação alélica.

Perante estes resultados criaram-se novos *pools* constituídos por um maior número de indivíduos e prosseguiu-se o estudo com a análise de SNPs identificados noutros genes, de forma a permitir averiguar a reprodutibilidade da técnica. O incremento no tamanho das amostras em *pool* mostrou resultados consistentes na determinação das frequências alélicas e respectiva comparação com as obtidas pelo método de genotipagem clássico (ver Tabela 9.2), sendo mais evidentes quando se comparam os valores após incremento, obtidos para as frequências alélicas do SNP no gene *OGG1*. Os resultados obtidos para este SNP, que numa primeira análise apresentavam grande variação, após aumento do número de indivíduos nos novos *pools*, apontaram para uma diminuição do valor de *cut-off* ($\Delta = 0,020$ e $\Delta = 0,040$, para os Pools E e F, respectivamente).

A validação da aplicabilidade desta metodologia continuou com a utilização de um estudo caso-controlo desenvolvido e publicado sobre o papel do gene *XRCC1* na susceptibilidade individual para cancro da mama (Silva *et al.*, 2007), com o qual se pretendeu, uma vez mais, avaliar os *pools* criados. Foram então criados três *pools* com maior número de amostras utilizando para tal amostras das populações descritas no estudo caso-controlo (Silva *et al.*, 2007), o que deu origem a um *Pool* constituído por Índios Xavante (*Pool_G*), um *Pool* de controlos (*Pool_H*) e um *Pool* de doentes (*Pool_I*) com cancro da mama. A determinação das frequências alélicas em cada um destes *pools* é concordante com os resultados obtidos através da genotipagem individual.

Os coeficientes de variação (desvio padrão/média) para os *pools* testados, relativamente às frequências obtidas, para o polimorfismo Arg194Trp foram de 0,03 para o *Pool_G*; 0,02 para o *Pool_H* e 0,01 para o *Pool_I*. Para o SNP Arg399Gln os coeficientes de variação obtidos, para os mesmos *pools*, foram 0,07; 0,03 e 0,04, respectivamente. Depreende-se que a percentagem de variação não excede os 5% o que valida a reprodutibilidade desta metodologia, à excepção do *Pool_G* para o SNP Arg399Gln que apresenta uma variação de 7%. O que é corroborado pela interpretação dos valores de *cut-off* (Δ apresentado na Tabela 9.2). Contudo, o *Pool_H*

para o SNP Arg194Trp do gene *XRCC1* apresenta um valor *borderline*, devendo estar relacionado com erros inerentes já referidos.

9.4 DISCUSSÃO

Com este trabalho conseguiu demonstrar-se que a técnica de *pools* de DNA tem potencial para estimar frequências alélicas com exactidão, reprodutibilidade e sensibilidade, o que significa que esta metodologia representa uma ferramenta prática e eficiente com aplicabilidade na selecção de polimorfismos relevantes para os estudos de associação. Outro aspecto relevante associado a esta técnica está relacionado com o *status* laboratorial e económico inferior quando comparado com a análise individual das amostras, o qual aumenta significativamente com o tamanho da amostra em estudo (Chowdari *et al.*, 2007; Gaustadnes *et al.*, 2006; Hoogendoorn *et al.*, 2000; Norton *et al.*, 2002; Norton *et al.*, 2004; Rollinson *et al.*, 2004).

Mesmo existindo um conjunto de ensaios de genotipagem quantitativa que podem ser utilizados em estudos que utilizem amostras em *pool* (*primer extension*, dHPLC, quebra no local do SNP, amplificação com *primers* de alelo específico (ASO), hibridação em *microarrays*, PCR em tempo real, etc), a técnica escolhida foi a que se baseia nos princípios de RFLP (Casado-Diaz *et al.*, 2007; Gaustadnes *et al.*, 2006; Hoogendoorn *et al.*, 2000; Rollinson *et al.*, 2004; Sham *et al.*, 2002), uma vez que assim não implicaria a aquisição de novos equipamentos. O maior problema inerente a esta técnica reside na possibilidade de ocorrência de digestão incompleta. Qualquer tendência de digestão parcial resulta numa avaliação errada e sistemática do alelo que corresponde ao produto de PCR que não foi clivado (Sham *et al.*, 2002). Contudo, todos os estudos desenvolvidos até à data e que utilizam como ferramenta a abordagem por *pool*, mesmo em diferentes tipos de ensaios de genotipagem, são consistentes e mostram exactidão de resultados (Chowdari *et al.*, 2007; Hoogendoorn *et al.*, 2000; Le Hellard *et al.*, 2002; Sham *et al.*, 2002; Yang *et al.*, 2006).

Tabela 9.2 - Relação entre a altura relativa dos picos e a percentagem das frequências alélicas para o alelo mais comum, para cada um dos SNPs, obtidas para os diferentes *pools* (Δ corresponde à diferença entre o valor da frequência alélica obtida por genotipagem individual e o valor obtido através dos *pools*).

ID Pools	n	ERCC2 Lys751			OGG1 Ser326			XRCC1 Arg194			XRCC1 Arg399		
		GENOTIPAGEM PCR/RFLP	POOL DNA (\pm DP)	Δ	GENOTIPAGEM PCR/RFLP	POOL DNA (\pm DP)	Δ	GENOTIPAGEM PCR/RFLP	POOL DNA (\pm DP)	Δ	GENOTIPAGEM PCR/RFLP	POOL DNA (\pm DP)	Δ
POOL_A	20	90,0	81,4 \pm 5,1	0,086	72,5	75,0 \pm 0,4	-0,025						
POOL_B	20	80,0	78,2 \pm 1,8	0,018	70,0	77,3 \pm 1,6	-0,073						
POOL_C	20	70,0	69,4 \pm 7,1	0,006	75,0	80,2 \pm 2,2	-0,052						
POOL_D	20	50,0	49,7 \pm 5,4	0,003	77,5	70,8 \pm 5,3	0,067						
POOL_E	50	80,0	79,9 \pm 3,0	0,001	73,0	75,0 \pm 0,4	0,02	76,0	71,3 \pm 4,4	0,047	87,0	88,3 \pm 2,4	0,013
POOL_F	100	77,5	77,3 \pm 0,8	0,002	78,0	82,0 \pm 0,8	0,04	78,0	75,6 \pm 3,7	0,024	85,0	84,7 \pm 3,5	0,003
POOL_G	178							74,0	74,4 \pm 2,2	0,004	83,0	78,1 \pm 5,5	0,049
POOL_H	200							91,5	86,1 \pm 1,4	0,054	63,0	63,7 \pm 1,8	0,007
POOL_I	200							93,8	89,3 \pm 0,9	0,045	67,8	67,4 \pm 2,4	0,004

A abordagem por *pools* de DNA tem sido cada vez mais utilizada, aplicada principalmente na análise de SNPs e de haplótipos, como promessa de que esta metodologia possa acelerar a descoberta de genes associados com doenças poligénicas e de fenótipos complexos em doenças Mendelianas (Sham *et al.*, 2002). É também uma ferramenta muito útil para os estudos de associação *Genome Wide* (GWAS), já que possibilitam a determinação das frequências alélicas, em vez de genótipos distintos, permitindo analisar e classificar a probabilidade de associação com uma doença. A aplicação desta técnica em estudos GWAS possibilita a redução dos custos inerentes a estudos desta dimensão que são, na maioria das vezes, imperativos (Bosse *et al.*, 2009).

A utilização de resultados obtidos previamente através de um estudo caso-controlo (Silva *et al.*, 2007), permitiu estabelecer uma avaliação global entre os dois métodos de genotipagem. Os dados publicados em Silva *et al.* (Silva *et al.*, 2007), quando analisados individualmente para cada SNP estudado não revelaram qualquer associação entre eles e a susceptibilidade individual para cancro da mama, como foi discutido anteriormente no Capítulo 5 desta dissertação. Todavia, após estratificação da população de acordo com o *status* menopausico constatou-se que, mulheres pós-menopausicas apresentavam um risco aumentado para a neoplasia mamária associado a um genótipo específico de cada SNP (ver Capítulo 5, Tabelas 5.5 e 5.6). É então notório, que a abordagem por *pools* de DNA não permite a estratificação genotípica da amostra, mas sim uma contextualização global sobre qual a distribuição dos alelos de cada polimorfismo em cada população estudada, averiguando eventuais diferenças entre estas populações.

Por sua vez, a genotipagem individual permite uma abordagem muito mais direccionada e específica, sendo possível obter frequências alélicas e genotípicas para as populações em estudo. Quando se analisam os resultados obtidos por genotipagem individual (Silva *et al.*, 2007) e se avaliam as frequências alélicas obtidas para o conjunto total de indivíduos envolvidos nas populações em estudo, obtém-se para a população de doentes e para o alelo Arg194 do gene *XRCC1* uma frequência alélica de 90,7%, sendo na população controlo de 92,8%. Para o alelo Arg399 do mesmo gene, as frequências nas populações de Doentes e Controlos são de cerca de 68,0% e 65,0%, respectivamente. O que revela que, na globalidade, as populações não são muito

diferentes na sua distribuição alélica, já que quando comparadas com as frequências obtidas nas amostras *pool*, apresentadas na Tabela 9.2, as diferenças não são muito relevantes, sendo os valores de *cut-off* inferiores aos 5% estabelecidos. De salientar que os *pools* foram construídos com DNA de 200 indivíduos seleccionados aleatoriamente e a genotipagem individual foi efectuada na amostra total de indivíduos que constituem cada população. Assim se estabelecem as diferenças na interpretação de resultados obtidos pelas duas metodologias.

É portanto possível, e uma vez mais, num contexto global, utilizar-se a técnica de *pools* de DNA na pré-selecção de SNPs de interesse para estudos de associação. Pode ser uma ferramenta igualmente interessante na identificação de diferenças relacionadas com o perfil genético das populações. Os valores obtidos pelas frequências alélicas permitem comparar as frequências obtidas com as publicadas previamente noutros estudos de associação (Huang *et al.*, 2009; Li *et al.*, 2009; Saadat e Ansari-Lari, 2009). Esta comparação entre as diferenças alélicas obtidas e as descritas favorecem a caracterização populacional relativamente à distribuição dos alelos em estudo.

A metodologia de *pools* de DNA apresenta, contudo, algumas limitações quando comparada com os métodos de genotipagem individual. De acordo com a plataforma de genotipagem escolhida na determinação das frequências alélicas é possível encontrarem-se variações na precisão das frequências estimadas utilizando *pools* de DNA, incluindo tecnologia de elevado desempenho como *arrays* de genotipagem (Jawaid e Sham, 2009). No entanto, a montagem da técnica acarreta sempre alguns erros experimentais inerentes (erros de pipetagem durante a construção dos *pools*, amplificação preferencial e erros de estimativa) que influenciam a determinação das frequências alélicas.

Adicionalmente, os resultados obtidos através dos *pools* resultam na perda de capacidade para se estudarem sub-fenótipos, haplótipos e modelos genéticos específicos bem como empreender estudos de interacção gene-gene ou gene-ambiente. No entanto, e numa tentativa de se obter mais informação sobre as diferenças entre as populações, a construção de diferentes *pools*, ou seja *sub-pools* derivados dos *pools* populacionais, onde se agrupariam os indivíduos de forma estratificada e não aleatória, de acordo com o tipo de estudo que se pretendesse

efectuar, poderia ser uma abordagem a considerar. Todavia, e uma vez mais, os resultados obtidos com esta estratificação iriam reflectir apenas e só a distribuição alélica para cada *pool*. Ou seja, esta técnica contudo, não permite a detecção de estratificação e ajustes nas populações que constituem os *pools*, o que obriga a que a referida estratificação populacional seja efectuada num número limitado de indivíduos previamente genotipados individualmente (Bosse *et al.*, 2009; Jawaid e Sham, 2009).

Contudo, e independentemente de todas as limitações, a técnica de *pools* de DNA apresenta uma alternativa atractiva quando se equacionam os gastos inerentes a estudos de associação.

Em conclusão, o método pode ser aplicado com garantia de sucesso na identificação das principais associações genéticas com a doença, sendo uma abordagem eficaz na identificação de SNPs de interesse acrescido para estudos de associação efectuados posteriormente por genotipagem individual importantes na susceptibilidade genética para a doença.

CAPÍTULO 10

CONSIDERAÇÕES FINAIS

10. CONSIDERAÇÕES FINAIS

O cancro da mama é a doença oncológica mais frequente nas mulheres, sendo a primeira causa de morte por patologia neoplásica no sexo feminino (Bastos *et al.*, 2007). No entanto, nas últimas três décadas tem-se assistido a uma diminuição na taxa de mortalidade associada a esta neoplasia. Esta evidência está relacionada com a crescente disponibilidade de campanhas de sensibilização e métodos de detecção aos quais se aliam terapêuticas eficazes cada vez mais ajustados ao perfil clínico das pacientes. Contudo, mesmo tratando-se de uma patologia com uma incidência elevada, as causas exactas que contribuem para o seu desenvolvimento e progressão permanecem na generalidade desconhecidas.

A neoplasia mamária apresenta uma vertente hereditária, na qual se incluem apenas 5 a 10% dos casos, casos esses que estão relacionados com mutações em genes chave (*BRCA1* e *BRCA2*), e ainda uma vertente multifactorial, onde se incluem todos os outros casos designados de casos esporádicos. Os casos esporádicos podem ser associados com a susceptibilidade individual para a doença, ou seja com a variabilidade genética encontrada em genes críticos potencialmente envolvidos em diferentes vias mecanicistas que possam influenciar o desenvolvimento da neoplasia mamária.

O principal objectivo de qualquer forma de vida é transmitir o seu material genético, intacto e inalterado, à geração seguinte. O que deverá ser alcançado apesar das constantes lesões induzidas por agentes endógenos e exógenos face ao DNA. Para evitar tais lesões as células desenvolveram diversos sistemas de detecção da lesão de DNA, sinalizando a sua presença com o intuito de mediar a sua reparação. Tais respostas, que reflectem um impacto relevante num amplo conjunto de eventos celulares, são biologicamente significantes uma vez que podem ser úteis na prevenção de algumas doenças (Jackson e Bartek, 2009), nomeadamente no cancro da mama, a nível populacional. Com a recente sequenciação do genoma humano, o estudo e a descoberta de polimorfismos genéticos emerge como uma das áreas de investigação da genética moderna (Singh *et al.*, 2008), na medida em que impulsiona os estudos de associação genética (Gaspar *et al.*, 2006).

