



UNIVERSIDADE DA BEIRA INTERIOR
Ciências da Saúde

Expressão de genes de baixa penetrância no prognóstico do cancro da mama

João Paulo Baptista Godinho

Dissertação para obtenção do Grau de Mestre em
Medicina
(ciclo de estudos integrado)

Orientador: Prof. Doutor José Alberto Fonseca Moutinho
Coorientadora: Prof^a. Doutora Luiza Breitenfeld Granadeiro
Coorientadora: Doutora Ana Cristina Ramalinho

Covilhã, abril de 2014

Agradecimentos

Qualquer trabalho reflete a dedicação e empenho que o seu autor lhe proporciona, sobretudo quando nos embrenhamos num tema fora da “área de conforto”.

Aos meus pais, que me continuam a proporcionar tudo o que está ao seu alcance para percorrer o meu caminho. Ao meu irmão que, mesmo longe, me apoia e motiva para trabalhar sempre e jamais baixar os braços. À Rita meras palavras de agradecimento não são suficientes, pela forma incansável com que me manteve coerente e disposto a superar qualquer adversidade que se interpôs. Sem vós os quatro seria muito mais difícil levantar-me de todos os tropeços que as dificuldades precipitaram.

A realização deste projeto não seria possível sem o apoio e toda a disponibilidade do orientador, Professor Doutor José Moutinho e coorientadoras Professora Doutora Luiza Granadeiro e Doutora Ana Cristina Ramalinho. A eles um grande obrigado por toda a paciência e dedicação, que me permitiram concluir este gratificante trabalho.

Ao Luís Crisóstomo, um obrigado pelo constante apoio nas mais difíceis técnicas laboratoriais que tive oportunidade de utilizar ao longo deste projeto.

Ao grupo de Genética/Endocrinologia, meus colegas de laboratório, um obrigado por me “socorrerem”, sempre que necessário, na azáfama do quotidiano laboratorial. Em particular à Maria Inês Alvéolos, sem ela teria sido impossível, pois mais que uma colega de bancada, foi a minha tutora, aconselhando-me desde o primeiro momento que cheguei ao Laboratório.

Não poderia deixar de agradecer à Professora Sara Nunes, que me ajudou a olhar para a análise estatística de uma forma mais fácil e simples.

Um agradecimento ao Serviço de Anatomia Patológica do CHCB, e muito especialmente à Doutora Rosa Tomé, técnica do Serviço de Anatomia Patológica do CHCB e à Doutora Helena Garcia do CEDAP (Centro de Diagnóstico Anátomo-Patológico) de Coimbra, que flexibilizaram e selecionaram as lâminas de tecido em parafina. Sem a participação de ambas não haveria amostras para realizar este trabalho.

Um sincero Obrigado a todos!

Resumo

O cancro da mama é uma patologia de grande impacto na população feminina, devido à grande incidência, prevalência e morbimortalidade a que está associada. O património genético representa um papel importante nesta entidade nosológica. No entanto, mutações que conferem alto risco para esta doença, apenas explicam uma diminuta porção dos casos. Genes de baixa penetrância evidenciam um importante papel na suscetibilidade e têm sido alvo de estudo nas diversas populações. Trabalhos desenvolvidos pelo nosso grupo, na população que o Centro Hospitalar Cova da Beira abrange, apontaram para uma associação entre polimorfismos de genes de baixa penetrância e cancro da mama. São eles “*CYP19A1* Trp39Arg (T/C)”, “*GSTT1 null*” e “*GSTM1 null*”.

Realizámos um trabalho do tipo caso-controlo, em que abordámos uma amostra de 37 mulheres pertencentes à população do Centro Hospitalar Cova da Beira. Foi estudado se os 3 polimorfismos (“*CYP19A1* Trp39Arg (T/C)”, “*GSTT1 null*” e “*GSTM1 null*”) apresentam diferentes níveis de expressão, entre mulheres com diagnóstico de cancro da mama, mas com prognóstico distinto. Focámo-nos assim em dois principais parâmetros, “Estadio clínico” e “Envolvimento ganglionar regional”. Neste trabalho de investigação foram utilizadas amostras de tecido de mama fixadas e incluídas em parafina. A técnica usada para quantificar a expressão génica foi PCR em tempo real. Para a análise estatística recorreu-se ao teste de qui-quadrado de *Pearson* e a testes não paramétricos. Para calcular as diferenças na expressão foi utilizado o método comparativo de C_T , $2^{-\Delta\Delta CT}$.

Os resultados obtidos, para os polimorfismos de *GSTT1* e *GSTM1*, foram influenciados pela diminuta quantidade de RNA extraída do tecido em parafina. Contudo, apesar da difícil extração de RNA, obteve-se uma elevada eficácia nas metodologias laboratoriais que dependeram deste procedimento, cerca de 81,1%.

No que concerne ao polimorfismo *CYP19A1* Trp39Arg, detetou-se expressão em cerca de 36,7% dos casos estudados. Após validarmos os resultados obtidos verificámos que a expressão entre 2 grupos de diferente estadio clínico é distinta. O grupo com “Estadio III” expressa 4,33 vezes mais *CYP19A1* Trp39Arg que o grupo no “Estadio I e II”, sendo que a presença ou ausência dos Recetores de estrogénio tem que ser considerada para a correta interpretação dos resultados. Especulamos que a expressão deste polimorfismo da aromatase poderá ter algum tipo de impacto na determinação do prognóstico do cancro da mama, nesta população estudada.

Palavras-chave

Cancro da mama; Prognóstico; Expressão; *CYP19A1* Trp39Arg (T/C); Glutathione-S-transferase

Abstract

Breast cancer is a pathology with great impact in female population, because it is associated with considerable incidence, prevalence, morbidity and mortality. Genetic heritage represents an important role in this nosological entity. However, mutations that confer high risk of developing this disease only explain a small proportion of the cases. Low penetrance genes are an important susceptibility factor and are being the target of several studies in populations. Research accomplished by our group within the population that Centro Hospitalar Cova da Beira covers, pointed to an association between low penetrance genes polymorphisms and the risk of breast cancer. These polymorphisms are “*CYP19A1* Trp39Arg (T/C)”, “*GSTT1 null*” and “*GSTM1 null*”.

We conducted a case-control study in a sample of 37 women belonging to the population covered by Centro Hospitalar Cova da Beira. Our main goal was to determine if the 3 polymorphisms (“*CYP19A1* Trp39Arg (T/C)”, “*GSTT1 null*” and “*GSTM1 null*”) have different expression levels, between women diagnosed with breast cancer, but with distinct prognoses. Our focus was based in two main parameters, “Clinical staging” and “Regional lymph node involvement”. In this research work we used paraffin embedded tissues of breast. Expression quantification was accomplished by Real Time PCR. Statistical analysis was performed recurring to the Pearson’s chi-squared test and nonparametric tests. To estimate expression levels differences we used the C_T comparative method, $2^{-\Delta\Delta CT}$.

GSTT1 and *GSTM1* results were influenced by small RNA quantities extracted from the paraffin embedded tissues. Nevertheless, despite difficult RNA extraction, we obtained a high efficacy in the laboratory methodologies depending on this procedure, about 81.1%.

Regarding *CYP19A1* Trp39Arg expression, 36.7% of the studied cases were detected. After result validation we noticed that this expression was different between two groups of distinct clinical staging. “Stage III” expresses 4.33 more *CYP19A1* Trp39Arg than “Stages I and II”, being aware that the presence or absence of the Estrogen receptor has to be considered to interpret the results correctly. We speculate that this aromatase polymorphism expression may have some kind of impact in the determination of breast cancer prognoses in the studied population.

Keywords

Breast cancer; Prognoses; Expression; *CYP19A1* Trp39Arg (T/C); Glutathione S-transferase

Índice

Agradecimentos	iii
Resumo	v
Abstract.....	vii
Índice.	ix
Lista de tabelas... ..	xi
Lista de gráficos.....	xv
Lista de figuras	xvii
Lista de acrónimos.....	xix
1. Introdução.....	1
2. Materiais e métodos	5
2.1. Desenho do estudo.....	5
2.2. Seleção de casos e recolha de dados clínicos e histopatológicos	7
2.3. Seleção das amostras de tecido em parafina	8
2.4. Componente laboratorial	9
2.4.1. Remoção de parafina e reidratação da amostra de tecido.....	9
2.4.2. Extração de RNA	10
2.4.3. Síntese de cDNA.....	12
2.4.4. Estudo de expressão por RT-PCR (<i>Real Time PCR</i> - PCR em tempo real)	12
2.5. Validação dos resultados obtidos em âmbito laboratorial.....	16
2.6. Cálculo de diferenças de expressão, em termos relativos	16
2.7. Descrição das variáveis e métodos estatísticos.....	18
2.7.1. Testes de independência - qui-quadrado de <i>Pearson</i>	19
2.7.2. Testes não paramétricos	20
3. Resultados.....	23
3.1. Validação dos resultados obtidos em âmbito laboratorial.....	23
3.1.1. Resultados obtidos para a expressão do Gene <i>housekeeping</i> , <i>B-actina</i>	23
3.1.2. Resultados obtidos para a expressão da variante polimórfica da aromatase, <i>CYP19A1</i> Trp39Arg (T/C).....	27
3.1.3. Resultados obtidos para a expressão dos restantes genes (<i>GSTM1</i> e <i>GSTT1</i>)	30
3.2. Análise descritiva da amostra	32

3.3.	Estudo da associação entre a expressão de <i>CYP19A1</i> Trp39Arg (T/C) e variáveis qualitativas clínicas e histopatológicas	36
3.4.	Diferenças de expressão da variante polimórfica, <i>CYP19A1</i> Trp39Arg (T/C), entre grupos	38
3.4.1.	Estudo de média/distribuição de ΔC_T de <i>CYP19A1</i> Trp39Arg (T/C) entre as categorias das variáveis	38
3.4.2.	Cálculo do rácio de expressão de <i>CYP19A1</i> Trp39Arg entre as categorias das variáveis	47
4.	Discussão/Conclusão	51
	Referências	63
	Glossário.....	67
	Anexos	69
	Anexo 1	70
	Anexo 2	71
	Anexo 3	72

Lista de tabelas

Tabela 1 - Critérios de inclusão utilizados na seleção da amostra a estudar	7
Tabela 2 - Critérios de exclusão utilizados na seleção da amostra a estudar.....	7
Tabela 3 - Metodologia utilizada no processo de desparafinação e reidratação	10
Tabela 4 - Metodologia utilizada no processo de extração de RNA	11
Tabela 5 - <i>Primers</i> utilizados para o gene <i>housekeeping</i> e genes de interesse.....	13
Tabela 6 - Descrição da mistura utilizada para cada gene em estudo	14
Tabela 7 - Protocolos usados no termociclador <i>iQ™5 Multicolor Real-Time PCR Detection System (Bio-Rad®, Munich, Germany)</i> para os diferentes genes de interesse.....	15
Tabela 8 - Variáveis e suas categorias - consideradas na avaliação de diferenças de expressão em mRNA nos genes de interesse	18
Tabela 9 - Variáveis e suas categorias - consideradas para controlo do acaso e auxílio da validação da análise.....	19
Tabela 10 - Outras variáveis clínicas e histopatológicas e suas respectivas categorias - consideradas na avaliação de diferenças de expressão em mRNA nos genes de interesse	19
Tabela 11 - Casos processados pelas várias etapas descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR para gene <i>housekeeping</i> , <i>B-actina</i>	23
Tabela 12 - Descrição dos valores de tendência central da temperatura de <i>melting</i> da <i>B-actina</i> , antes e após a aplicação dos critérios de validação, para obtenção da amostragem definitiva	27
Tabela 13 - Casos validados (n=30) processados pelas várias etapas laboratoriais descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR, para <i>CYP19A1 Trp39Arg (T/C)</i>	27
Tabela 14 - Casos validados (n=30) processados pelas várias etapas laboratoriais descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR, para <i>GSTM1</i>	30

Tabela 15 - Casos validados (n=30) processados pelas várias etapas laboratoriais descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR, para <i>GSTT1</i>	31
Tabela 16 - Caraterísticas clínicas da amostra definitiva em estudo.....	33
Tabela 17 - Caraterísticas histopatológicas do tumor e gânglio sentinela e/ou outros gânglios da amostra definitiva em estudo	33
Tabela 18 - Recetores tumorais da amostra definitiva em estudo.....	34
Tabela 19 - Estadio clínico e avaliação TNM da amostra definitiva em estudo.....	34
Tabela 20 - Apresentação do nível de significância estatística (<i>p-value</i>) do teste de qui-quadrado de <i>Pearson</i> , que testa a associação da “Expressão de <i>CYP19A1</i> Trp39Arg” com as variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”	36
Tabela 21 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de <i>CYP19A1</i> Trp39Arg” e “Envolvimento do gânglio sentinela e/ou outros”.....	36
Tabela 22 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de <i>CYP19A1</i> Trp39Arg” e “Estadio clínico”	36
Tabela 23 - Apresentação do nível de significância estatística (<i>p-value</i>) do teste de qui-quadrado de <i>Pearson</i> , que testa a associação da “Expressão de <i>CYP19A1</i> Trp39Arg” com as variáveis “Mama afetada” e “Localização afetada”	37
Tabela 24 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de <i>CYP19A1</i> Trp39Arg” e “Mama afetada”	37
Tabela 25 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de <i>CYP19A1</i> Trp39Arg” e “Localização afetada”	37
Tabela 26 - Descrição das variáveis, nos casos em que ocorreu expressão de <i>CYP19A1</i> Trp39Arg, consideradas para cálculo de diferenças na expressão da variante polimórfica da aromatase.....	47
Tabela 27 - Diferença de expressão de <i>CYP19A1</i> Trp39Arg nas categorias das variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”	48
Tabela 28 - Diferença de expressão de <i>CYP19A1</i> Trp39Arg nas categorias da variável “RE (Recetores de estrogénio) ”	48