Nos últimos anos, o estudo de genes polimórficos tem sido a abordagem principal e mais frequente na descrição das diferenças inter-individuais e suas implicações na predisposição para o cancro. Diversos estudos desenvolvidos até ao momento demonstram que a avaliação de um único polimorfismo está apenas associada a uma pequena variação do factor de risco para a doença, no entanto, a predictibilidade do risco individual pode ser aumentada por associação de múltiplos polimorfismos (Gaspar *et al.*, 2006).

Assim, foi objectivo primordial deste projecto o desenvolvimento de estudos caso-controlo para avaliar o efeito dos polimorfismos genéticos presentes em genes intervenientes em alguns mecanismos celulares potencialmente relacionados com o desenvolvimento da neoplasia mamária. Na tentativa de se abranger um grande conjunto de vias celulares potencialmente envolvidas na cancerigénese mamária foram estudados polimorfismos em genes envolvidos na via de destoxificação de xenobióticos, e o efeito de polimorfismos em genes cuja acção principal afecta a reparação de diferentes tipos de lesão induzida por diferentes agentes cancerígenos. Os estudos foram efectuados na população Portuguesa tendo sido utilizadas duas populações, uma constituída por doentes com cancro da mama e outra por indivíduos controlo saudáveis recrutados de um hospital de referência da área de Lisboa.

Para além dos compostos exógenos a que se está expostos, também endogenamente, como produtos do metabolismo celular, são produzidos compostos com capacidade lesiva para o DNA, nomeadamente metabolitos derivados do stress oxidativo e/ou metabolitos derivados do metabolismo oxidativo de estrogénios. Estes metabolitos podem estar envolvidos no desenvolvimento do cancro da mama, sendo portanto relevante o estudo de proteínas relacionadas com a destoxificação destes mesmos cancerígenos, nomeadamente de proteínas da família das GSTs (Crooke *et al.*, 2006; Hayes *et al.*, 2005; Torresan *et al.*, 2008). Dentro desta grande família de GSTs, os genes mais estudados na susceptibilidade para o cancro da mama são, sem dúvida, os genes *GSTM1*, *GSTT1* e *GSTP1* (Egan *et al.*, 2004; McCarty *et al.*, 2009; Saxena *et al.*, 2008; Spurdle *et al.*, 2007; Steck *et al.*, 2007; Syamala *et al.*, 2008; Torresan *et al.*, 2008; Unlu *et al.*, 2008). Contudo, ainda não existe concordância quanto à potencial contribuição destes polimorfismos no desenvolvimento de cancro da mama, devido à falta de concordância dos resultados publicados. A dimensão das populações

envolvidas em cada estudo, falhas no ajuste dos potenciais factores confundentes, nomeadamente estratificação étnica devido ao padrão genético, entre outros factores, poderão estar na origem da disparidade de resultados.

Todavia, outros genes pertencentes à família das GSTs foram ainda pouco abordados não se podendo rejeitar a hipótese do seu potencial envolvimento no cancro da mama, uma vez que os perfis funcionais inerentes a cada GST podem correlacionar-se com esta neoplasia. A classe *Alpha* das GST é altamente expressa pelas glândulas supra-renais e ovários e também pela placenta, durante a gravidez, estando, então, envolvida na síntese de hormonas esteróides, de onde se salientam as hormonas sexuais femininas (estrogénio e progesterona). O envolvimento desta classe na defesa contra o stress oxidativo fez dela um alvo de estudo quanto ao seu envolvimento no cancro da mama. Deste modo, o Capítulo 3 referiu-se, ao estudo de três polimorfismos presentes no gene *GSTA2* (Pro110Ser; Ser112Thr; Glu210Ala). No entanto, os resultados obtidos não revelaram qualquer associação entre os três SNPs estudados e a susceptibilidade individual para cancro da mama (Silva *et al.*, 2009), o que é concordante com um estudo muito recentemente publicado (Andonova *et al.*, 2009).

O envolvimento da exposição a hormonas sexuais, nomeadamente estrogénio e metabolitos derivados do seu metabolismo, na neoplasia da mama é cada vez mais evidente. Deste modo, não faria sentido não se abordar esta vertente num projecto dedicado ao cancro da mama. Assim, no Capítulo 4 pretendeu avaliar-se o efeito dos polimorfismos presentes em genes envolvidos na biossíntese de estrogénios (*CYP17*, *HSD17 β 1*) e na inactivação de metabolitos reactivos derivados da via biossintética (*COMT* e *MnSOD*). A abordagem adoptada neste capítulo incluiu a avaliação de uma terceira população caracterizada por não apresentar casos de cancro de mama. Trata-se de um isolado populacional constituído por índios (Índios Xavante) residentes numa reserva situado em Mato Grosso, Brasil.

Nesta fase, pretendeu-se avaliar qual seria a grande diferença genética, se é que existia, entre a população indígena e as diferentes populações onde estes polimorfismos já tivessem sido estudados. Determinaram-se as diferenças alélicas para cada polimorfismo seleccionado, na população constituída pelos Índios Xavante e compararam-se com as mesmas frequências descritas em diferentes populações

publicadas até à data. Os resultados obtidos revelaram existir diferenças, apenas, para o polimorfismo do gene *MnSOD* (Val16Ala). Perante esta evidência, procedeu-se ao estudo de associação caso-controlo para a população Portuguesa apenas para este gene. As diferenças genótípicas entre Xavantes e população Portuguesa mantiveram-se significativamente diferentes. Contudo, o estudo de associação não revelou qualquer relação entre o polimorfismo e a susceptibilidade individual para cancro da mama. Todavia, após estratificação por *status* de amamentação (mulheres que amamentaram *versus* mulheres que nunca amamentaram), observou-se uma diminuição de risco para cancro da mama em mulheres que nunca tinham amamentado e que apresentassem pelo menos um alelo variante (Silva *et al.*, 2006b). Pensa-se que a amamentação estará associada à esfoliação das células ductais o que provocará a remoção de um número significativo de células com lesão genética, evitando, assim, a sua transformação em células neoplásicas (Thompson *et al.*, 2002). De facto, há evidências que associam o período de amamentação como sendo inversamente proporcional ao risco das mulheres que amamentam desenvolverem cancro da mama (Rea, 2004). Contudo, o efeito protector descrito poderá estar concomitantemente relacionado com uma maior capacidade de destoxificação por parte da enzima superóxido dismutase, já que a presença do alelo variante (Ala) poderá alterar a actividade catalítica da enzima (Sutton *et al.*, 2005) promovendo o referido efeito. É importante referir que os resultados obtidos e publicados (Silva *et al.*, 2006b) foram utilizados por outros autores na elaboração de estudos de meta-análise (Bag e Bag, 2008; Wang *et al.*, 2009). O estudo de meta-análise desenvolvido por Bag e Bag (Bag e Bag, 2008), não encontra qualquer associação entre os possíveis genótipos e o risco de cancro. Já a meta-análise desenvolvida por Wang *et al.* (Wang *et al.*, 2009) descreve um aumento de risco para mulheres pré-menopausicas e portadoras do alelo variante (Ala). Contudo, os critérios de estratificação utilizados pelos diferentes estudos revelaram-se inconsistentes impossibilitando interacções com potência estatística significativa, pelo que os resultados são, uma vez mais, inconclusivos.

Com o capítulo 5 iniciou-se a temática dedicada à ubiquidade associada aos sistemas de reparação da lesão genética no DNA. Estes sistemas têm sido alvo de grande interesse na temática da cancerigénese, e em particular no cancro da mama, uma vez que as evidências apontam para um défice na capacidade de reparação de

lesões de DNA em doentes com esta neoplasia. Tem sido referido ao longo desta dissertação que alguns compostos derivados do metabolismo celular endógeno, nomeadamente espécies reactivas de oxigénio, têm capacidade de lesar as células e deste modo aumentar a instabilidade genómica favorecendo o desenvolvimento do cancro da mama. As lesões provocadas por estes metabolitos vão dar origem a diferentes lesões algumas das quais podem e são reparadas pela via de reparação por excisão de bases (BER). Algumas das proteínas intervenientes nesta via são polimórficas, pelo que o estudo de alguns dos seus polimorfismos fez parte deste projecto. A abordagem foi idêntica à descrita no capítulo anterior, tendo o estudo de associação incidido sobre o efeito de dois polimorfismos identificados para o gene *XRCC1* na susceptibilidade individual para cancro da mama na população Portuguesa. Os resultados obtidos para ambos os SNPs (Arg194Trp e Arg399Gln) individualmente não revelaram qualquer associação entre estas variações polimórficas e a susceptibilidade individual para cancro da mama. Contudo, após estratificação para o *status* menopausico observou-se que em mulheres com idade de menopausa compreendida entre os 45 e os 54 anos apresentavam um aumento de risco para neoplasia mamária na presença de pelo menos um alelo variante do SNP Arg194Trp. Para o polimorfismo Arg399Gln foi descrito o mesmo efeito de aumento de risco em mulheres com genótipo homocigótico para a variante e com idade de menopausa superior a 55 anos. É importante referir que recentemente os dados obtidos neste estudo de associação foram incluídos em três estudos de meta-análise (Huang *et al.*, 2009; Li *et al.*, 2009; Saadat e Ansari-Lari, 2009). Todos os estudos revelaram a potencial associação entre o SNP Arg399Gln e o aumento de risco para cancro da mama em diferentes populações, salientando-se, uma vez mais, a importância do perfil genético de cada grupo étnico. A meta-análise desenvolvida por Huang *et al.* (Huang *et al.*, 2009) descreve um aumento de risco para neoplasia mamária em mulheres portadoras do genótipo homocigótico para a variante (Gln/Gln) do SNP Arg399Gln após estratificação por etnia, especialmente quando se consideram as populações Asiática e Africana. Resultados semelhantes foram descritos por Li *et al.* (Li *et al.*, 2009), associado ao genótipo (Gln/Gln) do mesmo SNP com um aumento de risco para cancro da mama em populações Asiáticas. Da mesma forma, o estudo desenvolvido por Saadat e Ansari-Lari (Saadat e Ansari-Lari, 2009) apresenta, uma vez

mais, uma associação do alelo Gln399 com aumento de risco para cancro da mama também em populações Asiáticas. Da comparação dos resultados obtidos para a população Portuguesa (Silva *et al.*, 2007) com os estudos meta-análise publicados, todos as conclusões parecem indicar no sentido de relacionar a presença do alelo Gln399 com aumento de risco para neoplasia mamária. Todos os estudos de meta-análise fazem referência à potencial discrepância de estratificação estabelecida pelos diferentes autores incluídos como potencial factor com capacidade de gerar resultados heterogêneos e conduzir a resultados contraditórios. No entanto, as evidências não podem ser negligenciadas e mais estudos serão imperativos, no sentido de averiguar com segurança a veracidade dos resultados publicados até à data.

Mantendo-se a temática da reparação de DNA, no Capítulo 6 estudaram-se variações polimórficas em genes envolvidos na via de reparação por excisão de nucleótidos (NER). Agentes cancerígenos como aductos de elevado peso molecular ou lesões induzidas por radiação ionizante são alguns dos tipos mais comuns de lesões reparadas por NER, tratando-se de agentes que podem influenciar a etiologia da neoplasia da mama, os genes que codificam para proteínas desta via poderiam exercer uma função na susceptibilidade individual. Contudo, os resultados obtidos não revelaram existir qualquer associação entre os polimorfismos do gene *ERCC2* seleccionados (Ile199Met; His201Tyr; Asp312Asn e Lys751Gln), quando avaliados individualmente, e o cancro da mama na população Portuguesa. Todavia, após combinação dos diferentes génotipos obtidos para os polimorfismos Asp312Asn e Lys751Gln, encontrou-se um significativo aumento de risco em mulheres portadoras dos génotipos Asp/Asn + Lys/Lys, sendo que para todas as outras combinações possíveis não se encontraram associações. No entanto, estes resultados podem estar relacionados com um erro estatístico tipo I, já que a estratificação do génotipo de risco inclui um número reduzido de indivíduos (n=19).

A instabilidade genómica característica dos processos de cancerigénese pode dever-se a erros recorrentes dos processos de replicação e recombinação de DNA. Ao não ocorrer o correcto processamento destes erros e consequente reparação, desencadeada pela via *mismatch* (MMR) as células ficam expostas a uma taxa de mutação espontânea muito elevada, podendo acelerar o desenvolvimento neoplásico. No Capítulo 7, e uma vez que a informação sobre o efeito de polimorfismos em genes

envolvidos nesta via de reparação e o cancro da mama é escassa, avaliaram-se 9 polimorfismos presentes em 7 genes MMR e a sua potencial associação com o cancro da mama.

Os resultados obtidos, quando analisados individualmente, revelaram uma potencial redução de risco na susceptibilidade individual para cancro da mama associado ao polimorfismo identificado no gene *MLH3* (Leu844Pro) em portadores dos genótipos heterozigótico e homozigótico para a variante. Todos os outros SNPs, individualmente, não revelaram qualquer associação. Sendo a via MMR, uma via em que intervêm diversas proteínas heterodiméricas, foi também objecto de análise as possíveis interacções gene-gene. Assim, os resultados obtidos revelaram duas interacções entre genes desta via e a susceptibilidade para cancro da mama. Uma das interacções está associada com uma diminuição de risco para esta neoplasia e corresponde à interacção *MSH3* Ala1045Thr/*MSH6* Gly39Glu (AA/TC). A segunda interacção descrita (*MSH4* Ala97Thr/*MLH3* Leu844Pro) revelou, por sua vez, um potencial aumento de risco para cancro da mama (Conde *et al.*, 2009). Como já foi referido, não existem muitos estudos de associação em cancro da mama e MMR, pelo que os resultados obtidos necessitarão de ser suportados por outros estudos, preferencialmente, em que a amostra populacional seja maior.

A influência de factores ambientais no desenvolvimento do cancro da mama, nomeadamente derivado da exposição à radiação ionizante, é responsável pela aquisição de lesões no DNA, sendo neste caso concreto responsável pelas quebras de cadeia dupla. No capítulo 8 foram estudados polimorfismos identificados em genes envolvidos no processo de reparação por recombinação homóloga. Esta via de reparação tem sido alvo de estudo em diferentes tipos de cancro (Bastos *et al.*, 2009; Dhillon *et al.*, 2009), não sendo o cancro da mama excepção. No entanto, os resultados publicados até à data têm-se revelado contraditórios. O trabalho descrito na presente dissertação, após avaliação individual dos polimorfismos seleccionados, não revelou, de *per se*, associação entre as variações polimórficas e a susceptibilidade individual para cancro da mama. Todavia, após estratificação de acordo com o *status* de amamentação e o *status* menopausico revelaram um efeito associativo com a neoplasia mamária. Em mulheres portadoras do alelo variante para o SNP identificado no gene *XRCC2* (Arg118His) e que nunca tivessem amamentado foi detectado um

efeito associado a uma diminuição de risco. Efeito semelhante foi descrito em Silva *et al.* (Silva *et al.*, 2006b). Também um efeito protector foi descrito para mulheres pós-menopausicas e portadoras de pelo menos um alelo variante do SNP presente no gene *XRCC3* (Thr241Met) (Silva *et al.*, 2010), o que poderá sugerir o envolvimento destes genes na susceptibilidade individual para cancro da mama.

Os estudos de associação envolvem a genotipagem individual de inúmeras amostras, o que representa um razão custo/mão-de-obra muito elevada. Deste modo, o desenvolvimento de técnicas que permitam uma rápida identificação das variações polimórficas mais relevantes no estudo do risco individual associado a uma patologia é crucial. No Capítulo 9 foi descrito o procedimento da abordagem por *pools* de DNA. Esta é uma técnica que permite determinar as frequências alélicas para um determinado polimorfismo, funcionando como pré-selecção das variantes alélicas que possam ser mais relevantes de serem abordadas em estudos de associação. Tem por base a construção de uma amostra constituída por uma mistura equitativa de DNA de diferentes indivíduos que constituem as diferentes populações em estudo. A utilização desta técnica permitir-nos-ia seleccionar de um conjunto de SNPs de potencial interesse aqueles que, e através de uma reacção simples e económica, revelassem diferenças alélicas significativas quando comparadas com as descritas na bibliografia disponível. Deste modo, os estudos de associação seriam mais direccionados. A abordagem adoptada, utilizando equipamentos disponíveis neste Departamento, permitiu obter resultados que atingiram a reprodutibilidade de resultados quando se compararam os resultados gerados pelos *pools* com os gerados, previamente, por genotipagem individual. Contudo, esta técnica não dispensa os estudos caso-controlo, apenas e só é útil como ferramenta de pré-selecção cujo resultado é expresso através das percentagens alélicas atribuídas às variações polimórficas pretendidas.

A título de conclusão, os resultados obtidos e que compõe esta dissertação sugerem que algumas das variações polimórficas estudadas e que fazem parte de diferentes vias mecanicistas, poderão constituir um importante factor de risco no desenvolvimento de cancro da mama. Pelo que, a avaliação de outros polimorfismos presentes tanto nas vias descritas, como em outras vias que seensem relevantes serão de igual importância. Dada a grande disparidade de resultados publicados até à data, para uma correcta validação dos resultados obtidos neste estudo, é necessário

realizar estudos em outras populações e, se possível, com um maior número de indivíduos, para que os resultados obtidos se possam validar e deste modo confirmar se os polimorfismos identificados estão realmente associados à susceptibilidade individual para neoplasia mamária.

A utilização de estudos de associação tem sido criticada por diversas razões, nomeadamente pelo facto da determinação retrospectiva de uma associação entre ser portador de um alelo ou genótipo e uma dada patologia, não querer necessariamente dizer que exista uma ligação etiológica entre eles. Apenas estudos prospectivos (estudos *cohort*) em grandes populações primariamente saudáveis podem fornecer uma evidência directa e fidedigna. Existem também várias outras formas de aumentar a credibilidade entre a associação de genótipos polimórficos e fenótipos, sendo a mais utilizada a meta-análise que permite separar as associações ao acaso das evidências fidedignas (Tempfer *et al.*, 2006).

Pode ainda perspectivar-se a continuação deste trabalho pelo estudo de vias celulares não menos importantes, nomeadamente a via de reparação não homóloga (NHEJ), tendo sido a única a não ser incluída neste estudo, e aquela sobre qual há menos estudos publicados. Não podem ser esquecidas, contudo, ainda outras vias ou mecanismos que podem estar envolvidos na etiologia do cancro da mama, nomeadamente a via apoptótica (Domingues, 2008), e que por não fazerem parte deste trabalho não constaram da revisão bibliográfica desta dissertação.

O impacto dos estudos de associação pode ser aumentado se se progredir para o estudo da genómica funcional dos alelos envolvidos na susceptibilidade para a neoplasia em estudo, o que poderá ser interpretado como mais uma abordagem futura na linha de continuação deste trabalho. Contudo, esta abordagem funcional exige sistemas flexíveis e robustos que sejam capazes de explorar diferentes actividades dos genes, de forma a justificar os papéis atribuídos às variantes alélicas em determinados tipos de cancro, e concretamente no cancro da mama (Rueff *et al.*, 2002).

REFERÊNCIAS BIBLIOGRÁFICAS

Referências Bibliográficas

- Ahmed S., Thomas G., Ghoussaini M., Healey C.S., Humphreys M.K., Platte R., Morrison J., Maranian M., Pooley K.A., Luben R., Eccles D., Evans D.G., Fletcher O., Johnson N., dos S.S., I, Peto J., Stratton M.R., Rahman N., Jacobs K., Prentice R., Anderson G.L., Rajkovic A., Curb J.D., Ziegler R.G., Berg C.D., Buys S.S., McCarty C.A., Feigelson H.S., Calle E.E., Thun M.J., Diver W.R., Bojesen S., Nordestgaard B.G., Flyger H., Dork T., Schurmann P., Hillemanns P., Karstens J.H., Bogdanova N.V., Antonenkova N.N., Zalutsky I.V., Bermisheva M., Fedorova S., Khusnutdinova E., Kang D., Yoo K.Y., Noh D.Y., Ahn S.H., Devilee P., van Asperen C.J., Tollenaar R.A., Seynaeve C., Garcia-Closas M., Lissowska J., Brinton L., Peplonska B., Nevanlinna H., Heikkinen T., Aittomaki K., Blomqvist C., Hopper J.L., Southey M.C., Smith L., Spurdle A.B., Schmidt M.K., Broeks A., van Hien R.R., Cornelissen S., Milne R.L., Ribas G., Gonzalez-Neira A., Benitez J., Schmutzler R.K., Burwinkel B., Bartram C.R., Meindl A., Brauch H., Justenhoven C., Hamann U., Chang-Claude J., Hein R., Wang-Gohrke S., Lindblom A., Margolin S., Mannermaa A., Kosma V.M., Kataja V., Olson J.E., Wang X., Fredericksen Z., Giles G.G., Severi G., Baglietto L., English D.R., Hankinson S.E., Cox D.G., Kraft P., Vatten L.J., Hveem K., Kumle M., Sigurdson A., Doody M., Bhatti P., Alexander B.H., Hooning M.J., van den Ouweland A.M., Oldenburg R.A., Schutte M., Hall P., Czene K., Liu J., Li Y., Cox A., Elliott G., Brock I., Reed M.W., Shen C.Y., Yu J.C., Hsu G.C., Chen S.T., nton-Culver H., Ziogas A., Andrulis I.L., Knight J.A., Beesley J., Goode E.L., Couch F., Chenevix-Trench G., Hoover R.N., Ponder B.A., Hunter D.J., Pharoah P.D., Dunning A.M., Chanock S.J., e Easton D.F. (2009) Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat.Genet.*, **41**, 585-590.
- Ahn J., Gammon M.D., Santella R.M., Gaudet M.M., Britton J.A., Teitelbaum S.L., Terry M.B., Neugut A.I., Eng S.M., Zhang Y., Garza C., e Ambrosone C.B. (2006) Effects of glutathione S-transferase A1 (GSTA1) genotype and potential modifiers on breast cancer risk. *Carcinogenesis*, **27**, 1876-1882.
- Al-Tassan N., Chmiel N.H., Maynard J., Fleming N., Livingston A.L., Williams G.T., Hodges A.K., Davies D.R., David S.S., Sampson J.R., e Cheadle J.P. (2002) Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nat.Genet.*, **30**, 227-232.
- Ambrosone C.B., Moysich K.B., Furberg H., Freudenheim J.L., Bowman E.D., Ahmed S., Graham S., Vena J.E., e Shields P.G. (2003) CYP17 genetic polymorphism, breast cancer, and breast cancer risk factors. *Breast Cancer Res.*, **5**, R45-R51.
- Andonova I.E., Justenhoven C., Winter S., Hamann U., Baisch C., Rabstein S., Spickenheuer A., Harth V., Pesch B., Bruning T., Ko Y.D., Ganev V., e Brauch H. (2009) No evidence for glutathione S-transferases GSTA2, GSTM2, GSTO1, GSTO2, and GSTZ1 in breast cancer risk. *Breast Cancer Res.Treat.*,

- Angelini S., Kumar R., Carbone F., Maffei F., Forti G.C., Violante F.S., Lodi V., Curti S., Hemminki K., e Hrelia P. (2005) Micronuclei in humans induced by exposure to low level of ionizing radiation: influence of polymorphisms in DNA repair genes. *Mutat.Res.*, **570**, 105-117.
- Arnheim N., Strange C., e Erlich H. (1985) Use of pooled DNA samples to detect linkage disequilibrium of polymorphic restriction fragments and human disease: studies of the HLA class II loci. *Proc.Natl.Acad.Sci.U.S.A*, **82**, 6970-6974.
- Badve S. e Nakshatri H. (2009) Oestrogen-receptor-positive breast cancer: towards bridging histopathological and molecular classifications. *J.Clin.Pathol.*, **62**, 6-12.
- Bag A. e Bag N. (2008) Target sequence polymorphism of human manganese superoxide dismutase gene and its association with cancer risk: a review. *Cancer Epidemiol.Biomarkers Prev.*, **17**, 3298-3305.
- Bastos H.N., Antao M.R., Silva S.N., Azevedo A.P., Manita I., Teixeira V., Pina J.E., Gil O.M., Ferreira T.C., Limbert E., Rueff J., e Gaspar J.F. (2009) Association of polymorphisms in genes of the homologous recombination DNA repair pathway and thyroid cancer risk. *Thyroid*, **19**, 1067-1075.
- Bastos J., Barros H., e Lunet N. (2007) Breast cancer mortality trend in Portugal (1955-2002). *Acta Med.Port.*, **20**, 139-144.
- Benhamou S. e Sarasin A. (2005) ERCC2 /XPD gene polymorphisms and lung cancer: a HuGE review. *Am.J Epidemiol.*, **161**, 1-14.
- Bergman-Jungstrom M., Gentile M., Lundin A.C., e Wingren S. (1999) Association between CYP17 gene polymorphism and risk of breast cancer in young women. *Int.J.Cancer*, **84**, 350-353.
- Berkovich E., Monnat R.J., Jr., e Kastan M.B. (2007) Roles of ATM and NBS1 in chromatin structure modulation and DNA double-strand break repair. *Nat.Cell Biol.*, **9**, 683-690.
- Bertram J.S. (2000) The molecular biology of cancer. *Mol.Aspects Med.*, **21**, 167-223.
- Bosse Y., Bacot F., Montpetit A., Rung J., Qu H.Q., Engert J.C., Polychronakos C., Hudson T.J., Froguel P., Sladek R., e Desrosiers M. (2009) Identification of susceptibility genes for complex diseases using pooling-based genome-wide association scans. *Hum.Genet.*, **125**, 305-318.
- Boyle P. (2005) Breast cancer control: signs of progress, but more work required. *Breast*, **14**, 429-438.

- Bradbury A.R. e Olopade O.I. (2007) Genetic susceptibility to breast cancer. *Rev.Endocr.Metab Disord.*, **8**, 255-267.
- Bray F., McCarron P., e Parkin D.M. (2004) The changing global patterns of female breast cancer incidence and mortality. *Breast Cancer Res.*, **6**, 229-239.
- Breast Cancer Association Consortium (2006) Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. *J.Natl.Cancer Inst.*, **98**, 1382-1396.
- Bristow R.G. e Hill R.P. (2008) Hypoxia and metabolism. Hypoxia, DNA repair and genetic instability. *Nat.Rev.Cancer*, **8**, 180-192.
- Brooks J., Shore R.E., Zeleniuch-Jacquotte A., Currie D., Afanasyeva Y., Koenig K.L., Arslan A.A., Toniolo P., e Wirgin I. (2008) Polymorphisms in RAD51, XRCC2, and XRCC3 are not related to breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **17**, 1016-1019.
- Bu D., Tomlinson G., Lewis C.M., Zhang C., Kildebeck E., e Euhus D.M. (2006) An intronic polymorphism associated with increased XRCC1 expression, reduced apoptosis and familial breast cancer. *Breast Cancer Res.Treat.*, **99**, 257-265.
- Buermeyer A.B., Deschenes S.M., Baker S.M., e Liskay R.M. (1999) Mammalian DNA mismatch repair. *Annu.Rev.Genet.*, **33**, 533-564.
- Butkiewicz D., Rusin M., Enewold L., Shields P.G., Chorazy M., e Harris C.C. (2001) Genetic polymorphisms in DNA repair genes and risk of lung cancer. *Carcinogenesis*, **22**, 593-597.
- Caggana M., Kilgallen J., Conroy J.M., Wiencke J.K., Kelsey K.T., Miike R., Chen P., e Wrensch M.R. (2001) Associations between ERCC2 polymorphisms and gliomas. *Cancer Epidemiol.Biomarkers Prev.*, **10**, 355-360.
- Cai Q., Shu X.O., Wen W., Cheng J.R., Dai Q., Gao Y.T., e Zheng W. (2004) Genetic polymorphism in the manganese superoxide dismutase gene, antioxidant intake, and breast cancer risk: results from the Shanghai Breast Cancer Study. *Breast Cancer Res.*, **6**, R647-R655.
- Caldecott K.W. (2003) XRCC1 and DNA strand break repair. *DNA Repair (Amst)*, **2**, 955-969.
- Calladine C.R., Drew H.R., Luisi B.F., e Travers A.A. (2004) Understanding DNA - The Molecule and How It Works. **3rd ed.**, 1-349.