Tabela 29 - Diferença de expressão de <i>CYP19A1</i> Trp39Arg nas categorias das variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”, apenas considerando casos RE+.....	48
Tabela 30 - Diferença de expressão de <i>CYP19A1</i> Trp39Arg em subcategorias da categoria “Metastização” da variável “Envolvimento do gânglio sentinela e/ou outros”	49
Tabela 31 - Diferença de expressão de <i>CYP19A1</i> Trp39Arg nas categorias das variáveis “controlo do acaso”, “Mama afetada” e “Localização afetada”.....	49

Lista de gráficos

Gráfico 1 - Exemplificação de curvas de deteção de fluorescência ao longo dos vários ciclos de amplificação em RT-PCR	15
Gráfico 2 - Diferença entre mediana de temperatura de <i>melting</i> e temperatura de <i>melting</i> de cada caso da β -actina.....	25
Gráfico 3 - <i>Boxplot</i> da distribuição das temperaturas de <i>melting</i> da β -actina dos 37 casos, antes da aplicação dos critérios de validação	26
Gráfico 4 - <i>Boxplot</i> da distribuição das temperaturas de <i>melting</i> da β -actina dos 30 casos, após a aplicação dos critérios de validação.....	26
Gráfico 5 - Representação das curvas de <i>melting</i> de um controlo - NTC (rosa) e um caso (azul) de <i>CYP19A1</i> Trp39Arg	29
Gráfico 6 - Gráfico circular ilustrando a expressão da variante polimórfica da aromatase, <i>CYP19A1</i> Trp39Arg (T/C), na amostragem definitiva em estudo	35
Gráfico 7 - <i>Boxplot</i> em que é ilustrada a distribuição de ΔC_T de <i>CYP19A1</i> Trp39Arg entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”	39
Gráfico 8 - <i>Boxplot</i> em que é ilustrada a distribuição de ΔC_T de <i>CYP19A1</i> Trp39Arg entre as categorias, “I e II” e “III”, da variável “Estadio clínico”	39
Gráfico 9 - <i>Boxplot</i> em que é ilustrada a distribuição de ΔC_T <i>CYP19A1</i> Trp39Arg entre as categorias, “Direita” e “Esquerda”, da variável “Mama afetada”.....	41
Gráfico 10 - <i>Boxplot</i> em que é ilustrada a distribuição de ΔC_T de <i>CYP19A1</i> Trp39Arg entre as categorias, “Quadrante Supero-externo” e “Outra”, da variável “Localização afetada”	41
Gráfico 11 - <i>Boxplot</i> em que é ilustrada a distribuição de ΔC_T de <i>CYP19A1</i> Trp39Arg entre as categorias, “RE+ (Positivo) ” e “RE- (Negativo) ”, da variável “RE (Recetores de estrogénio)”.....	43
Gráfico 12 - <i>Boxplot</i> em que é ilustrada a distribuição de ΔC_T de <i>CYP19A1</i> Trp39Arg entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”, quando RE+	44
Gráfico 13 - <i>Boxplot</i> em que é ilustrada a distribuição de ΔC_T de <i>CYP19A1</i> Trp39Arg entre as categorias, “I e II” e “III”, da variável “Estadio clínico”, quando RE+.....	45

Gráfico 14 - *Boxplot* em que é ilustrada a distribuição de ΔC_T de *CYP19A1* Trp39Arg entre as subcategorias, “Metastização ganglionar, N2-3” e “Metastização ganglionar, N1”, da categoria “Metastização” da variável “Envolvimento do gânglio sentinela e/ou outros”, para casos RE+.... 46

Lista de figuras

Figura 1 - Principais etapas desenvolvidas ao longo do trabalho	6
Figura 2 - Esquematização da ligação do <i>reverse primer</i> à sequência que contém o polimorfismo <i>CYP19A1</i> Trp39Arg (T/C) e não com a sequência <i>wild</i> do gene da aromatase.....	13
Figura 3 - Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”	38
Figura 4 - Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma entre as categorias, “I e II” e “III”, da variável “Estadio clínico”.	38
Figura 5 - Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma entre as categorias, “Direita” e “Esquerda”, da variável “Mama afetada”	40
Figura 6 - Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma entre as categorias, “Quadrante Supero-externo” e “Outra”, da variável “Localização afetada”	40
Figura 7 - Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma entre as categorias, “RE+ (Positivo) ” e “RE- (Negativo) ”, da variável “RE (Recetores de estrogénio)”	42
Figura 8 -Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”, quando RE+	43
Figura 9 - Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma entre as categorias, “I e II” e “III”, da variável “Estadio clínico”, quando RE+	44
Figura 10 - Teste <i>Mann-Whitney U</i> , em que a hipótese nula testada é que a média de ΔC_T de <i>CYP19A1</i> Trp39Arg é a mesma nas subcategorias, “Metastização ganglionar, N2-3” e “Metastização ganglionar, N1”, da categoria “Metastização” da variável “Envolvimento do gânglio sentinela e/ou outros”, quando RE+	46

Figura 11 - Hipotética formação de homo-dímeros no *forward primer* de *CYP19A1* Trp39Arg (5'-ATCTGTACTGTACAGCACC-3') com 6 pares de bases..... 54

Figura 12 - Hipotética formação de homo-dímeros no *forward primer* de *CYP19A1* Trp39Arg (5'-ATCTGTACTGTACAGCACC-3') com 8 pares de bases..... 54

Lista de acrónimos

ΔC_T	Diferença entre C_T do gene de interesse e C_T do gene <i>housekeeping</i>
cDNA	DNA complementar
CHCB	Centro Hospitalar Cova da Beira
C_T	Ciclo limiar
DEPC	Dicarbonato de dietila
GLOBOCAN	<i>Global Burden of Cancer Study</i>
GSH	Glutathiona
GST	Glutathiona S-transferase
GSTM1	Glutathiona S-transferase M1
GSTT1	Glutathiona S-transferase T1
M-MLV	<i>Moloney-Murine Leukaemia Virus</i>
mRNA	RNA mensageiro
qPCR	PCR quantitativo
RE	Recetores de estrogénio
RE+	Recetores de estrogénio positivos
RE-	Recetores de estrogénio negativos
Rpm	Rotações por minuto
RT-PCR	Reação em cadeia de polimerase em tempo real
SAM	Sistema de Apoio ao Médico
SNP	Polimorfismo de nucleótido único
TM	Temperatura de <i>melting</i>

1. Introdução

O cancro da mama é uma entidade nosológica de grande impacto mundial. De acordo com os dados publicados no GLOBOCAN em 2008, cujos resultados são referidos por *Jemal et al.* em 2011, o cancro da mama foi a patologia oncológica mais diagnosticada em mulheres, em todo o mundo, representando 14% de todas as mortes por cancro, no sexo feminino (1). São diversos os fatores de risco, quer endógenos, quer exógenos, que aumentam a suscetibilidade do seu desenvolvimento. O papel do património genético é considerado importante, estimando-se que cerca de 30% dos casos sejam de origem hereditária (2). Genes de alta penetrância, como *BRCA1* e *BRCA2*, conferem alto risco de desenvolver a doença (3), no entanto, apenas estão associados a cerca de 5% de todos os cancros da mama (4). Este facto tem suscitado um interesse crescente para o estudo de outros genes, que apesar de ditarem menor risco individual para desenvolver esta patologia, são mais comuns no património genómico das populações (5). Tais genes, denominados de baixa penetrância, incluem na maior parte dos casos polimorfismos, responsáveis por diferenças interindividuais no metabolismo de estrogénios, xenobióticos, espécies reativas de oxigénio, entre outras (6). A variabilidade da sua expressão pode explicar as diferentes respostas de cada indivíduo, podendo desta forma influenciar a suscetibilidade para uma progressão oncológica.

A importância dos estrogénios exógenos na carcinogénese do cancro da mama ficou bem estabelecida no estudo prospetivo *Women's Health Initiative*, no qual foi observado uma associação positiva entre a administração de estrogénios e o risco de desenvolver esta patologia (7, 8). Desde então, os estrogénios são considerados agentes carcinogénicos (9). A ação carcinogénica dos estrogénios ocorre tanto pela promoção da proliferação celular como pela inibição de mecanismos apoptóticos (9), para além de lesar o DNA através dos seus metabolitos e espécies reativas de oxigénio, as quais originam mutações que posteriormente despoletam a carcinogénese (10).

Tem-se constatado que mulheres com cancro da mama têm níveis relativamente altos de adutos de DNA de estrogénio. Enzimas antioxidantes, como a GST (glutathione S-transferase), são de grande importância porque neutralizam estes metabolitos por conjugação com a GSH (glutathione), conferindo assim proteção contra a lesão de DNA e subsequentemente contra a oncogénese (10, 11).

Também é de referir que fatores exógenos, como a radiação ionizante e xenobióticos (12) podem, de igual modo, originar processos de mutagénese por *stress* oxidativo (13). Estima-se que a frequência diária de lesões do DNA por *stress* oxidativo seja de 1000 lesões/célula/dia, sendo os radicais OH (hidroxilo) os predominantes agentes causais (12). As enzimas do complexo GST participam na eliminação destes radicais OH (14), assim como no metabolismo

de xenobióticos, pois ao participarem em reações de fase II permitem uma destoxificação em compostos hidrossolúveis que são mais facilmente eliminados (12).

Foi já demonstrado pelo nosso grupo de trabalho, numa amostra da população feminina inscrita no Centro Hospitalar Cova da Beira (CHCB), uma associação positiva entre genes de baixa penetrância e o risco aumentado de desenvolver cancro da mama (15, 16). Um desses genes codifica a aromatase e o polimorfismo estudado foi *CYP19A1* Trp39Arg (T/C). Este é um *missense SNP* (Polimorfismo de nucleótido único), em que há troca de timina por citosina, originando a substituição de um triptofano por arginina, no codão 39 (17). A aromatase participa no passo final da biossíntese de estrogénios, convertendo os androgénios a estrogénios (18). As mulheres portadoras do alelo variante da aromatase apresentaram um aumento do risco de cancro da mama de quase duas vezes (OR=1,770; p=0,036) (15). Os outros genes estudados foram a *GSTM1* e *GSTT1*. Nestes genes, os polimorfismos estudados foram deleções homocigóticas, i.e., genótipo *null* para ambos. A presença destes polimorfismos representou um aumento do risco para cancro da mama em cerca de 3,6 vezes para o genótipo *null* da *GSTM1* (OR= 3,597; p=0,0001) e cerca de 2,6 vezes para o genótipo *null* da *GSTT1* (OR= 2,592; p= 0,002) (16).

Desta forma, considerou-se oportuno aprofundar o conhecimento sobre estas variantes (*CYP19A1* Trp39Arg (T/C), *GSTM1 null* e *GSTT1 null*), nesta mesma população, estudando os seus níveis de expressão através de mRNA (RNA mensageiro).

Os objetivos deste trabalho preliminar foram:

- 1- Avaliar o interesse prognóstico da expressão dos polimorfismos *CYP19A1* Trp39Arg (T/C), *GSTM1 null* e *GSTT1 null* no cancro da mama.
- 2- Desenvolver/Otimizar uma técnica laboratorial que facilite o estudo da expressão através de mRNA dos polimorfismos *CYP19A1* Trp39Arg (T/C), *GSTM1 null* e *GSTT1 null* utilizando amostras de tecido de mama fixado e incluído em parafina.

Tem-se assim a expectativa de dotar o corpo clínico de possíveis meios, que acrescentem uma mais-valia na abordagem e tomada de decisões de novos casos de cancro da mama na região.

Tendo em conta os objetivos descritos, estabelecem-se as seguintes hipóteses:

- H₁ - O polimorfismo da aromatase, *CYP19A1* Trp39Arg (T/C), é mais expresso em grupos de pior prognóstico de cancro da mama.
- H₂- O polimorfismo *GSTM1 null* é mais expresso em grupos de pior prognóstico de cancro da mama.
- H₃- O polimorfismo *GSTT1 null* é mais expresso em grupos de pior prognóstico de cancro da mama.

A realização deste estudo (Projeto de Investigação nº22/2013) foi autorizada, em 13-03-2013, por reunião do Conselho de Administração do Centro Hospitalar Cova da Beira. Salienta-se ainda que está eticamente aprovado por pareceres favoráveis da Comissão de Ética, da mesma instituição, nº 24/2008 e nº 321/2011. (ver anexo 1)

2. Materiais e métodos

2.1. Desenho do estudo

Para alcançar os objetivos a que nos propusemos desenvolvemos um estudo transversal do tipo caso-controlo. Pretendemos comparar as possíveis diferenças de expressão génica, originadas pelos polimorfismos supramencionados, entre pacientes com cancro da mama com diferente prognóstico, numa amostra de mulheres tratadas e seguidas na unidade de Ginecologia Oncológica do CHCB. O prognóstico foi avaliado com base no estadió clínico e na metastização regional, determinada através da pesquisa do gânglio sentinela ou de linfadectomia axilar, que constituíram assim, as nossas variáveis de interesse. Com base nestas variáveis dividimos as doentes em grupos: para o estadió clínico incluímos as doentes com estadió I e II no grupo de melhor prognóstico e as doentes de estadió III no grupo de pior prognóstico; para a metastização ganglionar regional incluímos as doentes sem metastização num grupo de melhor prognóstico, enquanto que as doentes com metastização incluímos num de pior prognóstico. Os grupos caso, consistiram em grupos de mulheres de pior prognóstico, enquanto os grupos controlo corresponderam a grupos de mulheres de melhor prognóstico.

O trabalho consistiu em quatro momentos fundamentais, citados a seguir, estando sumariadas, no esquema da figura 1, as principais etapas envolvidas.

- 1 - Seleção de casos passíveis de entrar no estudo com respetiva análise e recolha de dados. Paralelamente foram desenvolvidos e otimizados os procedimentos metodológicos laboratoriais;
- 2 - Efetuada componente laboratorial, em que foi realizada desparafinação das amostras, extração de mRNA, com respetiva síntese de cDNA (DNA complementar) e realização de RT-PCR (Reação em cadeia de polimerase em tempo real);
- 3 - Validação dos resultados obtidos em âmbito laboratorial e conseqüente análise descritiva da casuística ratificada;
- 4 - Análise dos dados com recurso a métodos estatísticos e cálculo do rácio de expressão entre grupos.