- Cariveau M.J., Tang X., Cui X.L., e Xu B. (2007) Characterization of an NBS1 C-terminal peptide that can inhibit ataxia telangiectasia mutated (ATM)-mediated DNA damage responses and enhance radiosensitivity. *Mol.Pharmacol.*, **72**, 320-326.
- Carmichael A.R., Harbach L., e Cooke R. (2009) Breast clinic and life style study (BLLISS). *Int.Semin.Surg.Oncol.*, **6**, 12.
- Casado-Diaz A., Cuenca-Acevedo R., Quesada J.M., e Dorado G. (2007) Individual single tube genotyping and DNA pooling by allele-specific PCR to uncover associations of polymorphisms with complex diseases. *Clin.Chim.Acta*, **376**, 155-162.
- Castro G.B., Fonte C.J.E., De-Lamonica-Freire E.M., e Hamada G.S. (2005) The absence of death by breast cancer among indigenous women living in Mato Grosso, Brazil 2000: Protection or undernotification. *J Womens Cancer*, **3**, 109-112.
- Cavalieri E., Frenkel K., Liehr J.G., Rogan E., e Roy D. (2000) Estrogens as endogenous genotoxic agents--DNA adducts and mutations. *J Natl.Cancer Inst.Monogr*, 75-93.
- Chacko P., Rajan B., Joseph T., Mathew B.S., e Pillai M.R. (2005) Polymorphisms in DNA repair gene XRCC1 and increased genetic susceptibility to breast cancer. *Breast Cancer Res.Treat.*, **89**, 15-21.
- Chang-Claude J., Popanda O., Tan X.L., Kropp S., Helmbold I., von F.D., Haase W., Sautter-Bihl M.L., Wenz F., Schmezer P., e Ambrosone C.B. (2005) Association between polymorphisms in the DNA repair genes, XRCC1, APE1, and XPD and acute side effects of radiotherapy in breast cancer patients. *Clin.Cancer Res.*, **11**, 4802-4809.
- Cheng T.C., Chen S.T., Huang C.S., Fu Y.P., Yu J.C., Cheng C.W., Wu P.E., e Shen C.Y. (2005) Breast cancer risk associated with genotype polymorphism of the catechol estrogen-metabolizing genes: a multigenic study on cancer susceptibility. *Int.J.Cancer*, **113**, 345-353.
- Choi J.Y., Hamajima N., Tajima K., Yoo K.Y., Yoon K.S., Park S.K., Kim S.U., Lee K.M., Noh D.Y., Ahn S.H., Choe K.J., Han W., Hirvonen A., e Kang D. (2003) hOGG1 Ser326Cys polymorphism and breast cancer risk among Asian women. *Breast Cancer Res.Treat.*, **79**, 59-62.
- Chowdari K.V., Northup A., Pless L., Wood J., Joo Y.H., Mirnics K., Lewis D.A., Levitt P.R., Bacanu S.A., e Nimgaonkar V.L. (2007) DNA pooling: a comprehensive, multi-stage association analysis of ACSL6 and SIRT5 polymorphisms in schizophrenia. *Genes Brain Behav.*, **6**, 229-239.

- Cianfrocca M. e Gradishar W. (2009) New molecular classifications of breast cancer. *CA Cancer J Clin.*, **59**, 303-313.
- Cleaver J.E., Lam E.T., e Revet I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat.Rev.Genet.*, **10**, 756-768.
- Coles B.F., Anderson K.E., Doerge D.R., Churchwell M.I., Lang N.P., e Kadlubar F.F. (2000) Quantitative analysis of interindividual variation of glutathione S-transferase expression in human pancreas and the ambiguity of correlating genotype with phenotype. *Cancer Res.*, **60**, 573-579.
- Coles B.F. e Kadlubar F.F. (2005) Human alpha class glutathione S-transferases: genetic polymorphism, expression, and susceptibility to disease. *Methods Enzymol.*, **401**, 9-42.
- Coles B.F., Morel F., Rauch C., Huber W.W., Yang M., Teitel C.H., Green B., Lang N.P., e Kadlubar F.F. (2001) Effect of polymorphism in the human glutathione S-transferase A1 promoter on hepatic GSTA1 and GSTA2 expression. *Pharmacogenetics*, **11**, 663-669.
- Collins J.A., Blake J.M., e Crosignani P.G. (2005) Breast cancer risk with postmenopausal hormonal treatment. *Hum.Reprod.Update.*, **11**, 545-560.
- Conde J., Silva S.N., Azevedo A.P., Teixeira V., Pina J.E., Rueff J., e Gaspar J.F. (2009) Association of common variants in mismatch repair genes and breast cancer susceptibility: a multigene study. *BMC.Cancer*, **9**, 344.
- Conner P., Lundstrom E., e von S.B. (2008) Breast cancer and hormonal therapy. *Clin.Obstet.Gynecol.*, **51**, 592-606.
- Costa S., Pinto D., Pereira D., Rodrigues H., Cameselle-Teijeiro J., Medeiros R., e Schmitt F. (2007) DNA repair polymorphisms might contribute differentially on familial and sporadic breast cancer susceptibility: a study on a Portuguese population. *Breast Cancer Res.Treat.*, **103**, 209-217.
- Couch F.J., Rasmussen L.J., Hofstra R., Monteiro A.N., Greenblatt M.S., e de W.N. (2008) Assessment of functional effects of unclassified genetic variants. *Hum.Mutat.*, **29**, 1314-1326.
- Crooke P.S., Ritchie M.D., Hachey D.L., Dawling S., Roodi N., e Parl F.F. (2006) Estrogens, enzyme variants, and breast cancer: a risk model. *Cancer Epidemiol.Biomarkers Prev.*, **15**, 1620-1629.
- Cuzick J. (2008) Assessing risk for breast cancer. *Breast Cancer Res.*, **10 (Suppl 4)**, S13.

- Danoy P., Sonoda E., Lathrop M., Takeda S., e Matsuda F. (2007) A naturally occurring genetic variant of human XRCC2 (R188H) confers increased resistance to cisplatin-induced DNA damage. *Biochem.Biophys.Res.Commun.*, **352**, 763-768.
- Deligezer U. e Dalay N. (2004) Association of the XRCC1 gene polymorphisms with cancer risk in Turkish breast cancer patients. *Exp.Mol.Med.*, **36**, 572-575.
- Dhillon V.S., Yeoh E., e Fenech M. (2009) DNA repair gene polymorphisms and prostate cancer risk in South Australia-results of a pilot study. *Urol.Oncol.*,
- Dialyna I.A., Arvanitis D.A., e Spandidos D.A. (2001) Genetic polymorphisms and transcriptional pattern analysis of CYP1A1, AhR, GSTM1, GSTP1 and GSTT1 genes in breast cancer. *Int.J.Mol.Med.*, **8**, 79-87.
- Dianov G.L., Sleeth K.M., Dianova I.I., e Allinson S.L. (2003) Repair of abasic sites in DNA. *Mutat.Res.*, **531**, 157-163.
- Difilippantonio S., Celeste A., Kruhlak M.J., Lee Y., Difilippantonio M.J., Feigenbaum L., Jackson S.P., McKinnon P.J., e Nussenzweig A. (2007) Distinct domains in Nbs1 regulate irradiation-induced checkpoints and apoptosis. *J.Exp.Med.*, **204**, 1003-1011.
- Domingues G.A. (2008) Polimorfismos em genes associados à apoptose e susceptibilidade genética para o cancro da mama. *Tese de Mestrado em Biologia Molecular e Genética*,
- Dudas A. e Chovanec M. (2004) DNA double-strand break repair by homologous recombination. *Mutat.Res.*, **566**, 131-167.
- Duell E.J., Millikan R.C., Pittman G.S., Winkel S., Lunn R.M., Tse C.K., Eaton A., Mohrenweiser H.W., Newman B., e Bell D.A. (2001) Polymorphisms in the DNA repair gene XRCC1 and breast cancer. *Cancer Epidemiol.Biomarkers Prev.*, **10**, 217-222.
- Duell E.J., Wiencke J.K., Cheng T.J., Varkonyi A., Zuo Z.F., Ashok T.D., Mark E.J., Wain J.C., Christiani D.C., e Kelsey K.T. (2000) Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells. *Carcinogenesis*, **21**, 965-971.
- Dufloth R.M., Costa S., Schmitt F., e Zeferino L.C. (2005) DNA repair gene polymorphisms and susceptibility to familial breast cancer in a group of patients from Campinas, Brazil. *Genet.Mol.Res.*, **4**, 771-782.
- Dumitrescu R.G. e Cotarla I. (2005) Understanding breast cancer risk -- where do we stand in 2005? *J.Cell Mol.Med.*, **9**, 208-221.

- Easton D.F. e Eeles R.A. (2008) Genome-wide association studies in cancer. *Hum.Mol.Genet.*, **17**, R109-R115.
- Egan K.M., Cai Q., Shu X.O., Jin F., Zhu T.L., Dai Q., Gao Y.T., e Zheng W. (2004) Genetic polymorphisms in GSTM1, GSTP1, and GSTT1 and the risk for breast cancer: results from the Shanghai Breast Cancer Study and meta-analysis. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 197-204.
- Egan K.M., Thompson P.A., Titus-Ernstoff L., Moore J.H., e Ambrosone C.B. (2003) MnSOD polymorphism and breast cancer in a population-based case-control study. *Cancer Lett.*, **199**, 27-33.
- Fackenthal J.D. e Olopade O.I. (2007) Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat.Rev.Cancer*, **7**, 937-948.
- Fan J. e Wilson III D.M. (2005) Protein-protein interactions and posttranslational modifications in mammalian base excision repair. *Free Radic.Biol.Med.*, **38**, 1121-1138.
- Feigelson H.S., Kean-Cowdin R., Coetzee G.A., Stram D.O., Kolonel L.N., e Henderson B.E. (2001) Building a multigenic model of breast cancer susceptibility: CYP17 and HSD17B1 are two important candidates. *Cancer Res.*, **61**, 785-789.
- Feuk L., Carson A.R., e Scherer S.W. (2006) Structural variation in the human genome. *Nat.Rev.Genet.*, **7**, 85-97.
- Figueiredo J.C., Knight J.A., Briollais L., Andrulis I.L., e Ozcelik H. (2004) Polymorphisms XRCC1-R399Q and XRCC3-T241M and the risk of breast cancer at the Ontario site of the Breast Cancer Family Registry. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 583-591.
- Ford B.N., Ruttan C.C., Kyle V.L., Brackley M.E., e Glickman B.W. (2000) Identification of single nucleotide polymorphisms in human DNA repair genes. *Carcinogenesis*, **21**, 1977-1981.
- Forsti A., Angelini S., Festa F., Sanyal S., Zhang Z., Grzybowska E., Pamula J., Pekala W., Zientek H., Hemminki K., e Kumar R. (2004) Single nucleotide polymorphisms in breast cancer. *Oncol.Rep.*, **11**, 917-922.
- Friedberg E.C. (2001) How nucleotide excision repair protects against cancer. *Nat.Rev.Cancer*, **1**, 22-33.
- Gaikwad N.W., Yang L., Muti P., Meza J.L., Pruthi S., Ingle J.N., Rogan E.G., e Cavalieri E.L. (2008) The molecular etiology of breast cancer: evidence from biomarkers of risk. *Int.J.Cancer*, **122**, 1949-1957.

- Gammon M.D. e Santella R.M. (2008) PAH, genetic susceptibility and breast cancer risk: an update from the Long Island Breast Cancer Study Project. *Eur.J Cancer*, **44**, 636-640.
- Garcia-Closas M., Egan K.M., Newcomb P.A., Brinton L.A., Titus-Ernstoff L., Chanock S., Welch R., Lissowska J., Peplonska B., Szeszenia-Dabrowska N., Zatonski W., Bardin-Mikolajczak A., e Struewing J.P. (2006) Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. *Hum.Genet.*, **119**, 376-388.
- Garinis G.A., van der Horst G.T., Vijg J., e Hoeijmakers J.H. (2008) DNA damage and ageing: new-age ideas for an age-old problem. *Nat.Cell Biol.*, **10**, 1241-1247.
- Garner E.I., Stokes E.E., Berkowitz R.S., Mok S.C., e Cramer D.W. (2002) Polymorphisms of the estrogen-metabolizing genes CYP17 and catechol-O-methyltransferase and risk of epithelial ovarian cancer. *Cancer Res.*, **62**, 3058-3062.
- Gaspar J., Faber A., e Rueff J. (2006) Variabilidade Genética e susceptibilidade individual para patologia multifactorial. *Revista Acção Médica*, **2**, 41-63.
- Gaustadnes M., Orntoft T.F., Jensen J.L., e Topping N. (2006) Validation of the use of DNA pools and primer extension in association studies of sporadic colorectal cancer for selection of candidate SNPs. *Hum.Mutat.*, **27**, 187-194.
- Geimba de Lima M., Koifman S., Scapulatempo I.L., Peixoto M., Naomi S., e Curado do A.M. (2001) [Risk factors for breast cancer among rural Terena Indian women in the State of Mato Grosso do Sul, Brazil]. *Cad.Saude Publica*, **17**, 1537-1544.
- Gemignani F., Landi S., Szeszenia-Dabrowska N., Zaridze D., Lissowska J., Rudnai P., Fabianova E., Mates D., Foretova L., Janout V., Bencko V., Gaborieau V., Gioia-Patricola L., Bellini I., Barale R., Canzian F., Hall J., Boffetta P., Hung R.J., e Brennan P. (2007) Development of lung cancer before the age of 50: the role of xenobiotic metabolizing genes. *Carcinogenesis*, **28**, 1287-1293.
- Gerson S.L. (2004) MGMT: its role in cancer aetiology and cancer therapeutics. *Nat.Rev.Cancer*, **4**, 296-307.
- Golubicic I., Borojevic N., e Pavlovic T. (2008) Risk factors for breast cancer: is ionizing radiation among them? *J.BUON.*, **13**, 487-494.
- Goode E.L., Ulrich C.M., e Potter J.D. (2002) Polymorphisms in DNA repair genes and associations with cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **11**, 1513-1530.