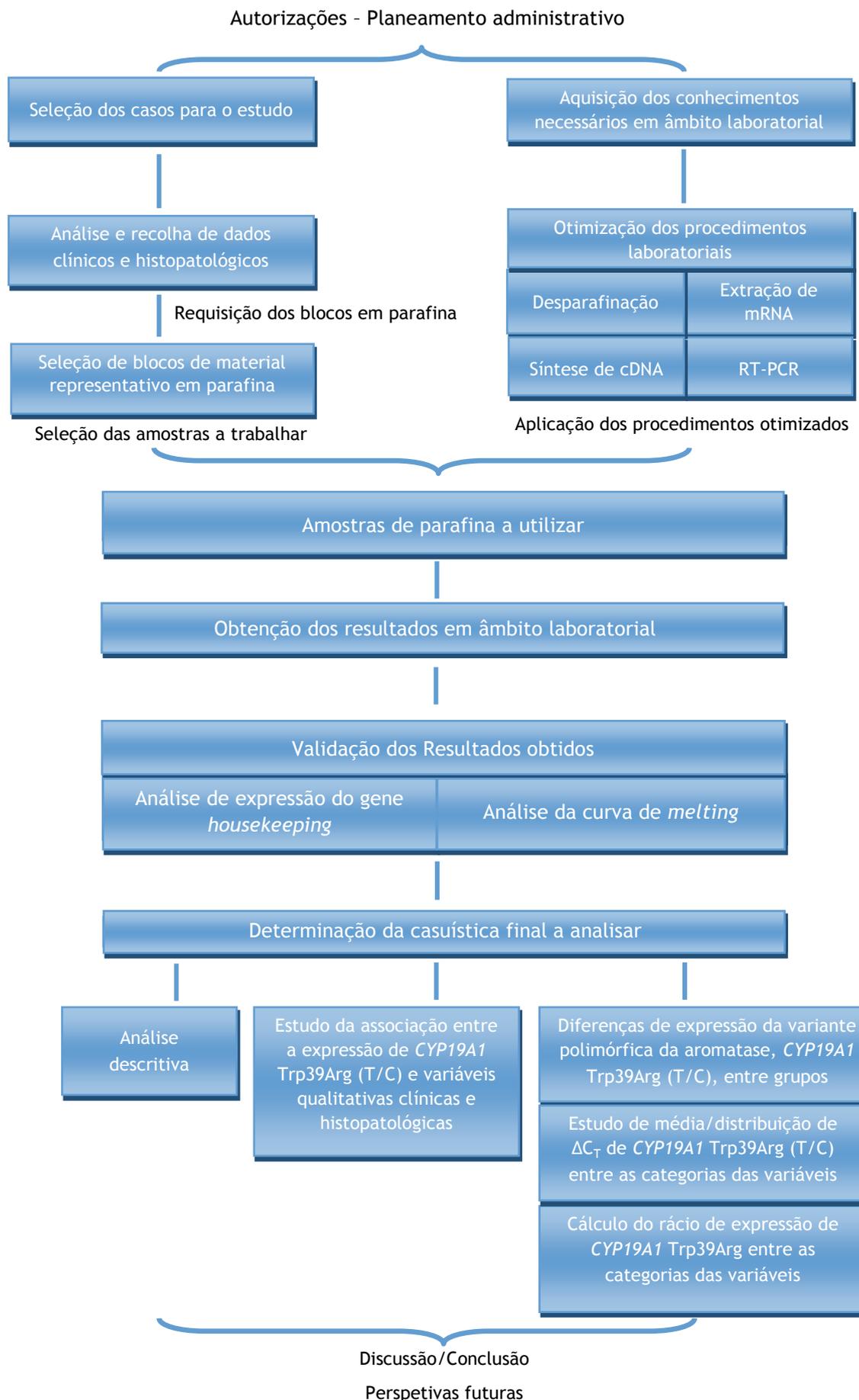


Figura 1 - Principais etapas desenvolvidas ao longo do trabalho

2.2. Seleção de casos e recolha de dados clínicos e histopatológicos

A população alvo deste estudo foi o grupo de mulheres inscritas no Centro Hospitalar Cova da Beira, com diagnóstico histopatológico de carcinoma da mama, entre os anos 2006 e 2011.

Foram previamente estabelecidos os critérios de inclusão e exclusão, listados nas tabelas 1 e 2, respetivamente.

Tabela 1 - Critérios de inclusão utilizados na seleção da amostra a estudar

Critérios de inclusão
Ser do género feminino
Estar inscrita no CHCB
Diagnóstico de cancro da mama, com ou sem metastização regional, ao gânglio sentinela e/ou outros, estabelecido por relatório histopatológico, entre 2006 a 2011
Realização de mastectomia ou tumorectomia com pesquisa de gânglio sentinela e/ou esvaziamento ganglionar regional, com relatório da anatomia patológica

Tabela 2 - Critérios de exclusão utilizados na seleção da amostra a estudar

Critérios de exclusão
Suspeita ou confirmação de metastização à distância
Realização de qualquer terapêutica neoadjuvante previamente aos procedimentos cirúrgicos citados na tabela 1

Para uma dada paciente ser considerada para o estudo, todos os quatro critérios de inclusão tiveram de ser cumpridos. Se pelo menos um critério de exclusão esteve presente, mesmo apesar de cumpridos todos os de inclusão, esse caso foi excluído.

A pesquisa dos casos foi realizada através da consulta de processos clínicos, em suporte informático, com recurso à ferramenta SAM (Sistema de Apoio ao Médico), nas infraestruturas do CHCB. Inicialmente foi realizado um levantamento de todas as mulheres inscritas no CHCB com o referido diagnóstico, entre os anos 2006 e 2011 e que efetuaram tumorectomia ou mastectomia com pesquisa de gânglio sentinela e/ou esvaziamento ganglionar regional. Posteriormente foram analisados os relatórios histo-anátomo-patológicos das peças excisadas na cirurgia e de biópsias diagnósticas, realizadas previamente à cirurgia. Diversas informações sobre o tumor foram recolhidas, tipo histológico, recetores de estrogénio, progesterona e c-erb-b2, apresentação uni ou multifocal, tamanho, grau de diferenciação, invasão linfovascular, envolvimento da pele e/ou músculo, componente *in situ*, calcificação tumoral e envolvimento do gânglio sentinela e/ou outros.

Quando existiu suspeita clínica ou comprovação de metastização à distância, por exames complementares de diagnóstico realizados para investigação do estadio clínico da paciente, esta foi excluída do estudo. A realização de qualquer terapêutica neoadjuvante também motivou exclusão.

As considerações clínicas relativas a cada paciente alvo de registo foram a idade de diagnóstico, mama afetada e localização afetada.

Foram assim elegíveis para participar no estudo 78 mulheres.

2.3. Seleção das amostras de tecido em parafina

A requisição de amostras em parafina de tecido de mama “normal”, de peças excisadas cirurgicamente na unidade de Ginecologia Oncológica do CHCB, seguiu-se à seleção das pacientes a participar no estudo. Este pedido foi efetuado ao Serviço de Anatomia Patológica do CHCB. A decisão de requisitar mama histologicamente “normal”, i.e., que não contivesse células tumorais, baseou-se no facto de pretendermos estudar quais as influências que o tecido de mama normal, envolvente a um processo tumoral, realiza ou está sujeito.

Prepararam-se diversas lâminas para visualizar a histologia do tecido, com recurso a microscopia ótica, de modo a identificar blocos de parafina que contivessem amostras de tecido de mama normal.

No Serviço de Anatomia Patológica do CHCB foram realizados cortes por micrótomo com 10µm de espessura nos blocos de parafina selecionados. As amostras em parafina resultantes foram conservadas em tubos de *Eppendorf* e guardadas à temperatura ambiente.

Salienta-se ainda que os casos foram codificados pelo Serviço de Anatomia Patológica, para que o investigador, no decorrer do trabalho em âmbito laboratorial, não fosse conhecedor da correspondência amostra de tecido parafinado - caso estudado. Só no final da obtenção dos resultados laboratoriais e respetiva validação foi decifrada esta chave.

2.4. Componente laboratorial

O objetivo principal de todos os procedimentos realizados nesta etapa foi a quantificação relativa da expressão dos genes em estudo.

Para tal, procedeu-se à extração de RNA do tecido, previamente desparafinado e reidratado. Posteriormente foi efetuada a síntese de cDNA partindo do mRNA obtido, com subsequente amplificação por RT-PCR. A expressão foi avaliada de forma relativa, tendo como gene *housekeeping*, a β -actina, que codifica uma proteína presente no citoesqueleto das células.

Para o trabalho a apresentar nesta dissertação decidimos analisar, primeiro, os casos que continham amostras de tecido em parafina em duplicado (37 de 78 casos), uma vez que a Anatomia Patológica enviou amostras com 2 cortes de microtomia e outras com apenas 1 corte. Deste modo, foi assegurada a preservação do banco de dados para repetição e/ou continuação do estudo e, também, permitiu estabelecer critérios de seleção dos casos a trabalhar.

2.4.1. Remoção de parafina e reidratação da amostra de tecido

Esta metodologia baseou-se na utilização do solvente orgânico, xilol, responsável por dissolver a parafina envolvente ao tecido, com posterior reidratação da amostra. Esta última foi efetuada com soluções de concentração decrescente de etanol em água DEPC (dicarbonato de dietila). Um passo intermédio consistiu na utilização de etanol absoluto para remoção dos excedentes de xilol (19).

De forma a otimizar a remoção, da maior quantidade possível, de parafina, sem danificar o tecido por este solvente orgânico, foram tentados diversos tempos e temperaturas de incubação. Vários processos de homogeneização do xilol com a amostra de parafina também foram testados.

A reidratação progressiva da amostra foi também alvo de otimização, com a utilização de diversas sequências de soluções de concentração decrescente de etanol em água DEPC.

Em baixo, na tabela 3, é descrito todo o encadeamento metodológico utilizado neste procedimento laboratorial, estando já expostos todos os passos de otimização realizados.

Tabela 3 - Metodologia utilizada no processo de desparafinação e reidratação

Descrição das etapas
1- Colocou-se a amostra de tecido de parafina num tubo de <i>Eppendorf</i> de 1,5ml
2- Adicionou-se 1,2ml de xilol à amostra
3- Ressuspendeu-se com o auxílio de uma ponta de pipeta cortada
4- Agitou-se durante 15segundos em vórtex a 1500rpm (rotações por minuto)
5- Incubou-se durante 10minutos a temperatura ambiente, sendo as amostras agitadas de modo suave, manualmente, em intervalos de 2-3minutos
6- Incubou-se durante 5min a 50 °C num bloco de aquecimento
7- Centrifugou-se a 13000rpm, durante 5minutos
8- Retirou-se o máximo de sobrenadante possível
9- Repetição dos passos 2-8, por mais 2 vezes
10- Adicionou-se 1,2ml de etanol absoluto
11- Agitou-se por 15segundos em vórtex a 1500rpm
12- Incubou-se durante 10 minutos a temperatura ambiente
13- Centrifugou-se a 13000rpm, durante 5minutos
14- Retirou-se o máximo sobrenadante possível
15- Adicionou-se solução de etanol 95% em água DEPC
16- Agitou-se por 15segundos em vórtex a 1500rpm
17- Incubou-se durante 10 minutos a temperatura ambiente
18- Centrifugou-se a 13000rpm, durante 5minutos
19- Retirou-se o máximo sobrenadante possível
20- Adicionou-se solução de etanol 75% em água DEPC
21- Agitou-se por 15segundos em vórtex a uma velocidade de 1500rpm
22- Incubou-se durante 10 minutos a temperatura ambiente
23- Centrifugou-se a 13000rpm, durante 5minutos
24- Retirou-se o máximo sobrenadante possível
25- Após evaporação durante um período de 12-15h a temperatura ambiente e em local protegido de fluxo de ar horizontal e vertical descendente, procedeu-se à extração de RNA

2.4.2. Extração de RNA

A extração de RNA foi realizada com recurso ao reagente TRIzol® (*Ambion*®, *Life Technologies*™, *Paisley, UK*), que permite o isolamento de RNA com base no método clássico de extração por tiocianato de guanidina - fenol - clorofórmio.

Uma vez que os métodos de fixação de tecidos e subsequente parafinação são altamente agressivos para os ácidos nucleicos, sobretudo RNA, a quantidade expectável a ser extraída é diminuta (20). Desta forma o método utilizado, baseou-se no protocolo descrito pelo nosso grupo para a extração de RNA (21), a partir de sangue colhido em cartão FTA®, tendo sido alterado por forma a aumentar a eficiência de extração deste ácido nucleico.

Estas principais modificações foram efetuadas no método de mistura de TRIzol® com o tecido desparafinado e nos tempos de centrifugação. Foi também acrescentado o passo 16 da tabela 4, que resume a metodologia utilizada no processo de extração de RNA, para assegurar que reagentes, como solventes orgânicos e álcoois, fossem eliminados por evaporação, garantindo assim uma menor interferência na posterior síntese de cDNA.

Tabela 4 - Metodologia utilizada no processo de extração de RNA

Descrição das etapas
1- Adicionou-se 1ml de reagente TRIzol® à amostra de tecido desparafinado e reidratada, após confirmação da total evaporação de etanol
2- Ressuspendeu-se para maior homogeneização da amostra
3- Incubou-se por 8 minutos a temperatura ambiente
4- Adicionou-se 0,2ml de clorofórmio
5- Agitou-se por 15 segundos em vórtex a 1500rpm
6- Incubou-se por 3 minutos a temperatura ambiente
7- Centrifugou-se a 14000rpm, durante 15 minutos, a 4 °C
8- Retirou-se 0,5ml da fase aquosa para um novo tubo de <i>Eppendorf</i> de 1,5ml
9- Adicionou-se 0,5ml de isopropanol seguido de uma ligeira agitação dos tubos de <i>Eppendorf</i>
10- Incubou-se por 10 minutos a temperatura ambiente
11- Centrifugou-se a 14000rpm, por 20 minutos, a 4 °C
12- Eliminou-se o sobrenadante
13- Adicionou-se 0,5ml de solução de etanol a 75% em água DEPC
14- Centrifugou-se durante 5 minutos a 11000rpm, a 4 °C
15- Eliminou-se o sobrenadante
16- Colocaram-se amostras a secar, por cerca de 5 minutos a temperatura ambiente e reaquecidas a 56°C, se necessário ¹
17- Adicionou-se 20µl de água tratada com DEPC, previamente aquecida a 57 °C
18- Procedeu-se à síntese de cDNA.

¹-Quando evidência de etanol residual.

Nota: A quantificação do RNA extraído foi tentada diversas vezes através da leitura da absorvência a 260nm, no aparelho *NanoPhotometer™* (*Implen GmbH, Munchen, Germany*). No entanto os resultados obtidos por estas medições foram díspares, existindo em algumas amostras grande variabilidade dos resultados. Optámos assim pela não utilização sistemática deste aparelho que permite quantificar ácidos nucleicos, colocando em causa a sua sensibilidade e/ou funcionamento.

2.4.3. Síntese de cDNA

Para a síntese de cDNA efetuada a partir do mRNA extraído, este foi transcrito com M-MLV (*Moloney-Murine Leukaemia Virus*), uma transcriptase reversa comercial. Num primeiro passo efetuou-se uma mistura de reação com os *Random primer*, que continha:

- 2µl de *Random primer* (Promega®, Madison, USA)
- 1µl de dNTPs (100 nM, Sigma-Aldrich®, St. Louis, USA)
- 9µl de extrato de RNA

Esta etapa decorreu a 70°C durante 5min garantindo um arrefecimento final a 4°C. Este processo foi realizado no termociclador *TProfessional Basic* (Biometra®, Gottingen, Germany).

Para o segundo passo efetuou-se uma mistura constituída por:

- n x 4µl de solução tampão apropriada - 5xM-MLV “*Reverse Transcriptase (RT) Buffer*” (Promega®, Madison, USA)
- n x 2µl Ditioneitol (DTT) (0,1M) (Sigma-Aldrich®, St. Louis, USA)
- n x 1µl de inibidor de RNase (Promega®, Madison, USA)
- n x 0,5µl M-MLV RT (Promega®, Madison, USA)

Sendo que (n) representa o número de reações de síntese de cDNA a realizar.

Posteriormente 7,5µl desta solução foi adicionada a cada mistura do primeiro passo.

A síntese de cDNA foi realizada no mesmo termociclador, com o seguinte programa:

- 10min a 25°C - incubação inicial
- 60min a 37°C - fase de extensão
- 15min a 70°C - inativação da enzima

O cDNA sintetizado foi armazenado em arca frigorífica a uma temperatura de 4°C.

2.4.4. Estudo de expressão por RT-PCR (*Real time PCR* - PCR em tempo real)

De modo a conseguir avaliar as diferenças na expressão recorreu-se à técnica PCR em tempo real, que permite quantificar o cDNA, à medida que ocorre a reação de amplificação do mesmo. Por isso denomina-se PCR em tempo real, também sendo chamado qPCR (PCR quantitativo). A quantificação do número de cópias presentes é obtida através de um marcador que, ao ligar-se ao cDNA na forma de cadeia dupla, emite um sinal de fluorescência. O marcador fluorescente que usámos foi *SYBR® Green*.

Por forma a amplificar as variantes polimórficas em estudo, bem como o gene escolhido para referência, gene *housekeeping*, foram usados os *primers* descritos na tabela 5.

Tabela 5 - *Primers* utilizados para o gene *housekeeping* e genes de interesse

Gene	Primer	TM (°C)	Proveniência/Fabricante
β-actina	<i>Forward</i>	5'-TTCCTGGGCATGGAGTCCT-3'	53,2 Stab Vida®, Setúbal, Portugal
	<i>Reverse</i>	5'-AGGAGGAGCAATGATCTTGATC-3'	53 Stab Vida®, Setúbal, Portugal
GSTM1	<i>Forward</i>	5'-GAACTCCCTGAAAAGCTAAAG-3'	50,5 Stab Vida®, Setúbal, Portugal
	<i>Reverse</i>	5'-GTTGGGCTCAAATATACGGTGG-3'	54,8 Stab Vida®, Setúbal, Portugal
GSTT1	<i>Forward</i>	5'-TTCCTTACTGGTCCTCACATCTC-3'	55,3 Stab Vida®, Setúbal, Portugal
	<i>Reverse</i>	5'-TCACCGGATCATGGCCAGCA-3'	55,9 Stab Vida®, Setúbal, Portugal
CYP19A1	<i>Forward</i>	5'-ATCTGTACTGTACAGCACC-3'	48,9 Stab Vida®, Setúbal, Portugal
Trp39Arg	<i>Reverse</i>	5'-ATGTGCCCTCATAATTCCG-3'	48,9 Stab Vida®, Setúbal, Portugal
CYP19A1	<i>Forward</i>	5'-GGCCTTTTCTCTTGGTGT-3'	48,9 Stab Vida®, Setúbal, Portugal
wild type	<i>Reverse</i>	5'-CTCCAAGTCCTCATTTGCT-3'	48,9 Stab Vida®, Setúbal, Portugal

TM - Temperatura de *melting*

Uma vez que estes *primers* tinham sido adquiridos para serem utilizados em técnicas de PCR com DNA genómico, considerámos importante confirmar através do *software* informático *Ensembl*, a localização dos mesmos, i.e., intrão ou exão.

Chegou-se à conclusão que o *reverse primer* para a determinação do genótipo *wild type* da aromatase não estava corretamente desenhado para a realização de RT-PCR.

Confirmámos ainda que o *reverse primer* utilizado para a determinação da variante *CYP19A1* Trp39Arg é específico para a mesma (figura 2).

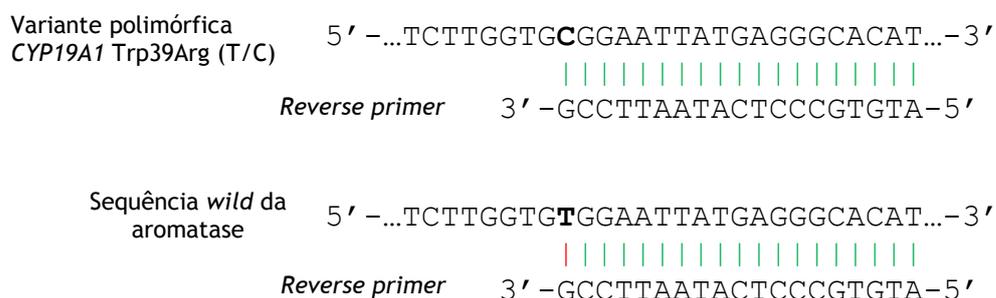


Figura 2 - Esquematização da ligação do *reverse primer* à sequência que contém o polimorfismo *CYP19A1* Trp39Arg (T/C) e não com a sequência *wild* do gene da aromatase

A negrito está a troca de nucleótidos T (*wild*) para C (variante polimórfica em estudo). A verde estão representadas as ligações entre bases complementares. A vermelho está a ligação não efetuada do *reverse primer* com a sequência *wild*.

Nos casos da *GSTT1* e *GSTM1*, como as variantes de interesse são deleções homozigóticas em ambos os casos, não existe nenhuma consideração especial a ter em conta com um determinado nucleótido dos seus *primers*.

Com o intuito de aumentar a eficiência de amplificação, procedemos à sua otimização para os genes *GSTT1*, *GSTM1* e *CYP19A1* Trp39Arg. Em relação às condições descritas pelo nosso grupo, foi necessário alterar a temperatura da fase de *annealing*.

Os genes de interesse e a β -actina (gene *housekeeping*) foram amplificados em simultâneo em cada amostra de cDNA. Deste modo foi-nos permitido comparar a expressão dos primeiros ao gene *housekeeping* e também garantir que tínhamos realmente cDNA.

Para o estudo de expressão por RT-PCR, a metodologia usada foi a seguinte:

- 1- Preparou-se uma mistura para cada gene, cujos constituintes estão descritos na tabela 6.

Tabela 6 - Descrição da mistura utilizada para cada gene em estudo

Mistura de reação para o RT-PCR
(n + 1) x 10 μ l de uma solução de SYBR® Green/ROX qPCR Master Mix (2x) (Fermentas - Thermo Fisher Scientific Inc., Vilnius, Lithuania)
(n + 1) x 0,5 μ l de uma solução 10nM de <i>forward primer</i>
(n + 1) x 0,5 μ l de uma solução de 10nM de <i>reverse primer</i>
(n + 1) x 6 μ l água livre de DNase (<i>DNase free</i>) (Promega®, Madison, USA)

n - número de amostras a realizar

- 2- Adicionou-se 3 μ l de cDNA sintetizado a 17 μ l da mistura referida na tabela 6, num poço da placa de RT-PCR (Bio-Rad®, Munich, Germany).

Para cada gene foi adicionado um controlo negativo, também com 17 μ l da mistura e, em vez de cDNA, foram colocados 3 μ l água livre de DNase (Promega®, Madison, USA).

- 3- Uma vez selada com uma película, a placa foi colocada no termociclador *iQ™5 Multicolor Real-Time PCR Detection System* (Bio-Rad®, Munich, Germany).

Tendo em conta que existe uma diferença significativa na temperatura de *annealing* dos *primers* de *CYP19A1*Trp39Arg comparativamente a *GSTT1* e *GSTM1* foram usados, em simultâneo, dois protocolos distintos, tal como se verifica na tabela 7.

Tabela 7 - Protocolos usados no termociclador *iQ™5 Multicolor Real-Time PCR Detection System (Bio-Rad®, Munich, Germany)* para os diferentes genes de interesse

CYP19A1 Trp39Arg		GSTT1 e GSTM1	
Desnaturação inicial de DNA a 95 °C por 30 minutos			
60 ciclos de amplificação das 3 fases (1 - desnaturação, 2 - <i>annealing</i> e 3 - extensão):			
1 - Desnaturação a 95 °C por 10 segundos			
2- <i>Annealing</i> a 55 °C por 30 segundos		2- <i>Annealing</i> a 58 °C por 30 segundos	
3 - Extensão a 72 °C por 30 segundos			

A leitura de fluorescência foi sendo detetada, ciclo após ciclo, no decorrer da reação de amplificação e registada em gráfico, com a respetiva determinação do valor de C_T (ciclo limiar), para cada amostra. Esta deteção foi visualizada com recurso ao *software iQ™5 v 2.1 (Bio-Rad®, Munich, Germany)*, tal como se pode observar no gráfico 1.

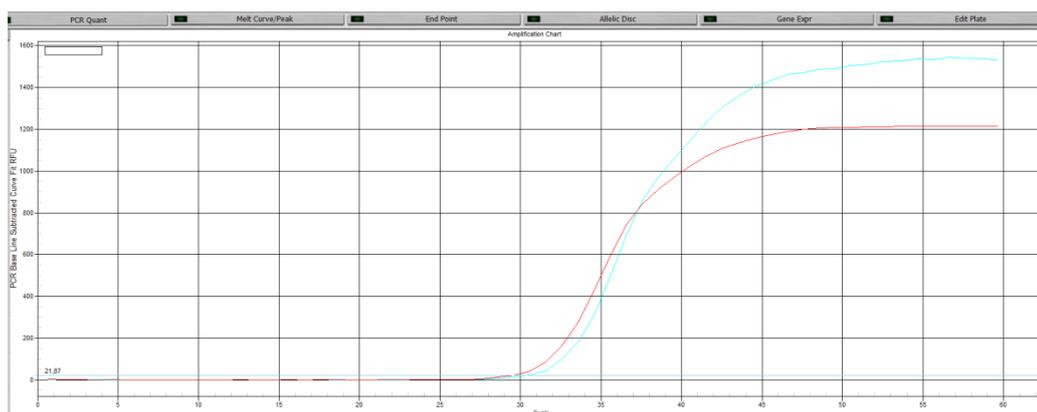


Gráfico 1 - Exemplicação de curvas de detecção de fluorescência ao longo dos vários ciclos de amplificação em RT-PCR

No eixo das abcissas está o número de ciclos, enquanto no eixo das ordenadas o valor de fluorescência. São evidenciados as curvas de 2 amostras para o gene *CYP19A1 Trp39Arg (T/C)*.

Foi efetuada uma curva de *melting* de forma a validar os resultados obtidos. Para a obtenção dessa curva recorreremos, no final de cada amplificação, a um programa com uma amplitude de temperatura dos 55 °C aos 95 °C e aumentos graduais de 0,5 °C a cada 10 segundos.

2.5. Validação dos resultados obtidos em âmbito laboratorial

Uma fase fundamental do nosso trabalho consistiu na validação dos resultados obtidos no decorrer do trabalho laboratorial. Para tal foi efetuada uma análise exaustiva destes. Foram assim excluídos os resultados, e consequentemente os casos, que considerámos dúbios e só então definimos a amostra final.

Neste subcapítulo apenas relatamos a sequência dos procedimentos gerais utilizados nesta etapa, sem alusão a dados obtidos, uma vez que existe um capítulo em que estão apresentados os resultados de forma pormenorizada. Pretendemos com isso evidenciar de forma mais clara qual a amostra definitiva em estudo e também encadear de uma forma mais lógica os resultados obtidos, com a posterior análise estatística.

A sequência usada na validação dos resultados foi a seguinte:

- 1- Análise da expressão do gene *housekeeping*, *B-actina*: com esta pretendemos determinar quais os casos, em que foi possível concluir de forma segura que todos os procedimentos laboratoriais, em cima descritos, foram eficazes. Assim, foi-nos possível identificar em que casos conseguimos sintetizar cDNA, partindo do respetivo mRNA, e amplificá-lo. O que nos garantiu que o mRNA do qual partimos estava íntegro, mesmo após ser sujeito aos processos de parafinação/desparafinação.
- 2- Análise das curvas de *melting*: com esta pretendemos excluir os valores resultantes de amplificações inespecíficas e garantir que avaliámos a presença ou ausência de expressão dos genes de interesse.