- Gorlewska-Roberts K., Green B., Fares M., Ambrosone C.B., e Kadlubar F.F. (2002) Carcinogen-DNA adducts in human breast epithelial cells. *Environ.Mol.Mutagen.*, **39**, 184-192.
- Guy C.A., Hoogendoorn B., Smith S.K., Coleman S., O'Donovan M.C., e Buckland P.R. (2004) Promoter polymorphisms in glutathione-S-transferase genes affect transcription. *Pharmacogenetics*, **14**, 45-51.
- Haiman C.A., Hankinson S.E., Spiegelman D., Colditz G.A., Willett W.C., Speizer F.E., Kelsey K.T., e Hunter D.J. (1999) The relationship between a polymorphism in CYP17 with plasma hormone levels and breast cancer. *Cancer Res.*, **59**, 1015-1020.
- Han J., Hankinson S.E., De V., I, Spiegelman D., Tamimi R.M., Mohrenweiser H.W., Colditz G.A., e Hunter D.J. (2003) A prospective study of XRCC1 haplotypes and their interaction with plasma carotenoids on breast cancer risk. *Cancer Res.*, **63**, 8536-8541.
- Han J., Hankinson S.E., Ranu H., De V., I, e Hunter D.J. (2004a) Polymorphisms in DNA double-strand break repair genes and breast cancer risk in the Nurses' Health Study. *Carcinogenesis*, **25**, 189-195.
- Han W., Kang D., Park I.A., Kim S.W., Bae J.Y., Chung K.W., e Noh D.Y. (2004b) Associations between breast cancer susceptibility gene polymorphisms and clinicopathological features. *Clin.Cancer Res.*, **10**, 124-130.
- Hanahan D. e Weinberg R.A. (2000) The hallmarks of cancer. *Cell*, **100**, 57-70.
- Harfe B.D. e Jinks-Robertson S. (2000) Mismatch repair proteins and mitotic genome stability. *Mutat.Res.*, **451**, 151-167.
- Hayes J.D., Flanagan J.U., e Jowsey I.R. (2005) Glutathione transferases. *Annu.Rev.Pharmacol.Toxicol.*, **45**, 51-88.
- Helleday T., Petermann E., Lundin C., Hodgson B., e Sharma R.A. (2008) DNA repair pathways as targets for cancer therapy. *Nat.Rev.Cancer*, **8**, 193-204.
- Hemminki K., Xu G., Angelini S., Snellman E., Jansen C.T., Lambert B., e Hou S.M. (2001) XPD exon 10 and 23 polymorphisms and DNA repair in human skin in situ. *Carcinogenesis*, **22**, 1185-1188.
- Hergueta-Redondo M., Palacios J., Cano A., e Moreno-Bueno G. (2008) "New" molecular taxonomy in breast cancer. *Clin.Transl.Oncol.*, **10**, 777-785.

- Hilakivi-Clarke L. (2000) Estrogens, BRCA1, and breast cancer. *Cancer Res.*, **60**, 4993-5001.
- Hoeijmakers J.H. (2001) Genome maintenance mechanisms for preventing cancer. *Nature*, **411**, 366-374.
- Hoogendoorn B., Norton N., Kirov G., Williams N., Hamshere M.L., Spurlock G., Austin J., Stephens M.K., Buckland P.R., Owen M.J., e O'Donovan M.C. (2000) Cheap, accurate and rapid allele frequency estimation of single nucleotide polymorphisms by primer extension and DHPLC in DNA pools. *Hum.Genet.*, **107**, 488-493.
- Hopper J.L. (2001) More breast cancer genes? *Breast Cancer Res.*, **3**, 154-157.
- Houtgraaf J.H., Versmissen J., e van der Giessen W.J. (2006) A concise review of DNA damage checkpoints and repair in mammalian cells. *Cardiovasc.Revasc.Med.*, **7**, 165-172.
- Hsu H.M., Wang H.C., Chen S.T., Hsu G.C., Shen C.Y., e Yu J.C. (2007) Breast cancer risk is associated with the genes encoding the DNA double-strand break repair Mre11/Rad50/Nbs1 complex. *Cancer Epidemiol.Biomarkers Prev.*, **16**, 2024-2032.
- Hu J.J., Smith T.R., Miller M.S., Lohman K., e Case L.D. (2002) Genetic regulation of ionizing radiation sensitivity and breast cancer risk. *Environ.Mol.Mutagen.*, **39**, 208-215.
- Hu Z., Ma H., Chen F., Wei Q., e Shen H. (2005) XRCC1 polymorphisms and cancer risk: a meta-analysis of 38 case-control studies. *Cancer Epidemiol.Biomarkers Prev.*, **14**, 1810-1818.
- Hu Z., Wei Q., Wang X., e Shen H. (2004) DNA repair gene XPD polymorphism and lung cancer risk: a meta-analysis. *Lung Cancer*, **46**, 1-10.
- Huang C.S., Chern H.D., Chang K.J., Cheng C.W., Hsu S.M., e Shen C.Y. (1999) Breast cancer risk associated with genotype polymorphism of the estrogen-metabolizing genes CYP17, CYP1A1, and COMT: a multigenic study on cancer susceptibility. *Cancer Res.*, **59**, 4870-4875.
- Huang Y., Li L., e Yu L. (2009) XRCC1 Arg399Gln, Arg194Trp and Arg280His polymorphisms in breast cancer risk: a meta-analysis. *Mutagenesis*, **24**, 331-339.

- Huber K.E., Carey L.A., e Wazer D.E. (2009) Breast cancer molecular subtypes in patients with locally advanced disease: impact on prognosis, patterns of recurrence, and response to therapy. *Semin.Radiat.Oncol.*, **19**, 204-210.
- Hung R.J., Hall J., Brennan P., e Boffetta P. (2005) Genetic polymorphisms in the base excision repair pathway and cancer risk: a HuGE review. *Am.J Epidemiol.*, **162**, 925-942.
- Irvin W.J., Jr. e Carey L.A. (2008) What is triple-negative breast cancer? *Eur.J Cancer*, **44**, 2799-2805.
- Isidro G., Laranjeira F., Pires A., Leite J., Regateiro F., Castro e Sousa, Soares J., Castro C., Gíria J., Brito M.J., Medeira A., Teixeira R., Morna H., Gaspar I., Marinho C., Jorge R., Brehm A., Ramos J.S., e Boavida M.G. (2004) Germline MUTYH (MYH) mutations in Portuguese individuals with multiple colorectal adenomas. *Hum.Mutat.*, **24**, 353-354.
- Jackson S.P. e Bartek J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071-1078.
- Jascur T. e Boland C.R. (2006) Structure and function of the components of the human DNA mismatch repair system. *Int.J.Cancer*, **119**, 2030-2035.
- Jawaid A. e Sham P. (2009) Impact and quantification of the sources of error in DNA pooling designs. *Ann.Hum.Genet.*, **73**, 118-124.
- Jiricny J. (2006) The multifaceted mismatch-repair system. *Nat.Rev.Mol.Cell Biol.*, **7**, 335-346.
- Jorgensen T.J., Helzlsouer K.J., Clipp S.C., Bolton J.H., Crum R.M., e Visvanathan K. (2009) DNA repair gene variants associated with benign breast disease in high cancer risk women. *Cancer Epidemiol.Biomarkers Prev.*, **18**, 346-350.
- Justenhoven C., Hamann U., Pesch B., Harth V., Rabstein S., Baisch C., Vollmert C., Illig T., Ko Y.D., Bruning T., e Brauch H. (2004) ERCC2 genotypes and a corresponding haplotype are linked with breast cancer risk in a German population. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 2059-2064.
- Keen J.C. e Davidson N.E. (2003) The biology of breast carcinoma. *Cancer*, **97**, 825-833.
- Kelley M.K., Engqvist-Goldstein A., Montali J.A., Wheatley J.B., Schmidt D.E., Jr., e Kauvar L.M. (1994) Variability of glutathione S-transferase isoenzyme patterns in matched normal and cancer human breast tissue. *Biochem.J.*, **304 (Pt 3)**, 843-848.

- Ketterer B. (1998) Glutathione S-transferases and prevention of cellular free radical damage. *Free Radic.Res.*, **28**, 647-658.
- Khanna K.K. e Jackson S.P. (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat.Genet.*, **27**, 247-254.
- Kim S.U., Park S.K., Yoo K.Y., Yoon K.S., Choi J.Y., Seo J.S., Park W.Y., Kim J.H., Noh D.Y., Ahn S.H., Choe K.J., Strickland P.T., Hirvonen A., e Kang D. (2002) XRCC1 genetic polymorphism and breast cancer risk. *Pharmacogenetics*, **12**, 335-338.
- Kiuru A., Lindholm C., Heinavaara S., Illus T., Jokinen P., Haapasalo H., Salminen T., Christensen H.C., Feychting M., Johansen C., Lonn S., Malmer B., Schoemaker M.J., Swerdlow A.J., e Auvinen A. (2008) XRCC1 and XRCC3 variants and risk of glioma and meningioma. *J.Neurooncol.*, **88**, 135-142.
- Kocabas N.A., Sardas S., Cholerton S., Daly A.K., Elhan A.H., e Karakaya A.E. (2005) Genetic polymorphism of manganese superoxide dismutase (MnSOD) and breast cancer susceptibility. *Cell Biochem.Funct.*, **23**, 73-76.
- Kocabas N.A., Sardas S., Cholerton S., Daly A.K., e Karakaya A.E. (2002) Cytochrome P450 CYP1B1 and catechol O-methyltransferase (COMT) genetic polymorphisms and breast cancer susceptibility in a Turkish population. *Arch.Toxicol.*, **76**, 643-649.
- Kolonel L.N., Altshuler D., e Henderson B.E. (2004) The multiethnic cohort study: exploring genes, lifestyle and cancer risk. *Nat.Rev.Cancer*, **4**, 519-527.
- Kruglyak L. e Nickerson D.A. (2001) Variation is the Spice of Life. *Nature Genetics*, **27**, 234-236.
- Krupa R., Synowiec E., Pawlowska E., Morawiec Z., Sobczuk A., Zadrozny M., Wozniak K., e Blasiak J. (2009) Polymorphism of the homologous recombination repair genes RAD51 and XRCC3 in breast cancer. *Exp.Mol.Pathol.*, **87**, 32-35.
- Kunkel T.A. e Erie D.A. (2005) DNA mismatch repair. *Annu.Rev.Biochem.*, **74**, 681-710.
- Kuschel B., Auranen A., McBride S., Novik K.L., Antoniou A., Lipscombe J.M., Day N.E., Easton D.F., Ponder B.A., Pharoah P.D., e Dunning A. (2002) Variants in DNA double-strand break repair genes and breast cancer susceptibility. *Hum.Mol.Genet.*, **11**, 1399-1407.
- Kuschel B., Chenevix-Trench G., Spurdle A.B., Chen X., Hopper J.L., Giles G.G., McCredie M., Chang-Claude J., Gregory C.S., Day N.E., Easton D.F., Ponder B.A., Dunning A.M., e Pharoah P.D. (2005) Common polymorphisms in ERCC2 (Xeroderma

- pigmentosum D) are not associated with breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **14**, 1828-1831.
- Lacroix M. e Leclercq G. (2005) The "portrait" of hereditary breast cancer. *Breast Cancer Res.Treat.*, **89**, 297-304.
- Landi S., Gemignani F., Moreno V., Gioia-Patricola L., Chabrier A., Guino E., Navarro M., de Oca J., Capella G., e Canzian F. (2005) A comprehensive analysis of phase I and phase II metabolism gene polymorphisms and risk of colorectal cancer. *Pharmacogenet.Genomics*, **15**, 535-546.
- Landi S., Gemignani F., Neri M., Barale R., Bonassi S., Bottari F., Canessa P.A., Canzian F., Ceppi M., Filiberti R., Ivaldi G.P., Mencoboni M., Scaruffi P., Tonini G.P., Mutti L., e Puntoni R. (2007) Polymorphisms of glutathione-S-transferase M1 and manganese superoxide dismutase are associated with the risk of malignant pleural mesothelioma. *Int.J.Cancer*, **120**, 2739-2743.
- Lange K., Cantor R., Horvath S., Perola M., Sabatti C., Sinsheimer J., e Sobel E. (2001) Mendel version 4.0: A complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. *Am.J.Hum.Genet.*, **69 (Supplement)**, 504.
- Lavigne J.A., Helzlsouer K.J., Huang H.Y., Strickland P.T., Bell D.A., Selmin O., Watson M.A., Hoffman S., Comstock G.W., e Yager J.D. (1997) An association between the allele coding for a low activity variant of catechol-O-methyltransferase and the risk for breast cancer. *Cancer Res.*, **57**, 5493-5497.
- Le Hellard S., Ballereau S.J., Visscher P.M., Torrance H.S., Pinson J., Morris S.W., Thomson M.L., Semple C.A., Muir W.J., Blackwood D.H., Porteous D.J., e Evans K.L. (2002) SNP genotyping on pooled DNAs: comparison of genotyping technologies and a semi automated method for data storage and analysis. *Nucleic Acids Res.*, **30**, e74.
- Lee J.E., Choi J.H., Lee J.H., e Lee M.G. (2005a) Gene SNPs and mutations in clinical genetic testing: haplotype-based testing and analysis. *Mutat.Res.*, **573**, 195-204.
- Lee K.M., Choi J.Y., Kang C., Kang C.P., Park S.K., Cho H., Cho D.Y., Yoo K.Y., Noh D.Y., Ahn S.H., Park C.G., Wei Q., e Kang D. (2005b) Genetic polymorphisms of selected DNA repair genes, estrogen and progesterone receptor status, and breast cancer risk. *Clin.Cancer Res.*, **11**, 4620-4626.
- Lee S.A., Fowke J.H., Lu W., Ye C., Zheng Y., Cai Q., Gu K., Gao Y.T., Shu X.O., e Zheng W. (2008) Cruciferous vegetables, the GSTP1 Ile105Val genetic polymorphism, and breast cancer risk. *Am.J.Clin.Nutr.*, **87**, 753-760.

- Lee S.A., Lee K.M., Park S.K., Choi J.Y., Kim B., Nam J., Yoo K.Y., Noh D.Y., Ahn S.H., e Kang D. (2007) Genetic polymorphism of XRCC3 Thr241Met and breast cancer risk: case-control study in Korean women and meta-analysis of 12 studies. *Breast Cancer Res. Treat.*, **103**, 71-76.
- Lee S.A., Lee K.M., Park W.Y., Kim B., Nam J., Yoo K.Y., Noh D.Y., Ahn S.H., Hirvonen A., e Kang D. (2005c) Obesity and genetic polymorphism of ERCC2 and ERCC4 as modifiers of risk of breast cancer. *Exp.Mol.Med.*, **37**, 86-90.
- Lehmann A.R. (2001) The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev.*, **15**, 15-23.
- Lei Y.C., Hwang S.J., Chang C.C., Kuo H.W., Luo J.C., Chang M.J., e Cheng T.J. (2002) Effects on sister chromatid exchange frequency of polymorphisms in DNA repair gene XRCC1 in smokers. *Mutat.Res.*, **519**, 93-101.
- Levy-Lahad E. e Plon S.E. (2003) Cancer. A risky business--assessing breast cancer risk. *Science*, **302**, 574-575.
- Li G.M. (2008) Mechanisms and functions of DNA mismatch repair. *Cell Res.*, **18**, 85-98.
- Li H., Ha T.C., e Tai B.C. (2009) XRCC1 gene polymorphisms and breast cancer risk in different populations: a meta-analysis. *Breast*, **18**, 183-191.
- Li H., Kantoff P.W., Giovannucci E., Leitzmann M.F., Gaziano J.M., Stampfer M.J., e Ma J. (2005) Manganese superoxide dismutase polymorphism, prediagnostic antioxidant status, and risk of clinical significant prostate cancer. *Cancer Res.*, **65**, 2498-2504.
- Liang G., Xing D., Miao X., Tan W., Yu C., Lu W., e Lin D. (2003) Sequence variations in the DNA repair gene XPD and risk of lung cancer in a Chinese population. *Int.J Cancer*, **105**, 669-673.
- Lieber M.R., Ma Y., Pannicke U., e Schwarz K. (2003) Mechanism and regulation of human non-homologous DNA end-joining. *Nat.Rev.Mol.Cell Biol.*, **4**, 712-720.
- Lindh A.R., Rafii S., Schultz N., Cox A., e Helleday T. (2006) Mitotic defects in XRCC3 variants T241M and D213N and their relation to cancer susceptibility. *Hum.Mol.Genet.*, **15**, 1217-1224.
- Lipkin S.M., Wang V., Jacoby R., Banerjee-Basu S., Baxevanis A.D., Lynch H.T., Elliott R.M., e Collins F.S. (2000) MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nat.Genet.*, **24**, 27-35.

- Lipkin S.M., Wang V., Stoler D.L., Anderson G.R., Kirsch I., Hadley D., Lynch H.T., e Collins F.S. (2001) Germline and somatic mutation analyses in the DNA mismatch repair gene MLH3: Evidence for somatic mutation in colorectal cancers. *Hum.Mutat.*, **17**, 389-396.
- Little J., Bradley L., Bray M.S., Clyne M., Dorman J., Ellsworth D.L., Hanson J., Khoury M., Lau J., O'Brien T.R., Rothman N., Stroup D., Taioli E., Thomas D., Vainio H., Wacholder S., e Weinberg C. (2002) Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. *Am.J.Epidemiol.*, **156**, 300-310.
- Loizidou M.A., Michael T., Neuhausen S.L., Newbold R.F., Marcou Y., Kakouri E., Daniel M., Papadopoulos P., Malas S., Kyriacou K., e Hadjisavvas A. (2008) Genetic polymorphisms in the DNA repair genes XRCC1, XRCC2 and XRCC3 and risk of breast cancer in Cyprus. *Breast Cancer Res.Treat.*, **112**, 575-579.
- Lord C.J., Garrett M.D., e Ashworth A. (2006) Targeting the double-strand DNA break repair pathway as a therapeutic strategy. *Clin.Cancer Res.*, **12**, 4463-4468.
- Lu J., Wei Q., Bondy M.L., Li D., Brewster A., Shete S., Yu T.K., Sahin A., Meric-Bernstam F., Hunt K.K., Singletary S.E., Ross M.I., e Wang L.E. (2006) Polymorphisms and haplotypes of the NBS1 gene are associated with risk of sporadic breast cancer in non-Hispanic white women ≤ 55 years. *Carcinogenesis*, **27**, 2209-2216.
- MacMahon B. (2006) Epidemiology and the causes of breast cancer. *Int.J.Cancer*, **118**, 2373-2378.
- Mangia A., Chiarappa P., Tommasi S., Chiriatti A., Petroni S., Schittulli F., e Paradiso A. (2008) Genetic heterogeneity by comparative genomic hybridization in BRCAx breast cancers. *Cancer Genet.Cytogenet.*, **182**, 75-83.
- Manuguerra M., Saletta F., Karagas M.R., Berwick M., Veglia F., Vineis P., e Matullo G. (2006) XRCC3 and XPD/ERCC2 single nucleotide polymorphisms and the risk of cancer: a HuGE review. *Am.J.Epidemiol.*, **164**, 297-302.
- Masi C.M. e Olopade O.I. (2005) Racial and ethnic disparities in breast cancer: a multilevel perspective. *Med.Clin.North Am.*, **89**, 753-770.
- Mattarucchi E., Marsoni M., Binelli G., Passi A., Lo C.F., Pasquali F., e Porta G. (2005) Different real time PCR approaches for the fine quantification of SNP's alleles in DNA pools: assays development, characterization and pre-validation. *J.Biochem.Mol.Biol.*, **38**, 555-562.
- Matullo G., Guarrera S., Carturan S., Peluso M., Malaveille C., Davico L., Piazza A., e Vineis P. (2001a) DNA repair gene polymorphisms, bulky DNA adducts in white

- blood cells and bladder cancer in a case-control study. *Int.J Cancer*, **92**, 562-567.
- Matullo G., Palli D., Peluso M., Guarrera S., Carturan S., Celentano E., Krogh V., Munnia A., Tumino R., Polidoro S., Piazza A., e Vineis P. (2001b) XRCC1, XRCC3, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects. *Carcinogenesis*, **22**, 1437-1445.
- Maugard C.M., Charrier J., e Bignon Y.J. (1998) Allelic deletion at glutathione S-transferase M1 locus and its association with breast cancer susceptibility. *Chem.Biol.Interact.*, **111-112**, 365-375.
- Maynard S., Schurman S.H., Harboe C., de Souza-Pinto N.C., e Bohr V.A. (2009) Base excision repair of oxidative DNA damage and association with cancer and aging. *Carcinogenesis*, **30**, 2-10.
- McCarty K.M., Santella R.M., Steck S.E., Cleveland R.J., Ahn J., Ambrosone C.B., North K., Sagiv S.K., Eng S.M., Teitelbaum S.L., Neugut A.I., e Gammon M.D. (2009) PAH-DNA adducts, cigarette smoking, GST polymorphisms, and breast cancer risk. *Environ.Health Perspect.*, **117**, 552-558.
- McIlwain C.C., Townsend D.M., e Tew K.D. (2006) Glutathione S-transferase polymorphisms: cancer incidence and therapy. *Oncogene*, **25**, 1639-1648.
- McKinnon P.J. e Caldecott K.W. (2007) DNA strand break repair and human genetic disease. *Annu.Rev.Genomics Hum.Genet.*, **8**, 37-55.
- McPherson K., Steel C.M., e Dixon J.M. (2000) ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *BMJ*, **321**, 624-628.
- Michiels S., Danoy P., Dessen P., Bera A., Boulet T., Bouchardy C., Lathrop M., Sarasin A., e Benhamou S. (2007) Polymorphism discovery in 62 DNA repair genes and haplotype associations with risks for lung and head and neck cancers. *Carcinogenesis*, **28**, 1731-1739.
- Millikan R.C., Player J., de Cotret A.R., Moorman P., Pittman G., Vannappagari V., Tse C.K., e Keku T. (2004) Manganese superoxide dismutase Ala-9Val polymorphism and risk of breast cancer in a population-based case-control study of African Americans and whites. *Breast Cancer Res.*, **6**, R264-R274.
- Millikan R.C., Player J.S., Decotret A.R., Tse C.K., e Keku T. (2005) Polymorphisms in DNA repair genes, medical exposure to ionizing radiation, and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **14**, 2326-2334.

- Mitrunen K. e Hirvonen A. (2003) Molecular epidemiology of sporadic breast cancer. The role of polymorphic genes involved in oestrogen biosynthesis and metabolism. *Mutat.Res.*, **544**, 9-41.
- Mitrunen K., Sillanpaa P., Kataja V., Eskelinen M., Kosma V.M., Benhamou S., Uusitupa M., e Hirvonen A. (2001) Association between manganese superoxide dismutase (MnSOD) gene polymorphism and breast cancer risk. *Carcinogenesis*, **22**, 827-829.
- Miyoshi Y., Iwao K., Ikeda N., Egawa C., e Noguchi S. (2000) Genetic polymorphism in CYP17 and breast cancer risk in Japanese women. *Eur.J Cancer*, **36**, 2375-2379.
- Miyoshi Y. e Noguchi S. (2003) Polymorphisms of estrogen synthesizing and metabolizing genes and breast cancer risk in Japanese women. *Biomed.Pharmacother.*, **57**, 471-481.
- Mohrenweiser H.W., Wilson D.M., III, e Jones I.M. (2003) Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive genetic variation in human DNA repair genes. *Mutat.Res.*, **526**, 93-125.
- Moinfar F., Beham A., Friedrich G., Deutsch A., Hrzenjak A., Luschin G., e Tavassoli F.A. (2008) Macro-environment of breast carcinoma: frequent genetic alterations in the normal appearing skins of patients with breast cancer. *Mod.Pathol.*, **21**, 639-646.
- Mor P. e Oberle K. (2008) Ethical issues related to BRCA gene testing in orthodox Jewish women. *Nurs.Ethics*, **15**, 512-522.
- Morel F., Rauch C., Coles B., Le F.E., e Guillouzo A. (2002) The human glutathione transferase alpha locus: genomic organization of the gene cluster and functional characterization of the genetic polymorphism in the hGSTA1 promoter. *Pharmacogenetics*, **12**, 277-286.
- Moullan N., Cox D.G., Angele S., Romestaing P., Gerard J.P., e Hall J. (2003) Polymorphisms in the DNA repair gene XRCC1, breast cancer risk, and response to radiotherapy. *Cancer Epidemiol.Biomarkers Prev.*, **12**, 1168-1174.
- Moyer A.M., Salavaggione O.E., Wu T.Y., Moon I., Eckloff B.W., Hildebrandt M.A., Schaid D.J., Wieben E.D., e Weinshilboum R.M. (2008) Glutathione s-transferase p1: gene sequence variation and functional genomic studies. *Cancer Res.*, **68**, 4791-4801.
- Mullenders L., Atkinson M., Paretzke H., Sabatier L., e Bouffler S. (2009) Assessing cancer risks of low-dose radiation. *Nat.Rev.Cancer*, **9**, 596-604.

- Murata H., Khatrar N.H., Kang Y., Gu L., e Li G.M. (2002) Genetic and epigenetic modification of mismatch repair genes hMSH2 and hMLH1 in sporadic breast cancer with microsatellite instability. *Oncogene*, **21**, 5696-5703.
- Nagaraju G. e Scully R. (2007) Minding the gap: the underground functions of BRCA1 and BRCA2 at stalled replication forks. *DNA Repair (Amst)*, **6**, 1018-1031.
- Neuhausen S.L. (1999) Ethnic differences in cancer risk resulting from genetic variation. *Cancer*, **86**, 2575-2582.
- Ning B., Wang C., Morel F., Nowell S., Ratnasinghe D.L., Carter W., Kadlubar F.F., e Coles B. (2004) Human glutathione S-transferase A2 polymorphisms: variant expression, distribution in prostate cancer cases/controls and a novel form. *Pharmacogenetics*, **14**, 35-44.
- Norton N., Williams N.M., O'Donovan M.C., e Owen M.J. (2004) DNA pooling as a tool for large-scale association studies in complex traits. *Ann.Med.*, **36**, 146-152.
- Norton N., Williams N.M., Williams H.J., Spurlock G., Kirov G., Morris D.W., Hoogendoorn B., Owen M.J., e O'Donovan M.C. (2002) Universal, robust, highly quantitative SNP allele frequency measurement in DNA pools. *Hum.Genet.*, **110**, 471-478.
- Oldenburg R.A., Meijers-Heijboer H., Cornelisse C.J., e Devilee P. (2007) Genetic susceptibility for breast cancer: how many more genes to be found? *Crit Rev.Oncol.Hematol.*, **63**, 125-149.
- Olopade O.I., Grushko T.A., Nanda R., e Huo D. (2008) Advances in breast cancer: pathways to personalized medicine. *Clin.Cancer Res.*, **14**, 7988-7999.
- Onay V.U., Briollais L., Knight J.A., Shi E., Wang Y., Wells S., Li H., Rajendram I., Andrulis I.L., e Ozcelik H. (2006) SNP-SNP interactions in breast cancer susceptibility. *BMC.Cancer*, **6**, 114.
- Packer B.R., Yeager M., Burdett L., Welch R., Beerman M., Qi L., Sicotte H., Staats B., Acharya M., Crenshaw A., Eckert A., Puri V., Gerhard D.S., e Chanock S.J. (2006) SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. *Nucleic Acids Res.*, **34**, D617-D621.
- Peters J., Loud J., Dimond E., e Jenkins J. (2001) Cancer genetics fundamentals. *Cancer Nurs.*, **24**, 446-461.
- Pharoah P.D., Dunning A.M., Ponder B.A., e Easton D.F. (2004) Association studies for finding cancer-susceptibility genetic variants. *Nat.Rev.Cancer*, **4**, 850-860.

- Pharoah P.D., Tyrer J., Dunning A.M., Easton D.F., e Ponder B.A. (2007) Association between common variation in 120 candidate genes and breast cancer risk. *PLoS.Genet.*, **3**, e42.
- Phillips D.H., Martin F.L., Williams J.A., Wheat L.M., Nolan L., Cole K.J., e Grover P.L. (2002) Mutagens in human breast lipid and milk: the search for environmental agents that initiate breast cancer. *Environ.Mol.Mutagen.*, **39**, 143-149.
- Ponder B.A. (2001) Cancer genetics. *Nature*, **411**, 336-341.
- Pooley K.A., Baynes C., Driver K.E., Tyrer J., Azzato E.M., Pharoah P.D., Easton D.F., Ponder B.A., e Dunning A.M. (2008) Common single-nucleotide polymorphisms in DNA double-strand break repair genes and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **17**, 3482-3489.
- Popanda O., Tan X.L., Ambrosone C.B., Kropp S., Helmbold I., von F.D., Haase W., Sautter-Bihl M.L., Wenz F., Schmezer P., e Chang-Claude J. (2006) Genetic polymorphisms in the DNA double-strand break repair genes XRCC3, XRCC2, and NBS1 are not associated with acute side effects of radiotherapy in breast cancer patients. *Cancer Epidemiol.Biomarkers Prev.*, **15**, 1048-1050.
- Poplawski T., Zadrozny M., Kolacinska A., Rykala J., Morawiec Z., e Blasiak J. (2005) Polymorphisms of the DNA mismatch repair gene HMSH2 in breast cancer occurrence and progression. *Breast Cancer Res.Treat.*, **94**, 199-204.
- Qiao Y., Spitz M.R., Shen H., Guo Z., Shete S., Hedayati M., Grossman L., Mohrenweiser H., e Wei Q. (2002) Modulation of repair of ultraviolet damage in the host-cell reactivation assay by polymorphic XPC and XPD/ERCC2 genotypes. *Carcinogenesis*, **23**, 295-299.
- Qu T., Morii E., Oboki K., Lu Y., e Morimoto K. (2005) Micronuclei in EM9 cells expressing polymorphic forms of human XRCC1. *Cancer Lett.*, **221**, 91-95.
- Rafii S., O'Regan P., Xinarianos G., Azmy I., Stephenson T., Reed M., Meuth M., Thacker J., e Cox A. (2002) A potential role for the XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer. *Hum.Mol.Genet.*, **11**, 1433-1438.
- Rea M.F. (2004) [Benefits of breastfeeding and women's health]. *J Pediatr.(Rio J)*, **80**, S142-S146.
- Reis-Filho J.S. e Tutt A.N. (2008) Triple negative tumours: a critical review. *Histopathology*, **52**, 108-118.