2.6. Cálculo de diferenças de expressão, em termos relativos

A expressão relativa dos genes de interesse foi realizada com base no método comparativo de C_T , também conhecido como método $2^{-\Delta\Delta CT}$ (22, 23).

Este não tem em consideração as eficiências de amplificação do gene de referência nem dos genes de interesse, mas preconiza que sejam idênticas. Nós assumimos essa semelhança, porque devido à baixa quantidade de mRNA extraído e consequentemente diminuto cDNA sintetizado, não nos foi possível realizar uma série de diluições. Desta forma não conseguimos determinar a reta de eficiência de amplificação.

Por isso optámos por este método em detrimento de métodos que consideram as eficiências de amplificação, como o método de *Pfaffl* (24).

O cálculo do C_T foi realizado a partir da curva traçada pelo *software* iQTM5 v 2.1 (Bio-Rad®, Munich, Germany). Estabeleceu-se um mesmo valor limiar de fluorescência para todos os genes, de forma a interseção o gráfico no início da fase exponencial da reação, tal como se pode verificar no gráfico 1.

Após determinar os C_T dos genes alvo do estudo e da β -actina, foi realizada a normalização de cada gene para o respetivo gene *housekeeping*. É explicada a seguir a sequência de equações que permite o cálculo de diferenças de expressão, em termos relativos, entre grupos:

$$\Delta C_T = C_T (\text{gene de interesse}) - C_T (\text{gene housekeeping})$$

Uma vez que o objetivo será a comparação da expressão relativa entre dois grupos considerados, realizámos o cálculo da média (\bar{x}) de ΔC_T de cada gene, para cada grupo de interesse e de controlo. O desvio padrão também foi calculado. A razão de se ter optado pela análise de expressão comparativa entre grupos baseou-se fundamentalmente na inexistência de casos e controlos pareados.

Para o cálculo de $\Delta \Delta C_T$ foi usada a seguinte fórmula:

$$\Delta \Delta C_T = \bar{x} \Delta C_T (\text{grupo de interesse}) - \bar{x} \Delta C_T (\text{grupo usado como controlo})$$

A título de exemplo:

$$\Delta \Delta C_T = \bar{x} \Delta C_T (\text{grupo com metastização ganglionar}) - \bar{x} \Delta C_T (\text{grupo sem metastização ganglionar})$$

Por último foi aplicada a fórmula da expressão relativa:

$$\text{Expressão relativa} = 2^{-\Delta \Delta C_T}$$

$$\text{Ou} = [2^{-(\bar{x} \Delta C_T (\text{grupo de interesse}))}] / [2^{-(\bar{x} \Delta C_T (\text{grupo usado como controlo}))}]$$

A título de exemplo:

$$= [2^{-(\bar{x} \Delta C_T (\text{grupo com metastização ganglionar}))}] / [2^{-(\bar{x} \Delta C_T (\text{grupo sem metastização ganglionar}))}]$$

O resultado obtido permite verificar o rácio de expressão entre os dois grupos, i.e., quantas vezes mais o gene considerado é expresso em mRNA no grupo de interesse, comparativamente ao grupo usado como controlo.

2.7. Descrição das variáveis e métodos estatísticos

Uma vez que o principal objetivo do estudo foi a avaliação de diferenças de expressão dos genes de interesse, em grupos de pacientes com prognóstico distinto, diagnosticadas com cancro da mama, devemos descrever as variáveis em estudo.

Os métodos utilizados permitiram a realização de dois tipos de análise. Uma delas consistiu na simples verificação da presença ou ausência de expressão dos genes de interesse, por mRNA. Já a outra permitiu determinar diferenças relativas na expressão de mRNA, i.e., um rácio de expressão entre dois grupos, quando foi confirmada a presença de expressão do gene de interesse.

Assim, no que concerne aos genes de interesse, considerámos dois tipos de variáveis: uma qualitativa, expressão do gene de interesse, que determina para cada caso a presença ou ausência de expressão em mRNA desse mesmo gene; e, uma numérica, ΔC_T do gene de interesse, valor proveniente de uma das equações/etapas utilizadas na fórmula de cálculo de diferenças de expressão, em termos relativos, entre grupos (subcapítulo 2.6).

Salienta-se que na variável numérica apenas são considerados os casos em que se detetou presença de expressão em mRNA do gene de interesse.

Ao nível do prognóstico considerámos as seguintes variáveis: “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros” para classificar os pacientes em grupos de diferentes prognóstico. Estas variáveis de interesse e respetivas categorias estão mencionadas, em baixo, na tabela 8.

Tabela 8 - Variáveis e suas categorias - consideradas na avaliação de diferenças de expressão em mRNA nos genes de interesse

Variável	Categorias da variável
Envolvimento do gânglio sentinela e/ou outros	Sem metastização vs Metastização
Estadio clínico	I e II vs III

Estando cientes da expectável baixa robustez dos dados, em consequência da diminuta amostragem, a influência do acaso poderá ser no nosso estudo mais proeminente do que em amostras maiores. Pode, assim, algum fator influenciar a análise e precipitar conclusões erradas. Desta forma, considerámos duas variáveis, que expectavelmente não se associam, nem tampouco se espera que apresentem diferenças de expressão entre as suas categorias com os genes de interesse. Pretendemos assim que as variáveis, “Mama afetada” e “Localização afetada”, apresentadas na tabela 9, “funcionem” como controlos de acaso e consequentemente auxiliem na validação da análise efetuada.

Tabela 9 - Variáveis e suas categorias - consideradas para controlo do acaso e auxílio da validação da análise

Variável	Categorias da variável
Mama afetada	Direita vs Esquerda
Localização afetada	Quadrante Supero-externo vs Outra

O estudo estatístico não se esgotou nas variáveis e respetivas categorias constantes nas tabelas 8 e 9. Foram também analisadas outras variáveis clínicas e histopatológicas, estando descritas, em baixo, na tabela 10.

Tabela 10 - Outras variáveis clínicas e histopatológicas e suas respetivas categorias - consideradas na avaliação de diferenças de expressão em mRNA nos genes de interesse

Variável	Categorias da variável
Idade	≤50 anos vs >50 anos
Tipo histológico	Ductal invasivo vs Lobular e Mucinoso invasivo
RE (Recetores de estrogénio)	Positivo vs Negativo
RP (Recetores de progesterona)	Positivo vs Negativo
C-erb-b2	Positivo vs Negativo
Tripló negativo	Sim vs Não
Subtipo molecular (25)	Luminal A vs Outros
Focalidade	Multifocal vs Unifocal
Tamanho	≤20mm vs >20mm
Grau de diferenciação ¹	Bem e moderado vs Pouco
Invasão linfo-vascular	Presente vs Não identificada
Envolvimento da pele e/ou músculo	Sim vs Não
Componente <i>in situ</i>	Presente vs Ausente
Calcificação tumoral	Presente vs Ausente

¹ -segundo classificação de *Scarff-Bloom-Richardson* (26)

2.7.1. Testes de independência - qui-quadrado de *Pearson*

Todas as variáveis mencionadas nas tabelas 8-10 são de índole qualitativa, assim como a variável expressão do gene de interesse. Desta forma, para testarmos a associação desta última com as restantes, i.e., se existe associação estatística entre variáveis qualitativas, recorreremos ao teste de qui-quadrado de *Pearson*.

Esta análise foi realizada recorrendo ao *software* informático *IBM® SPSS® Statistics* versão 21.0 e considerámos as seguintes hipóteses:

Hipótese nula, H_0 : As variáveis consideradas são independentes, i.e., não existe associação entre elas.

Hipótese alternativa, H_A : As variáveis consideradas estão relacionadas, i.e., existe associação entre elas.

Utilizámos um intervalo de confiança de 95%, ou seja, quando o *p-value* foi superior a 0,05 a hipótese nula não foi rejeitada e desta forma, considerámos não existir associação entre as variáveis.

Se *p-value* inferior a 0,05, a hipótese nula é rejeitada e assim considera-se existir associação estatisticamente significativa entre as variáveis. Caso esta proposição seja cumprida o cálculo de *odds ratio* será efetuado com recurso ao mesmo *software* informático supramencionado.

2.7.2. Testes não paramétricos

Como melhor explicitado no subcapítulo 2.6, “Cálculo de diferenças de expressão, em termos relativos”, uma das potencialidades do método RT-PCR usado, é a possibilidade de quantificar a expressão dos genes de interesse. Assim quando se observou a expressão dos genes de interesse, determinou-se o rácio de expressão relativa entre dois grupos distintos.

Atendendo ao método utilizado para este cálculo, quanto mais distinta for a média de ΔC_T do gene de interesse entre os grupos comparados (grupo interesse vs grupo controlo), maior tenderá a ser a diferença de expressão do gene de interesse entre esses grupos. Ou seja, previamente ao cálculo do rácio de expressão procedeu-se a uma análise da distribuição/média de ΔC_T , entre os grupos comparados. Com esta análise pretendemos prever em quais variáveis, as suas categorias poderiam apresentar diferenças significativas, no cálculo da expressão relativa dos polimorfismos de interesse. Foi também nosso objetivo identificar se outras variáveis poderiam exercer alguma influência no cálculo do rácio de expressão relativa das variáveis de interesse, devendo dessa forma ser tomadas em consideração. Para determinação dessa possível influência atentámos em variáveis, que não as de interesse, com média/distribuição de ΔC_T do gene de interesse muito díspar entre as suas categorias.

Nota: Não pretendemos atribuir significado fisiológico, a esta análise efetuada, aos valores da média/distribuição de ΔC_T dos genes de interesses entre os diferentes grupos. Apenas foi usado como um complemento à análise do rácio de expressão relativa, sendo sim os resultados desta última interpretados com significado fisiológico.

Sendo a variável ΔC_T do gene de interesse, de carácter numérico recorreu-se ao uso do teste não paramétrico, *Mann-Whitney U*, para verificar a distribuição/média de ΔC_T entre as

diversas categorias de cada variável. A utilização deste em detrimento de testes paramétricos teve como pilares a reduzida amostragem e o facto de esta não ser normalmente distribuída.

Considerámos as seguintes hipóteses:

Hipótese nula, H_0 : A distribuição de ΔC_T do gene de interesse é a mesma entre as categorias de uma variável.

Hipótese alternativa, H_A : A distribuição de ΔC_T do gene de interesse é diferente entre as categorias de uma variável.

Como veremos no subcapítulo 3.4.1, de certo modo, o que pretendemos testar na hipótese nula é se a média de ΔC_T de cada um dos genes de interesse é a mesma entre dois grupos.

À semelhança do que foi usado para o teste de qui-quadrado, foi definido um intervalo de confiança de 95%, i.e., quando o *p-value* foi superior a 0,05 a hipótese nula não foi rejeitada. Esta não rejeição leva a considerar que a distribuição de ΔC_T é semelhante entre as categorias de uma variável, sendo a média entre as duas categorias/grupos idêntica.

Para facilitar visualmente a observação da distribuição de ΔC_T entre as categorias consideradas em cada variável, foram traçados *boxplot*.

A análise por testes não paramétricos também foi efetuada com recurso ao *software* informático IBM® SPSS® Statistics versão 21.0.

3. Resultados

3.1. Validação dos resultados obtidos em âmbito laboratorial

3.1.1. Resultados obtidos para a expressão do gene *housekeeping*, β -actina

Numa primeira abordagem aos resultados obtidos em âmbito laboratorial, debruçamo-nos sobre a análise dos resultados da expressão do gene *housekeeping*, β -actina. Só após esta abordagem foi possível validar os casos, em que seria efetuada uma análise pormenorizada de *CYP19A1* Trp39Arg, *GSTT1* e *GSTM1*. Para tal, foram elaborados critérios para selecionar os casos, que nos permitiram garantir menor probabilidade de erro na análise dos genes de interesse. Na tabela 11, mostrada a seguir, são apresentados os resultados totais obtidos da realização do RT-PCR para o gene *housekeeping*.

Tabela 11 - Casos processados pelas várias etapas descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR para gene *housekeeping*, β -actina

Caso	C _T de NTC	C _T do gene <i>housekeeping</i>	Temperatura de Melting de NTC (°C)	Temperatura de Melting do gene <i>housekeeping</i> (°C)
1	N/A	44,15	N/A	82
2	N/A	49,93	N/A	77
3	N/A	47,9	N/A	84,5
4	N/A	35,69	N/A	82,5
5	N/A	N/A	N/A	N/A
6	N/A	N/A	N/A	N/A
7	N/A	40,68	N/A	84
8	N/A	N/A	N/A	N/A
9	N/A	38,09	N/A	83,5
10	N/A	38,12	N/A	83
11	N/A	44,58	N/A	84,5
12	N/A	39,66	N/A	83,5
13	N/A	39,36	N/A	82,5
14	N/A	43,56	N/A	82
15	N/A	40,69	N/A	83
16	N/A	36,73	N/A	83,5
17	N/A	36,3	N/A	82
18	N/A	35,55	N/A	83
19	N/A	42,31	N/A	83,5
20	N/A	56,93	N/A	73
21	N/A	N/A	N/A	N/A
22	N/A	35,34	N/A	84,5
23	N/A	37,37	N/A	83,5

24	N/A	37,07	N/A	84
25	N/A	42,23	N/A	81,5
26	N/A	43,61	N/A	82
27	N/A	42,28	N/A	81,5
28	N/A	38,26	N/A	83,5
29	N/A	47,07	N/A	85
30	N/A	39,34	N/A	83
31	N/A	38,97	N/A	84
32	N/A	N/A	N/A	N/A
33	N/A	40,93	N/A	83,5
34	N/A	41,38	N/A	83,5
35	N/A	37,1	N/A	84,5
36	N/A	41,19	N/A	83,5
37	N/A	38,45	N/A	83,5

São apresentados os valores do C_T e temperatura de *melting* (em °C) para o gene *housekeeping*, β -actina.

Também é apresentado C_T e temperatura de *melting* (em °C) para o respetivo controlo negativo (NTC). Quando não existiu deteção de fluorescência ou não foi definido um valor, usou-se a nomenclatura N/A.

Os critérios de exclusão usados fundamentaram-se essencialmente na comprovação da existência de cDNA para cada caso. Ou seja, se numa determinada amostra colocada no termociclador de RT-PCR, não foi possível determinar a presença do gene *housekeeping*, então não nos é possível concluir que exista cDNA. Deste modo não podemos concluir que as técnicas usadas foram eficazes para esse caso, devendo ser excluído da análise subsequente.

É de realçar que este tipo de validação dos resultados, baseada na presença de cDNA, apenas foi possível porque os *primers* de β -actina utilizados são altamente específicos, para a sequência alvo em cDNA em detrimento da sequência total de DNA genómico. Esta elevada especificidade advém das suas localizações em regiões fronteira de exões, estando o *forward primer* entre o 3º e 4º exões codificantes e o *reverse primer* entre o 4º e 5º exões codificantes. No anexo 2 descreve-se a localização de *primers* na sequência de cDNA.

Assim a leitura de fluorescência emitida é um valor que nos garante ser resultante de cDNA por nós sintetizado e não de DNA genómico.

Aplicámos desta forma os seguintes critérios:

- Quando não foi detetada qualquer fluorescência na amostra colocada no termociclador de RT-PCR (N/A na coluna C_T do gene *housekeeping* da tabela 11), essas amostras foram excluídas. Esta decisão baseia-se por não existirem dados que suportem a evidência de ter ocorrido a amplificação de cDNA do gene *housekeeping*, apontando assim a

favor da não existência de cDNA desse caso. Desta forma foram excluídos do estudo 5 casos (5; 6; 8; 21 e 32).

- Quando apesar de ter ocorrido deteção de fluorescência pelo procedimento descrito no RT-PCR, a temperatura de *melting* foi claramente discrepante à mediana de todos os casos, não nos foi possível concluir que a deteção de fluorescência correspondeu ao fragmento de interesse do gene *housekeeping*. Foram assim excluídos 2 casos (2 e 20), tal como é possível observar no gráfico 2, em baixo, por apresentarem um valor de temperatura de *melting* muito distante à mediana (6,5°C e 10,5°C respetivamente).

Diferença entre mediana de temperatura de *melting* e temperatura de *melting* de cada caso da β -actina

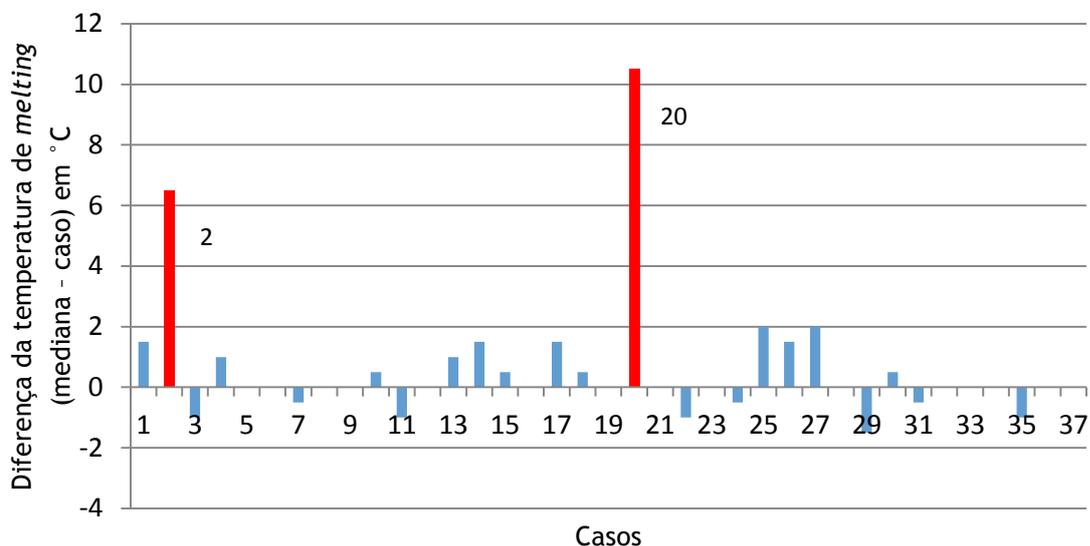


Gráfico 2 - Diferença entre mediana de temperatura de *melting* e temperatura de *melting* de cada caso da β -actina

Os casos 2 e 20, com barras a vermelho, têm valores claramente discrepantes relativamente à mediana. Quando o valor da temperatura de *melting* é de igual valor à mediana e nos casos 5; 6; 8; 21 e 32, que não apresentaram deteção de fluorescência, não foi traçada qualquer barra.

Consequentemente foram excluídos 7 casos (2; 5; 6; 8; 20; 21 e 32) por não ter sido possível validá-los, tendo os restantes 30 permanecido em estudo, por se concluir que nestes houve sucesso na extração de mRNA e respetiva síntese de cDNA. Será, assim, mais fidedigna a análise dos resultados obtidos para os genes de interesse, *CYP19A1* Trp39Arg (T/C), *GSTT1* e *GSTM1*.

Nos gráficos 3 e 4 e tabela 12, que se seguem, pretende-se ilustrar, que após a aplicação dos critérios de validação da amostra baseados na análise do gene *housekeeping*, a amostra ficou distribuída de forma mais uniforme.

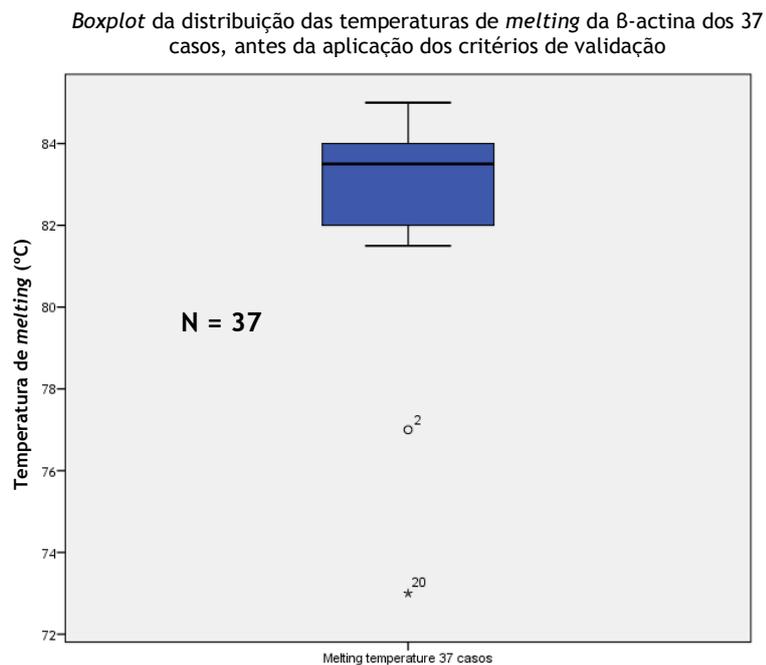


Gráfico 3 - *Boxplot* da distribuição das temperaturas de *melting* da β -actina dos 37 casos, antes da aplicação dos critérios de validação

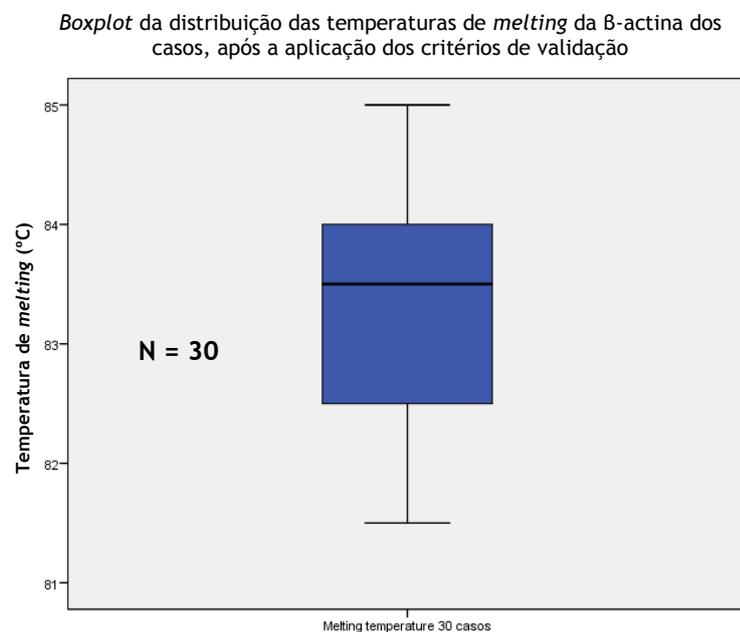


Gráfico 4 - *Boxplot* da distribuição das temperaturas de *melting* da β -actina dos 30 casos, após a aplicação dos critérios de validação

Tabela 12 - Descrição dos valores de tendência central da temperatura de *melting* da β -actina, antes e após a aplicação dos critérios de validação, para obtenção da amostragem definitiva

	Antes (37 casos)	Após (30 casos)
Mediana	83,5 °C	83,5 °C
Média	82,75 °C	83,27 °C
Desvio-padrão	2,282	0,935

3.1.2. Resultados obtidos para a expressão da variante polimórfica da aromatase, *CYP19A1* Trp39Arg (T/C)

Neste subcapítulo será elucidada a análise dos dados para o polimorfismo da aromatase, *CYP19A1* Trp39Arg, sendo explicitado como foi determinada a presença ou ausência de expressão em mRNA desta variante polimórfica. Na tabela 13 são expostos os resultados obtidos, mediante a realização de RT-PCR.

Tabela 13 - Casos validados (n=30) processados pelas várias etapas laboratoriais descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR, para *CYP19A1* Trp39Arg (T/C)

Caso	C _T de NTC	C _T de <i>CYP19A1</i> Trp39Arg	Temperatura de <i>Melting</i> de NTC (°C)	Temperatura de <i>Melting</i> de <i>CYP19A1</i> Trp39Arg (°C)	Expressão	ΔC_T
1	31,32	30,15	71	69,5 e 78,5	Sim	-14
3	30,79	38,7	72	69	Não	--
4	30,79	30,69	72	69,5	Não	--
7	27,91	N/A	72	N/A	Não	--
9	27,91	30,68	72	70	Não	--
10	32,72	32,88	72	70	Não	--
11	27,91	29,67	72	71	Não	--
12	27,91	28,31	72	71,5	Não	--
13	31,32	29,78	71	69,5 e 77	Sim	-9,58
14	31,32	30,03	71	70	Não	--
15	31,32	29,41	71	69,5 e 79	Sim	11,28
16	31,32	29,69	71	70 e 74	Sim	-7,04
17	31,32	28,69	71	70,50 e 75,5	Sim	-7,61
18	33,79	34,46	71,5	68,5	Não	--
19	33,79	31,51	71,5	69,5 e 75	Sim	-10,8
22	33,79	29,83	71,5	73 e 81,5	Sim	-5,51
23	33,79	29,85	71,5	70,5 e 76	Sim	-7,52
24	33,79	29,9	71,5	71 e 81,5	Sim	-7,17
25	32,18	35,06	71	68,5	Não	--

26	32,18	37,26	71	68,5	Não	--
27	32,18	31,23	71	69,5	Não	--
28	32,18	30,04	71	70	Não	--
29	32,18	30,93	71	71	Não	--
30	32,72	31,75	72	70,5	Não	--
31	32,72	30,48	72	71,5	Não	--
33	34,61	30,41	70,5	69,5	Não	--
34	34,61	34,19	70,5	69,5	Não	--
35	34,61	31,48	70,5	70,5	Não	--
36	34,61	29,86	70,5	70,5 e 79,5	Sim	11,33
37	34,61	29,73	70,5	71 e 79	Sim	-8,72

São apresentados os valores do C_T e temperatura de *melting* (em °C) para esta variante. Também é exposto C_T e temperatura de *melting* (em °C) para o respetivo controlo negativo (NTC). Quando não existiu deteção de fluorescência ou não foi definido um valor, usou-se a nomenclatura N/A. A coluna expressão define quando esta ocorreu, sob a forma de mRNA, para a variante da aromatase. ΔC_T , representa a normalização da expressão de *CYP19A1 Trp39Arg* relativamente ao gene *housekeeping*, (C_T (*CYP19A1 Trp39Arg*) - C_T (*B-ACTINA*)).

Tal como se pode verificar, em todos os casos, com a exceção de um (caso 7), houve leitura de fluorescência. Neste caso, à semelhança do que foi realizado na análise da *B-actina*, considerou-se não existir amplificação do fragmento desejado e deste modo a consequente ausência de expressão.

Após um atento olhar para as temperaturas de *melting* de todas as amostras, verificamos uma dicotomia, de semelhanças e diferenças dos casos com os controlos (NTC). Em todos os controlos houve deteção de fluorescência e determinação de valor de temperatura de *melting*, enquanto que, em alguns casos determinaram-se dois valores de temperatura de *melting*.

Como podemos observar em todos os controlos negativos (NTC), a temperatura de *melting* foi muito semelhante, com mediana de 71°C, média de 71,3°C e desvio padrão de 0,55. Dada a similitude apresentada e tendo sido todos colocados nas mesmas condições, i.e., na ausência de adição voluntária de qualquer ácido nucleico, colocamos a hipótese de estas leituras serem em consequência da formação de dímeros de *primers*.

Desta forma considerámos que valores de temperatura de *melting* entre +/- 3 desvios-padrão da média de NTC, poderiam estar dentro de intervalos próximos o suficiente dos controlos negativos (NTC) para serem tidos como uma amplificação distinta (27). Como a sequência de interesse é maior que a dos dímeros, vamos considerar que valores até + 3 desvios padrão da média das temperaturas de *melting* de NTC, i.e., até 72,95°C, serão considerados como não estando a ser detetada a sequência de interesse. Assim nos casos em que a temperatura de *melting* foi $\leq 72,95^\circ\text{C}$, considera-se como não tendo ocorrido amplificação da sequência de cDNA de interesse e desta forma que não expressam o gene em mRNA. Esta importante

conclusão assenta no facto de já ter sido provado pela análise do gene *housekeeping*, que em todos estes 30 casos, há cDNA na amostra.