- Rodrigue A., Lafrance M., Gauthier M.C., McDonald D., Hendzel M., West S.C., Jasin M., e Masson J.Y. (2006) Interplay between human DNA repair proteins at a unique double-strand break in vivo. *EMBO J.*, **25**, 222-231.
- Rollinson S., Allan J.M., Law G.R., Roddam P.L., Smith M.T., Skibola C., Smith A.G., Forrest M.S., Sibley K., Higuchi R., Germer S., e Morgan G.J. (2004) High-throughput association testing on DNA pools to identify genetic variants that confer susceptibility to acute myeloid leukemia. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 795-800.
- Ronen A. e Glickman B.W. (2001) Human DNA repair genes. *Environ.Mol.Mutagen.*, **37**, 241-283.
- Rossner P., Jr., Terry M.B., Gammon M.D., Zhang F.F., Teitelbaum S.L., Eng S.M., Sagiv S.K., Gaudet M.M., Neugut A.I., e Santella R.M. (2006) OGG1 polymorphisms and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **15**, 811-815.
- Rueff J., Gaspar J., e Kranendonk M. (2002) DNA polymorphisms as modulators of genotoxicity and cancer. *Biol.Chem.*, **383**, 923-932.
- Rundle A. (2006) Carcinogen-DNA adducts as a biomarker for cancer risk. *Mutat.Res.*, **600**, 23-36.
- Russo J., Hasan L.M., Balogh G., Guo S., e Russo I.H. (2003) Estrogen and its metabolites are carcinogenic agents in human breast epithelial cells. *J.Steroid Biochem.Mol.Biol.*, **87**, 1-25.
- Russo J., Mailo D., Hu Y.F., Balogh G., Sheriff F., e Russo I.H. (2005a) Breast differentiation and its implication in cancer prevention. *Clin.Cancer Res.*, **11**, 931s-936s.
- Russo J., Moral R., Balogh G.A., Mailo D., e Russo I.H. (2005b) The protective role of pregnancy in breast cancer. *Breast Cancer Res.*, **7**, 131-142.
- Ruttan C.C. e Glickman B.W. (2002) Coding variants in human double-strand break DNA repair genes. *Mutat.Res.*, **509**, 175-200.
- Saadat M. e Ansari-Lari M. (2009) Polymorphism of XRCC1 (at codon 399) and susceptibility to breast cancer, a meta-analysis of the literatures. *Breast Cancer Res.Treat.*, **115**, 137-144.
- Sachidanandam R., Weissman D., Schmidt S.C., Kakol J.M., Stein L.D., Marth G., Sherry S., Mullikin J.C., Mortimore B.J., Willey D.L., Hunt S.E., Cole C.G., Coghill P.C., Rice C.M., Ning Z., Rogers J., Bentley D.R., Kwok P.Y., Mardis E.R., Yeh R.T., Schultz B., Cook L., Davenport R., Dante M., Fulton L., Hillier L., Waterston R.H.,

- McPherson J.D., Gilman B., Schaffner S., Van Etten W.J., Reich D., Higgins J., Daly M.J., Blumenstiel B., Baldwin J., Stange-Thomann N., Zody M.C., Linton L., Lander E.S., e Altshuler D. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, **409**, 928-933.
- Sakiyama T., Kohno T., Mimaki S., Ohta T., Yanagitani N., Sobue T., Kunitoh H., Saito R., Shimizu K., Hiramata C., Kimura J., Maeno G., Hirose H., Eguchi T., Saito D., Ohki M., e Yokota J. (2005) Association of amino acid substitution polymorphisms in DNA repair genes TP53, POLI, REV1 and LIG4 with lung cancer risk. *Int.J Cancer*, **114**, 730-737.
- Salzano F.M., Franco M.H., Weimer T.A., Callegari-Jacques S.M., Mestriner M.A., Hutz M.H., Flowers N.M., Santos R.V., e Coimbra C.E., Jr. (1997) The Brazilian Xavante Indians revisited: new protein genetic studies. *Am.J Phys.Anthropol.*, **104**, 23-34.
- Sancar A., Lindsey-Boltz L.A., Unsal-Kacmaz K., e Linn S. (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu.Rev.Biochem.*, **73**, 39-85.
- Sangrajrang S., Schmezer P., Burkholder I., Boffetta P., Brennan P., Woelfelschneider A., Bartsch H., Wiangnon S., Cheisilpa A., e Popanda O. (2007) The XRCC3 Thr241Met polymorphism and breast cancer risk: a case-control study in a Thai population. *Biomarkers*, **12**, 523-532.
- Sangrajrang S., Schmezer P., Burkholder I., Waas P., Boffetta P., Brennan P., Bartsch H., Wiangnon S., e Popanda O. (2008) Polymorphisms in three base excision repair genes and breast cancer risk in Thai women. *Breast Cancer Res.Treat.*, **111**, 279-288.
- Santen R., Cavalieri E., Rogan E., Russo J., Guttenplan J., Ingle J., e Yue W. (2009) Estrogen mediation of breast tumor formation involves estrogen receptor-dependent, as well as independent, genotoxic effects. *Ann.N.Y.Acad.Sci.*, **1155**, 132-140.
- Santucci-Darmanin S., Neyton S., Lespinasse F., Saunieres A., Gaudray P., e Paquis-Flucklinger V. (2002) The DNA mismatch-repair MLH3 protein interacts with MSH4 in meiotic cells, supporting a role for this MutL homolog in mammalian meiotic recombination. *Hum.Mol.Genet.*, **11**, 1697-1706.
- Savage S.A. (2008) Cancer genetic association studies in the genome-wide age. *Per.Med.*, **5**, 589-597.
- Saxena A., Dhillon V.S., Raish M., Asim M., Rehman S., Shukla N.K., Deo S.V., Ara A., e Husain S.A. (2008) Detection and relevance of germline genetic polymorphisms

- in glutathione S-transferases (GSTs) in breast cancer patients from northern Indian population. *Breast Cancer Res.Treat.*, **115**, 537-543.
- Sazci A., Ergul E., Utkan N.Z., Canturk N.Z., e Kaya G. (2004) Catechol-O-methyltransferase Val 108/158 Met polymorphism in premenopausal breast cancer patients. *Toxicology*, **204**, 197-202.
- Schabath M.B., Delclos G.L., Grossman H.B., Wang Y., Lerner S.P., Chamberlain R.M., Spitz M.R., e Wu X. (2005) Polymorphisms in XPD exons 10 and 23 and bladder cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **14**, 878-884.
- Schedin P. (2006) Pregnancy-associated breast cancer and metastasis. *Nat.Rev.Cancer*, **6**, 281-291.
- Schofield M.J. e Hsieh P. (2003) DNA mismatch repair: molecular mechanisms and biological function. *Annu.Rev.Microbiol.*, **57**, 579-608.
- Schroering A.G., Edelbrock M.A., Richards T.J., e Williams K.J. (2007) The cell cycle and DNA mismatch repair. *Exp.Cell Res.*, **313**, 292-304.
- Setiawan V.W., Hankinson S.E., Colditz G.A., Hunter D.J., e De V., I (2004) HSD17B1 gene polymorphisms and risk of endometrial and breast cancer. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 213-219.
- Sham P., Bader J.S., Craig I., O'donovan M., e Owen M. (2002) DNA Pooling: a tool for large-scale association studies. *Nat.Rev.Genet.*, **3**, 862-871.
- Shaw S.H., Carrasquillo M.M., Kashuk C., Puffenberger E.G., e Chakravarti A. (1998) Allele frequency distributions in pooled DNA samples: applications to mapping complex disease genes. *Genome Res.*, **8**, 111-123.
- Shen J., Ambrosone C.B., e Zhao H. (2009) Novel genetic variants in microRNA genes and familial breast cancer. *Int.J.Cancer*, **124**, 1178-1182.
- Shen J., Gammon M.D., Terry M.B., Wang L., Wang Q., Zhang F., Teitelbaum S.L., Eng S.M., Sagiv S.K., Gaudet M.M., Neugut A.I., e Santella R.M. (2005) Polymorphisms in XRCC1 modify the association between polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking, dietary antioxidants, and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **14**, 336-342.
- Shen M., Hung R.J., Brennan P., Malaveille C., Donato F., Placidi D., Carta A., Hautefeuille A., Boffetta P., e Porru S. (2003) Polymorphisms of the DNA repair genes XRCC1, XRCC3, XPD, interaction with environmental exposures, and bladder cancer risk in a case-control study in northern Italy. *Cancer Epidemiol.Biomarkers Prev.*, **12**, 1234-1240.

- Shi Q., Wang L.E., Bondy M.L., Brewster A., Singletary S.E., e Wei Q. (2004) Reduced DNA repair of benzo[a]pyrene diol epoxide-induced adducts and common XPD polymorphisms in breast cancer patients. *Carcinogenesis*, **25**, 1695-1700.
- Shimada M. e Nakanishi M. (2006) DNA damage checkpoints and cancer. *J.Mol.Histol.*, **37**, 253-260.
- Shin M.H., Lee K.M., Yang J.H., Nam S.J., Kim J.W., Yoo K.Y., Park S.K., Noh D.Y., Ahn S.H., Kim B., e Kang D. (2005) Genetic polymorphism of CYP17 and breast cancer risk in Korean women. *Exp.Mol.Med.*, **37**, 11-17.
- Shu X.O., Cai Q., Gao Y.T., Wen W., Jin F., e Zheng W. (2003) A population-based case-control study of the Arg399Gln polymorphism in DNA repair gene XRCC1 and risk of breast cancer. *Cancer Epidemiol.Biomarkers Prev.*, **12**, 1462-1467.
- Silva S.N., Azevedo A.P., Teixeira V., Pina J.E., Rueff J., e Gaspar J.F. (2009) The role of GSTA2 polymorphisms and haplotypes in breast cancer susceptibility: a case-control study in the Portuguese population. *Oncol.Rep.*, **22**, 593-598.
- Silva S.N., Bezerra de C.G., Faber A., Pires M., Oliveira V.C., Azevedo A.P., Cabral M.N., Manita I., Pina J.E., Rueff J., e Gaspar J. (2006a) The role of ERCC2 polymorphisms in breast cancer risk. *Cancer Genet.Cytogenet.*, **170**, 86-88.
- Silva S.N., Cabral M.N., Bezerra de C.G., Pires M., Azevedo A.P., Manita I., Pina J.E., Rueff J., e Gaspar J. (2006b) Breast cancer risk and polymorphisms in genes involved in metabolism of estrogens (CYP17, HSD17beta1, COMT and MnSOD): possible protective role of MnSOD gene polymorphism Val/Ala and Ala/Ala in women that never breast fed. *Oncol.Rep.*, **16**, 781-788.
- Silva S.N., Gil O.M., Oliveira V.C., Cabral M.N., Azevedo A.P., Faber A., Manita I., Ferreira T.C., Limbert E., Pina J.E., Rueff J., e Gaspar J. (2005) Association of polymorphisms in ERCC2 gene with non-familial thyroid cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **14**, 2407-2412.
- Silva S.N., Moita R., Azevedo A.P., Gouveia R., Manita I., Pina J.E., Rueff J., e Gaspar J. (2007) Menopausal age and XRCC1 gene polymorphisms: role in breast cancer risk. *Cancer Detect.Prev.*, **31**, 303-309.
- Silva S.N., Tomar M., Paulo C., Gomes B.C., Azevedo A.P., Teixeira V., Pina J.E., Rueff J., e Gaspar J.F. (2010) Breast cancer risk and common single nucleotide polymorphisms in homologous recombination DNA repair pathway genes XRCC2, XRCC3, NBS1 and RAD51. *Cancer Epidemiol.*, **34**, 85-92.

- Singh V., Parmar D., e Singh M.P. (2008) Do single nucleotide polymorphisms in xenobiotic metabolizing genes determine breast cancer susceptibility and treatment outcomes? *Cancer Invest*, **26**, 769-783.
- Singletary S.E. (2003) Rating the risk factors for breast cancer. *Ann.Surg.*, **237**, 474-482.
- Singletary S.E. e Connolly J.L. (2006) Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. *CA Cancer J.Clin.*, **56**, 37-47.
- Smedby K.E., Lindgren C.M., Hjalgrim H., Humphreys K., Schollkopf C., Chang E.T., Roos G., Ryder L.P., Falk K.I., Palmgren J., Kere J., Melbye M., Glimelius B., e Adami H.O. (2006) Variation in DNA repair genes ERCC2, XRCC1, and XRCC3 and risk of follicular lymphoma. *Cancer Epidemiol.Biomarkers Prev.*, **15**, 258-265.
- Smith T.R., Levine E.A., Freimanis R.I., Akman S.A., Allen G.O., Hoang K.N., Liu-Mares W., e Hu J.J. (2008) Polygenic Model of DNA-Repair Genetic Polymorphisms in Human Breast Cancer Risk. *Carcinogenesis*, **29**, 2132-2138.
- Smith T.R., Levine E.A., Perrier N.D., Miller M.S., Freimanis R.I., Lohman K., Case L.D., Xu J., Mohrenweiser H.W., e Hu J.J. (2003a) DNA-repair genetic polymorphisms and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **12**, 1200-1204.
- Smith T.R., Miller M.S., Lohman K., Lange E.M., Case L.D., Mohrenweiser H.W., e Hu J.J. (2003b) Polymorphisms of XRCC1 and XRCC3 genes and susceptibility to breast cancer. *Cancer Lett.*, **190**, 183-190.
- Sole X., Guino E., Valls J., Iriarte R., e Moreno V. (2006) SNPStats: a web tool for the analysis of association studies. *Bioinformatics.*, **22**, 1928-1929.
- Spurdle A.B., Chang J.H., Byrnes G.B., Chen X., Dite G.S., McCredie M.R., Giles G.G., Southey M.C., Chenevix-Trench G., e Hopper J.L. (2007) A systematic approach to analysing gene-gene interactions: polymorphisms at the microsomal epoxide hydrolase EPHX and glutathione S-transferase GSTM1, GSTT1, and GSTP1 loci and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **16**, 769-774.
- Steck S.E., Gaudet M.M., Britton J.A., Teitelbaum S.L., Terry M.B., Neugut A.I., Santella R.M., e Gammon M.D. (2007) Interactions among GSTM1, GSTT1 and GSTP1 polymorphisms, cruciferous vegetable intake and breast cancer risk. *Carcinogenesis*, **28**, 1954-1959.
- Stern M.C., Johnson L.R., Bell D.A., e Taylor J.A. (2002) XPD codon 751 polymorphism, metabolism genes, smoking, and bladder cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **11**, 1004-1011.

- Strange R.C., Jones P.W., e Fryer A.A. (2000) Glutathione S-transferase: genetics and role in toxicology. *Toxicol.Lett.*, **112-113**, 357-363.
- Sturgis E.M., Castillo E.J., Li L., Eicher S.A., Strom S.S., Spitz M.R., e Wei Q. (2002a) XPD/ERCC2 EXON 8 Polymorphisms: rarity and lack of significance in risk of squamous cell carcinoma of the head and neck. *Oral Oncol.*, **38**, 475-477.
- Sturgis E.M., Dahlstrom K.R., Spitz M.R., e Wei Q. (2002b) DNA repair gene ERCC1 and ERCC2/XPD polymorphisms and risk of squamous cell carcinoma of the head and neck. *Arch.Otolaryngol.Head Neck Surg.*, **128**, 1084-1088.
- Sturgis E.M., Zheng R., Li L., Castillo E.J., Eicher S.A., Chen M., Strom S.S., Spitz M.R., e Wei Q. (2000) XPD/ERCC2 polymorphisms and risk of head and neck cancer: a case-control analysis. *Carcinogenesis*, **21**, 2219-2223.
- Sutton A., Imbert A., Igoudjil A., Descatoire V., Cazanave S., Pessayre D., e Degoul F. (2005) The manganese superoxide dismutase Ala16Val dimorphism modulates both mitochondrial import and mRNA stability. *Pharmacogenet.Genomics*, **15**, 311-319.
- Syamala V.S., Sreeja L., Syamala V., Raveendran P.B., Balakrishnan R., Kuttan R., e Ankathil R. (2008) Influence of germline polymorphisms of GSTT1, GSTM1, and GSTP1 in familial versus sporadic breast cancer susceptibility and survival. *Fam.Cancer*, **7**, 213-220.
- Synowiec E., Stefanska J., Morawiec Z., Blasiak J., e Wozniak K. (2008) Association between DNA damage, DNA repair genes variability and clinical characteristics in breast cancer patients. *Mutat.Res.*, **648**, 65-72.
- Tabor H.K., Risch N.J., e Myers R.M. (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat.Rev.Genet.*, **3**, 391-397.
- Tang D., Cho S., Rundle A., Chen S., Phillips D., Zhou J., Hsu Y., Schnabel F., Estabrook A., e Perera F.P. (2002) Polymorphisms in the DNA repair enzyme XPD are associated with increased levels of PAH-DNA adducts in a case-control study of breast cancer. *Breast Cancer Res.Treat.*, **75**, 159-166.
- Tchatchou S., Jung A., Hemminki K., Sutter C., Wappenschmidt B., Bugert P., Weber B.H., Niederacher D., Arnold N., Varon-Mateeva R., Ditsch N., Meindl A., Schmutzler R.K., Bartram C.R., e Burwinkel B. (2009) A variant affecting a putative miRNA target site in estrogen receptor (ESR) 1 is associated with breast cancer risk in premenopausal women. *Carcinogenesis*, **30**, 59-64.

- Tempfer C.B., Hefler L.A., Schneeberger C., e Huber J.C. (2006) How valid is single nucleotide polymorphism (SNP) diagnosis for the individual risk assessment of breast cancer? *Gynecol.Endocrinol.*, **22**, 155-159.
- Terry M.B., Gammon M.D., Zhang F.F., Eng S.M., Sagiv S.K., Paykin A.B., Wang Q., Hayes S., Teitelbaum S.L., Neugut A.I., e Santella R.M. (2004) Polymorphism in the DNA repair gene XPD, polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking, and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 2053-2058.
- Tetlow N. e Board P.G. (2004) Functional polymorphism of human glutathione transferase A2. *Pharmacogenetics*, **14**, 111-116.
- Tetlow N., Coggan M., Casarotto M.G., e Board P.G. (2004) Functional polymorphism of human glutathione transferase A3: effects on xenobiotic metabolism and steroid biosynthesis. *Pharmacogenetics*, **14**, 657-663.
- Tetlow N., Liu D., e Board P. (2001) Polymorphism of human Alpha class glutathione transferases. *Pharmacogenetics*, **11**, 609-617.
- The International HapMap Consortium (2003) The International HapMap Project. *Nature*, **426**, 789-796.
- Thomas G., Jacobs K.B., Kraft P., Yeager M., Wacholder S., Cox D.G., Hankinson S.E., Hutchinson A., Wang Z., Yu K., Chatterjee N., Garcia-Closas M., Gonzalez-Bosquet J., Prokunina-Olsson L., Orr N., Willett W.C., Colditz G.A., Ziegler R.G., Berg C.D., Buys S.S., McCarty C.A., Feigelson H.S., Calle E.E., Thun M.J., Diver R., Prentice R., Jackson R., Kooperberg C., Chlebowski R., Lissowska J., Peplonska B., Brinton L.A., Sigurdson A., Doody M., Bhatti P., Alexander B.H., Buring J., Lee I.M., Vatten L.J., Hveem K., Kumle M., Hayes R.B., Tucker M., Gerhard D.S., Fraumeni J.F., Jr., Hoover R.N., Chanock S.J., e Hunter D.J. (2009) A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat.Genet.*, **41**, 579-584.
- Thompson L.H. e Schild D. (2001) Homologous recombinational repair of DNA ensures mammalian chromosome stability. *Mutat.Res.*, **477**, 131-153.
- Thompson P.A., DeMarini D.M., Kadlubar F.F., McClure G.Y., Brooks L.R., Green B.L., Fares M.Y., Stone A., Joseph P.D., e Ambrosone C.B. (2002) Evidence for the presence of mutagenic arylamines in human breast milk and DNA adducts in exfoliated breast ductal epithelial cells. *Environ.Mol.Mutagen.*, **39**, 134-142.
- Torresan C., Oliveira M.M., Torrezan G.T., de Oliveira S.F., Abuazar C.S., Losi-Guembarovski R., Lima R.S., Urban C.A., Cavalli I.J., e Ribeiro E.M. (2008) Genetic polymorphisms in oestrogen metabolic pathway and breast cancer: a

- positive association with combined CYP/GST genotypes. *Clin.Exp.Med.*, **8**, 65-71.
- Tranah G.J., Giovannucci E., Ma J., Fuchs C., Hankinson S.E., e Hunter D.J. (2004) XRCC2 and XRCC3 polymorphisms are not associated with risk of colorectal adenoma. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 1090-1091.
- Tworoger S.S., Chubak J., Aiello E.J., Ulrich C.M., Atkinson C., Potter J.D., Yasui Y., Stapleton P.L., Lampe J.W., Farin F.M., Stanczyk F.Z., e McTiernan A. (2004) Association of CYP17, CYP19, CYP1B1, and COMT polymorphisms with serum and urinary sex hormone concentrations in postmenopausal women. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 94-101.
- Unlu A., Ates N.A., Tamer L., e Ates C. (2008) Relation of glutathione S-transferase T1, M1 and P1 genotypes and breast cancer risk. *Cell Biochem.Funct.*, **26**, 643-647.
- van Brabant A.J., Stan R., e Ellis N.A. (2000) DNA helicases, genomic instability, and human genetic disease. *Annu.Rev.Genomics Hum.Genet.*, **1**, 409-459.
- Van Emburgh B.O., Hu J.J., Levine E.A., Mosley L.J., Perrier N.D., Freimanis R.I., Allen G.O., Rubin P., Sherrill G.B., Shaw C.S., Carey L.A., Sawyer L.R., e Miller M.S. (2008) Polymorphisms in CYP1B1, GSTM1, GSTT1 and GSTP1, and susceptibility to breast cancer. *Oncol.Rep.*, **19**, 1311-1321.
- Vargo-Gogola T. e Rosen J.M. (2007) Modelling breast cancer: one size does not fit all. *Nat.Rev.Cancer*, **7**, 659-672.
- Venkitaraman A.R. (2002) Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*, **108**, 171-182.
- Vogel U., Nexø B.A., Olsen A., Thomsen B., Jacobsen N.R., Wallin H., Overvad K., e Tjønneland A. (2003) No association between OGG1 Ser326Cys polymorphism and breast cancer risk. *Cancer Epidemiol.Biomarkers Prev.*, **12**, 170-171.
- Vogel V.G. (2008) Epidemiology, genetics, and risk evaluation of postmenopausal women at risk of breast cancer. *Menopause.*, **15**, 782-789.
- Vogl F.D., Taioli E., Maugard C., Zheng W., Pinto L.F., Ambrosone C., Parl F.F., Nedelcheva-Kristensen V., Rebbeck T.R., Brennan P., e Boffetta P. (2004) Glutathione S-transferases M1, T1, and P1 and breast cancer: a pooled analysis. *Cancer Epidemiol.Biomarkers Prev.*, **13**, 1473-1479.
- Wajapeyee N. e Somasundaram K. (2004) Pharmacogenomics in breast cancer: current trends and future directions. *Curr.Opin.Mol.Ther.*, **6**, 296-301.

- Wang S., Wang F., Shi X., Dai J., Peng Y., Guo X., Wang X., Shen H., e Hu Z. (2009) Association between manganese superoxide dismutase (MnSOD) Val-9Ala polymorphism and cancer risk - A meta-analysis. *Eur.J Cancer*, **45**, 2874-2881.
- Wang Y., Spitz M.R., Zhu Y., Dong Q., Shete S., e Wu X. (2003) From genotype to phenotype: correlating XRCC1 polymorphisms with mutagen sensitivity. *DNA Repair (Amst)*, **2**, 901-908.
- Webb P.M., Hopper J.L., Newman B., Chen X., Kelemen L., Giles G.G., Southey M.C., Chenevix-Trench G., e Spurdle A.B. (2005) Double-strand break repair gene polymorphisms and risk of breast or ovarian cancer. *Cancer Epidemiol.Biomarkers Prev.*, **14**, 319-323.
- Wedren S., Rudqvist T.R., Granath F., Weiderpass E., Ingelman-Sundberg M., Persson I., e Magnusson C. (2003) Catechol-O-methyltransferase gene polymorphism and post-menopausal breast cancer risk. *Carcinogenesis*, **24**, 681-687.
- Weyandt J., Ellsworth R.E., Hooke J.A., Shriver C.D., e Ellsworth D.L. (2008) Environmental chemicals and breast cancer risk--a structural chemistry perspective. *Curr.Med.Chem.*, **15**, 2680-2701.
- Wilson III D.M. e Bohr V.A. (2007) The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair (Amst)*, **6**, 544-559.
- Wong E.M., Tesoriero A.A., Pupo G.M., McCredie M.R., Giles G.G., Hopper J.L., Mann G.J., Goldgar D.E., e Southey M.C. (2008) Is MSH2 a breast cancer susceptibility gene? *Fam.Cancer*, **7**, 151-155.
- Wood R.D., Mitchell M., e Lindahl T. (2005) Human DNA repair genes, 2005. *Mutat.Res.*, **577**, 275-283.
- Wu A.H., Seow A., Arakawa K., Van Den B.D., Lee H.P., e Yu M.C. (2003) HSD17B1 and CYP17 polymorphisms and breast cancer risk among Chinese women in Singapore. *Int.J.Cancer*, **104**, 450-457.
- Xu J., Zheng S.L., Turner A., Isaacs S.D., Wiley K.E., Hawkins G.A., Chang B.L., Bleecker E.R., Walsh P.C., Meyers D.A., e Isaacs W.B. (2002) Associations between hOGG1 sequence variants and prostate cancer susceptibility. *Cancer Res.*, **62**, 2253-2257.
- Yager J.D. (2000) Endogenous estrogens as carcinogens through metabolic activation. *J.Natl.Cancer Inst.Monogr*, 67-73.
- Yang H.C., Pan C.C., Lin C.Y., e Fann C.S. (2006) PDA: Pooled DNA analyzer. *BMC.Bioinformatics.*, **7**, 233.

- Yim D.S., Parkb S.K., Yoo K.Y., Yoon K.S., Chung H.H., Kang H.L., Ahn S.H., Noh D.Y., Choe K.J., Jang I.J., Shin S.G., Strickland P.T., Hirvonen A., e Kang D. (2001) Relationship between the Val158Met polymorphism of catechol O-methyl transferase and breast cancer. *Pharmacogenetics*, **11**, 279-286.
- Zhang L., Zhang Z., e Yan W. (2005) Single nucleotide polymorphisms for DNA repair genes in breast cancer patients. *Clin.Chim.Acta*, **359**, 150-155.
- Zhang Q.M. e Dianov G.L. (2005) DNA repair fidelity of base excision repair pathways in human cell extracts. *DNA Repair (Amst)*, **4**, 263-270.
- Zhang Y., Newcomb P.A., Egan K.M., Titus-Ernstoff L., Chanock S., Welch R., Brinton L.A., Lissowska J., Bardin-Mikolajczak A., Peplonska B., Szeszenia-Dabrowska N., Zatonski W., e Garcia-Closas M. (2006) Genetic polymorphisms in base-excision repair pathway genes and risk of breast cancer. *Cancer Epidemiol.Biomarkers Prev.*, **15**, 353-358.
- Zhao H., Pfeiffer R., e Gail M.H. (2003) Haplotype analysis in population genetics and association studies. *Pharmacogenomics.*, **4**, 171-178.
- Zheng T., Holford T.R., Zahm S.H., Owens P.H., Boyle P., Zhang Y., Zhang B., Wise J.P., Sr., Stephenson L.P., e li-Osman F. (2003) Glutathione S-transferase M1 and T1 genetic polymorphisms, alcohol consumption and breast cancer risk. *Br.J.Cancer*, **88**, 58-62.
- Zheng W., Wen W.Q., Gustafson D.R., Gross M., Cerhan J.R., e Folsom A.R. (2002) GSTM1 and GSTT1 polymorphisms and postmenopausal breast cancer risk. *Breast Cancer Res.Treat.*, **74**, 9-16.
- Zuo Y., Zou G., e Zhao H. (2006) Two-stage designs in case-control association analysis. *Genetics*, **173**, 1747-1760.