Em 11 casos houve determinação de dois valores distintos de temperatura de *melting*. Foi assim necessário recorrer à análise por curva de *melting*, de modo a melhor deslindar os resultados obtidos. Quando a temperatura de *melting* foi superior a 72,95°C, considerou-se que ocorreu amplificação de uma sequência distinta da detetada nos controlos negativos (NTC), podendo com grande probabilidade ser a sequência de interesse. Deste modo considerámos como presente a expressão da *CYP19A1* Trp39Arg em mRNA. Salienta-se que nestes casos, também houve deteção de uma temperatura de *melting*, que se situa até 72,95°C, sugerindo que nestas reações também pode ter ocorrido formação de dímeros de *primers*. No entanto a presença de uma outra temperatura distinta, >72,95°C, sugere que ocorreu amplificação do fragmento de interesse. O gráfico 5 pretende evidenciar isso mesmo, mostrando-se que nestes casos, apesar de ocorrer deteção de uma temperatura de *melting* semelhante a NTC, há uma outra temperatura de *melting* claramente distinta.

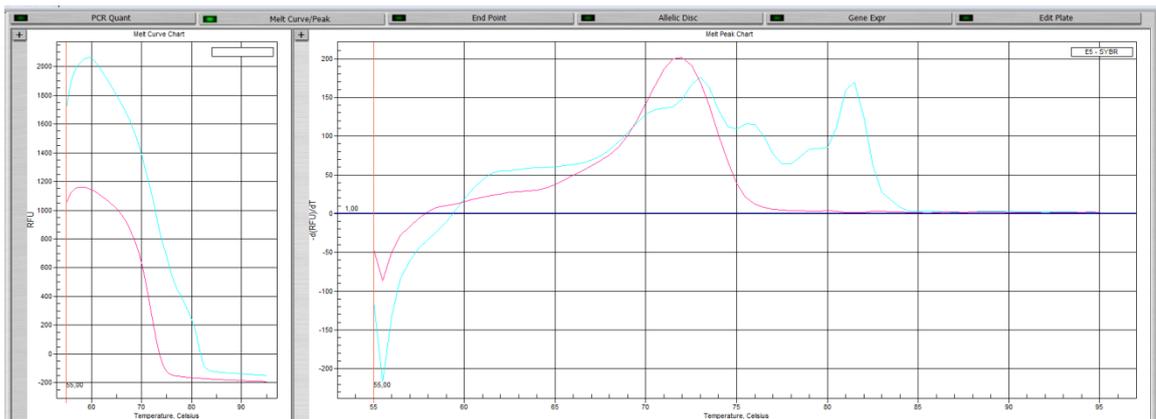


Gráfico 5 - Representação das curvas de *melting* de um controlo -NTC (rosa) e um caso (azul) de *CYP19A1* Trp39Arg

Apesar de uma temperatura de *melting* ser semelhante ao controlo, a outra é claramente distinta, fazendo-nos crer que corresponde ao fragmento de interesse.

No eixo das abcissas está a temperatura (°C), enquanto no eixo das ordenadas o valor de fluorescência.

Desta análise realizada, verifica-se a deteção da expressão de mRNA da variante polimórfica da aromatase, *CYP19A1* Trp39Arg, em 11 dos 30 casos considerados como válidos. Apenas estes 11 serão tidos em conta para o cálculo do rácio de expressão relativa entre grupos.

3.1.3. Resultados obtidos para a expressão dos restantes genes (*GSTM1* e *GSTT1*)

A expressão de *GSTM1* e *GSTT1* foi analisada em concordância com os critérios e métodos usados no gene *housekeeping* e *CYP19A1* Trp39Arg. Portanto, aqui, será com maior brevidade que se elucidará acerca dos seus resultados, uma vez que as bases de análise dos dados de RT-PCR já foram explicitadas.

Na *GSTM1*, como podemos observar na tabela 14 que se segue, muitos dos casos não obtiveram leitura de fluorescência. Uma vez que, já foi provada a existência de cDNA nessas amostras, estes resultados favorecem a ausência de expressão de *GSTM1* para esses casos.

Já nos casos em que ocorreu leitura de fluorescência, constatámos que a temperatura de *melting* é muito similar aos controlos (NTC), em que também foi detetada leitura. Deste modo, não existe evidência suficiente que prove que estes casos tenham uma sequência maior e diferente, que em NTC. Assim os resultados apontam para uma ausência de expressão de *GSTM1* em todos os casos.

Tabela 14 - Casos validados (n=30) processados pelas várias etapas laboratoriais descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR, para *GSTM1*

Caso	C _T de NTC	C _T de <i>GSTM1</i>	Temperatura de <i>Melting</i> de NTC (°C)	Temperatura de <i>Melting</i> de <i>GSTM1</i> (°C)
1	N/A	41,11	N/A	71,5
3	37,77	47,12	73,5	72
4	37,77	37,1	73,5	72,5
7	N/A	48,83	N/A	71
9	N/A	38,98	N/A	73
10	41,41	N/A	73	N/A
11	N/A	41,68	N/A	73,5
12	N/A	41,79	N/A	73,5
13	N/A	N/A	N/A	N/A
14	N/A	N/A	N/A	N/A
15	N/A	N/A	N/A	N/A
16	N/A	N/A	N/A	N/A
17	N/A	N/A	N/A	N/A
18	N/A	N/A	N/A	N/A
19	N/A	N/A	N/A	N/A
22	N/A	N/A	N/A	N/A
23	N/A	N/A	N/A	N/A
24	N/A	37,84	N/A	73,5
25	N/A	N/A	N/A	68,5
26	N/A	N/A	N/A	N/A
27	N/A	N/A	N/A	N/A

28	N/A	N/A	N/A	N/A
29	N/A	N/A	N/A	N/A
30	41,41	34,83	73	73,5
31	37,77	35,81	73,5	73,5
33	N/A	N/A	N/A	N/A
34	N/A	N/A	N/A	N/A
35	N/A	52,97	N/A	72
36	N/A	N/A	N/A	N/A
37	N/A	N/A	N/A	N/A

São apresentados os valores do C_T e temperatura de *melting* (em °C) para este gene. Também é exposto C_T e temperatura de *melting* (em °C) para o respetivo controlo negativo (NTC). Quando não existiu deteção de fluorescência, ou não foi definido um valor, usou-se a nomenclatura N/A.

Para a *GSTT1* o que se verifica é algo semelhante, tal como é mostrado na tabela 15. Porém, em todos os casos, sem qualquer exceção, ocorreu leitura de fluorescência, assim como na maioria dos controlos (NTC). Tendo de novo em consideração a temperatura de *melting*, também constatámos que nos casos, as temperaturas são muito semelhantes aos controlos (NTC). Assim, uma vez mais, não existe qualquer evidência de ter ocorrido amplificação de um fragmento distinto de NTC, i.e., amplificação do fragmento do gene. Assim os resultados também apontam para uma ausência de expressão de *GSTT1* em todos os casos.

Tabela 15 - Casos validados (n=30) processados pelas várias etapas laboratoriais descritas em materiais e métodos desde desparafinação a síntese de cDNA, com respetiva amplificação em RT-PCR, para *GSTT1*

Caso	C_T de NTC	C_T de <i>GSTT1</i>	Temperatura de <i>Melting</i> de NTC (°C)	Temperatura de <i>Melting</i> de <i>GSTT1</i> (°C)
1	32,73	31,75	75	74,5
3	34,72	33,68	75,5	74,5
4	34,72	33,92	75,5	75
7	N/A	35,63	N/A	73
9	N/A	34,12	N/A	74,5
10	36,12	32,26	75,5	73
11	N/A	34,45	N/A	75,5
12	N/A	31,9	N/A	75,5
13	32,73	31	75	74,5
14	32,73	31,62	75	74,5
15	32,73	31,38	75	74,5
16	32,73	31,53	75	74,5
17	32,73	30,99	75	75
18	32,28	33,53	75,5	73
19	32,28	31,9	75,5	74
22	32,28	31,18	75,5	75,5

23	32,28	30,81	75,5	76
24	32,28	30,47	75,5	76
25	N/A	34,85	N/A	73
26	N/A	32,07	N/A	73,5
27	N/A	30,85	N/A	73,5
28	N/A	30,82	N/A	74,5
29	N/A	31,62	N/A	76
30	36,12	34,88	75,5	74,5
31	34,72	28,47	75,5	75,5
33	N/A	33,26	N/A	73,5
34	N/A	34,16	N/A	73,5
35	N/A	33,07	N/A	74,5
36	N/A	31,17	N/A	76
37	N/A	31,22	N/A	76

São apresentados os valores do C_T e temperatura de *melting* (em °C) para este gene. Também é exposto C_T e temperatura de *melting* (em °C) para o respetivo controlo negativo (NTC). Quando não existiu deteção de fluorescência, ou não foi definido um valor, usou-se a nomenclatura N/A.

Atendendo aos resultados obtidos, em nenhum caso existe evidência que sustente a presença de expressão destes dois genes. Deste modo não seria sensato prosseguir com o seu estudo estatístico e consequente avaliação de diferenças de expressão entre grupos distintos, uma vez que em todos os casos parece ocorrer ausência de expressão.

Os intrigantes resultados laboratoriais obtidos para estes genes serão alvo de uma análise mais abrangente no capítulo da Discussão/Conclusão. Nos próximos subcapítulos serão apresentadas a análise descritiva e estatística da amostragem definitiva, assim como a determinação do rácio de expressão relativa. Porém em nenhuma delas serão considerados *GSTT1* e *GSTM1*.

3.2. Análise descritiva da amostra

A amostragem final é neste subcapítulo objeto de análise descritiva, mediante as tabelas 16 a 19, gráfico 6 e ao longo do texto. Também serão mencionadas algumas particularidades da amostra. Todos os dados recolhidos no SAM e obtidos em meio laboratorial, para os 30 casos que compõem a amostra definitiva, estão no anexo 3.

A amostra em estudo, após validação, é assim constituída por 30 mulheres, diagnosticadas histologicamente com carcinoma da mama às idades compreendidas entre os 34 e os 86 anos. A média de idades ao diagnóstico foi de 64,83 anos e σ (desvio padrão) de 14,75.

Tabela 16 - Características clínicas da amostra definitiva em estudo

Parâmetro clínico		Freq. Abs.	Percentagem
Idade ao diagnóstico (em anos)	≤50	6	20%
	>50	24	80%
Mama afetada	Direita	15	50%
	Esquerda	15	50%
Localização afetada	Quadrante Supero-externo	10	33,3%
	Outra localização	20	66,7%

Freq. Abs. - Frequência Absoluta

A variável “Localização afetada” foi dividida em 2 categorias, “Quadrante Supero-externo” e “Outra”, uma vez que a primeira corresponde ao local mais frequentemente afetado.

Tabela 17 - Características histopatológicas do tumor e gânglio sentinela e/ou outros gânglios da amostra definitiva em estudo

Parâmetro histopatológico		Freq. Abs.	Percentagem
Tipo Histológico	Carcinoma Ductal Invasivo	26	86,7%
	Carcinoma Lobular Invasivo	3	10%
	Carcinoma Mucinoso Invasivo	1	3,3%
Componente <i>in situ</i>	Sim	16	53,3%
	Não	14	46,7%
Verificada calcificação	Sim	5	16,7%
	Não	25	83,3%
Focalidade	Unifocal	27	90%
	Multifocal	3	10%
Tamanho do tumor ¹ (em mm)	≤ 20	16	53,3%
	>20	14	46,7%
Envolvimento da pele e/ou músculo	Sim	5	16,7%
	Não	25	83,3%
Grau de diferenciação (segundo classificação de <i>Scarff-Bloom-Richardson</i>) (26)	Bem diferenciado	10	33,3%
	Moderadamente diferenciado	16	53,3%
	Pouco diferenciado	4	13,4%
Invasão linfo-vascular	Sim	2	6,7%
	Não observado	28	93,3%
Envolvimento de gânglio sentinela e/ou outros	Com metastização	9	30%
	Sem metastização	21	70%

¹ - Quando multifocal, o de maior tamanho
Freq. Abs. - Frequência Absoluta

A média de tamanhos dos tumores excisados, ou do maior quando multifocal, foi de 24,48 mm, com σ (desvio-padrão) de 18,47. O mais pequeno de 1,5mm e o maior 90mm.

Como já foi sendo explicitado, quando nos referimos ao “Envolvimento do gânglio sentinela e/ou outros”, estamos a considerar metastização regional.

No caso 13, a Anatomia Patológica não conseguiu determinar se houve ou não envolvimento do gânglio sentinela, por metastização. Porém esta foi considerada como ausente, uma vez que após 4 anos e meio de seguimento em consultas regulares de Ginecologia Oncológica não foi detetada qualquer recidiva local, sinais de envolvimento ganglionar ou até mesmo metástase à distância.

Em dois dos casos, 1 e 19, existiu invasão e disrupção da cápsula dos gânglios linfáticos afetados.

Tabela 18 - Recetores tumorais da amostra definitiva em estudo

Recetores tumorais		Freq. Abs.	Percentagem
Recetores de estrogénio	Positivo	23	76,7%
	Negativo	7	23,3%
Recetores de progesterona	Positivo	18	60%
	Negativo	12	40%
C-erb-B2	Positivo	4	13,3%
	Negativo	26	86,7%
Triplo Negativo	Sim	3	10%
	Não	27	90%
Subtipo molecular (25)	Luminal A	23	76,7%
	Luminal B	1	3,3%
	HER2+ (ER-)	3	10%
	<i>Basal-like</i>	3	10%

Freq. Abs. - Frequência Absoluta.

Tabela 19 - Estadio clínico e avaliação TNM (28, 29) da amostra definitiva em estudo

Avaliação TNM e Estadio Clínico		Freq. Abs.	Percentagem
Tumor	T1	15	50%
	T2	9	30%
	T3	1	3,3%
	T4	5	16,7%
Gânglios linfáticos Regionais	N0	21	70%
	N1	7	23,4%
	N2	1	3,3%
	N3	1	3,3%

Estadio Clínico	I	12	40%
	IIA	9	30%
	IIB	1	3,3%
	IIIA	2	6,7%
	IIIB	5	16,7%
	IIIC	1	3,3%

Freq. Abs. - Frequência Absoluta

Como já foi referido em materiais e métodos não foram incluídas no estudo, mulheres com suspeita ou confirmação de metástases à distância, devendo ser sempre considerado M0.

Expressão da variante da aromatase *CYP19A1* Trp39Arg

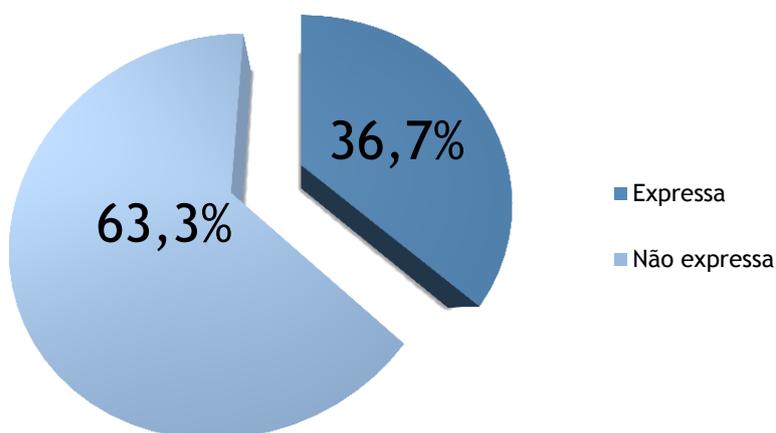


Gráfico 6 - Gráfico circular ilustrando a expressão da variante polimórfica da aromatase, *CYP19A1* Trp39Arg, na amostragem definitiva em estudo

A expressão da variante polimórfica da aromatase, *CYP19A1* Trp39Arg, foi detetada em 11 dos 30 casos que compõem a amostra definitiva estudada, correspondendo a cerca de 36,7% desta amostragem.

3.3. Estudo da associação entre a expressão de *CYP19A1* Trp39Arg (T/C) e variáveis qualitativas clínicas e histopatológicas

Aqui foi testado se a variável “Expressão de *CYP19A1* Trp39Arg” se associa de forma estatisticamente significativa com as variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”. Como podemos observar pela tabela 20, tal associação não se verificou, pois o *p-value* foi superior a 0,05, não sendo rejeitada a hipótese nula. As tabelas 21 e 22 pretendem mostrar a distribuição de frequências absolutas destas variáveis testadas.

Tabela 20 - Apresentação do nível de significância estatística (*p-value*) do teste de qui-quadrado de *Pearson*, que testa a associação da “Expressão de *CYP19A1* Trp39Arg” com as variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”

Variável	<i>p-value</i>
Envolvimento do gânglio sentinela e/ou outros	0,563
Estadio clínico	0,954

Tabela 21 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de *CYP19A1* Trp39Arg” e “Envolvimento do gânglio sentinela e/ou outros”

		Expressão de <i>CYP19A1</i> Trp39Arg (T/C)		Total
		Sim	Não	
Envolvimento do gânglio sentinela e/ou outros	Metastização	4	5	9
	Sem Metastização	7	14	21
Total		11	19	30

Tabela 22 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de *CYP19A1* Trp39Arg” e “Estadio clínico”

		Expressão de <i>CYP19A1</i> Trp39Arg (T/C)		Total
		Sim	Não	
Estadio clínico	III	3	5	8
	I e II	8	14	22
Total		11	19	30

Também nas “outras variáveis clínicas e histopatológicas”, constantes na tabela 10 do subcapítulo 2.7 de materiais e métodos, não foi encontrada qualquer associação estatisticamente significativa com expressão da variante polimórfica da aromatase - *CYP19A1* Trp39Arg.

O cálculo de *odds ratio* não foi assim efetuado, uma vez não ter ocorrido significância estatística em nenhum teste de associação entre as diversas variáveis qualitativas e “Expressão de *CYP19A1* Trp39Arg”.

Também nos testes realizados para as variáveis “controlo do acaso”, “Mama afetada” e “Localização afetada”, o *p-value* foi superior a 0,05, como se pode verificar na tabela 23. De igual forma, as tabelas 24 e 25, pretendem mostrar a distribuição de frequências absolutas destas duas variáveis testadas.

Tabela 23 - Apresentação do nível de significância estatística (*p-value*) do teste de qui-quadrado de *Pearson*, que testa a associação da “Expressão de *CYP19A1* Trp39Arg” com as variáveis “Mama afetada” e “Localização afetada”

Variável	<i>p-value</i>
Mama afetada	0,705
Localização afetada	0,789

Tabela 24 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de *CYP19A1* Trp39Arg” e “Mama afetada”

		Expressão de <i>CYP19A1</i> Trp39Arg (T/C)		Total
		Sim	Não	
Mama afetada	Direita	5	10	15
	Esquerda	6	9	15
Total		11	19	30

Tabela 25 - Tabela de contingência que mostra a distribuição de frequências absolutas das variáveis “Expressão de *CYP19A1* Trp39Arg” e “Localização afetada”

		Expressão de <i>CYP19A1</i> Trp39Arg (T/C)		Total
		Sim	Não	
Localização afetada	Outra	7	13	20
	Quadrante Supero-externo	4	6	10
Total		11	19	30

3.4. Diferenças de expressão da variante polimórfica, *CYP19A1* Trp39Arg (T/C), entre grupos

3.4.1. Estudo de média/distribuição de ΔC_T de *CYP19A1* Trp39Arg (T/C) entre as categorias das variáveis

Tal como explicitámos em 2.7.2, o presente subcapítulo, “Estudo de média/distribuição de ΔC_T de *CYP19A1* Trp39Arg (T/C) entre as categorias das variáveis”, pretende auxiliar na validação e interpretação dos resultados obtidos no posterior cálculo do rácio de expressão, em termos relativos, entre as categorias das variáveis.

Nesta análise verificou-se a distribuição de ΔC_T de *CYP19A1* Trp39Arg (T/C) pelas categorias das variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”. Como podemos verificar pelas figuras 3 e 4, em ambas as variáveis, o teste não paramétrico obteve um $p\text{-value} > 0,05$. Este dado levaria a não rejeitar a hipótese nula e assim dizer que a distribuição de ΔC_T para a variante polimórfica da aromatase seria semelhante em ambas as categorias de cada variável analisada, i.e., com uma média semelhante. Porém a realização dos respetivos *boxplot*, gráficos 7 e 8, que ilustram a distribuição de ΔC_T entre cada categoria das duas variáveis, permite constatar visualmente uma tendência para uma desigualdade nesta distribuição. Pode assim ser de prever que no posterior cálculo do rácio de expressão, possa existir uma diferença apreciável entre as categorias de cada variável aqui mencionada.

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔC_T <i>CYP19A1</i> Trp39Arg is the same across categories of Metastização vs Sem metastização.	Independent-Samples Mann-Whitney U Test	,412 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 3 - Teste *Mann-Whitney U*, em que a hipótese nula testada é que a média de ΔC_T de *CYP19A1* Trp39Arg é a mesma entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔC_T <i>CYP19A1</i> Trp39Arg is the same across categories of Estadio Clínico (I e II vs III).	Independent-Samples Mann-Whitney U Test	,194 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 4 - Teste *Mann-Whitney U*, em que a hipótese nula testada é que a média de ΔC_T de *CYP19A1* Trp39Arg é a mesma entre as categorias, “I e II” e “III”, da variável “Estadio clínico”

Boxplot com distribuição de ΔCT de CYP19A1 Trp39Arg entre as categorias da variável Envolvimento do gânglio sentinela e/ou outros

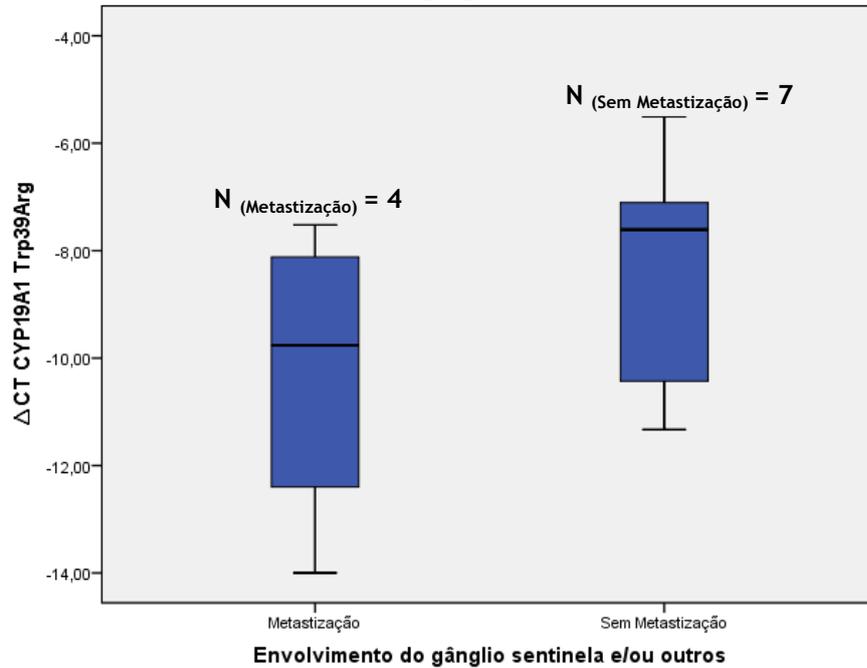


Gráfico 7 - *Boxplot* em que é ilustrada a distribuição de ΔCT de CYP19A1 Trp39Arg entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”. A frequência absoluta em cada categoria é apresentada no gráfico.

Boxplot com distribuição de ΔCT de CYP19A1 Trp39Arg entre as categorias da variável Estadio clínico

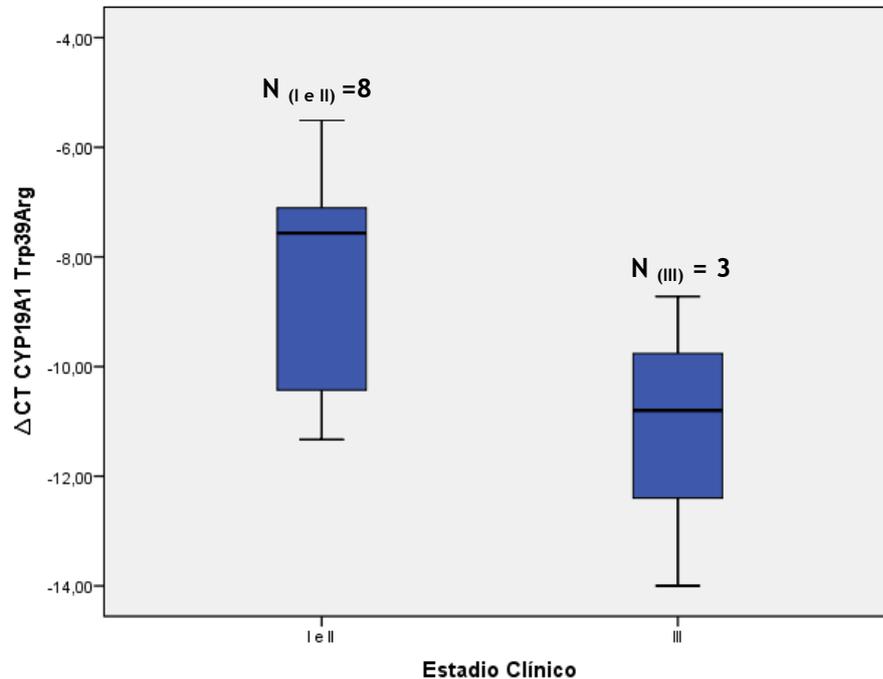


Gráfico 8 - *Boxplot* em que é ilustrada a distribuição de ΔCT de CYP19A1 Trp39Arg entre as categorias, “I e II” e “III”, da variável “Estadio clínico”. A frequência absoluta em cada categoria é apresentada no gráfico.

As variáveis utilizadas para “controlo do acaso”, “Mama afetada” e “Localização afetada”, têm $p\text{-value} = 1$, figuras 5 e 6. Este valor é indicativo de que a distribuição de ΔC_T de *CYP19A1* Trp39Arg é muito similar entre as categorias das variáveis controlo, i.e., média semelhante. Esta tendência a uma distribuição semelhante de ΔC_T é confirmada pelos gráficos 9 e 10. Portanto será de prever que, quando for determinada o rácio de expressão entre as categorias “mama direita vs mama esquerda afetada”, as diferenças de expressão do polimorfismo da aromatase tenderão a ser inexistentes. O mesmo ocorrendo entre as categorias “quadrante supero-externo vs outra localização afetada”.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔC_T <i>CYP19A1</i> Trp39Arg is the same across categories of Direita vs Esquerda.	Independent-Samples Mann-Whitney U Test	1,000 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 5 - Teste *Mann-Whitney U*, em que a hipótese nula testada é que a média de ΔC_T de *CYP19A1* Trp39Arg é a mesma entre as categorias, “Direita” e “Esquerda”, da variável “Mama afetada”

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔC_T <i>CYP19A1</i> Trp39Arg is the same across categories of Quadrante Supero-externo vs Outra.	Independent-Samples Mann-Whitney U Test	1,000 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 6 - Teste *Mann-Whitney U*, em que a hipótese nula testada é que a média de ΔC_T de *CYP19A1* Trp39Arg é a mesma entre as categorias, “Quadrante Supero-externo” e “Outra”, da variável “Localização afetada”

Boxplot com distribuição de ΔCT de CYP19A1 Trp39Arg entre as categorias da variável Mama afetada

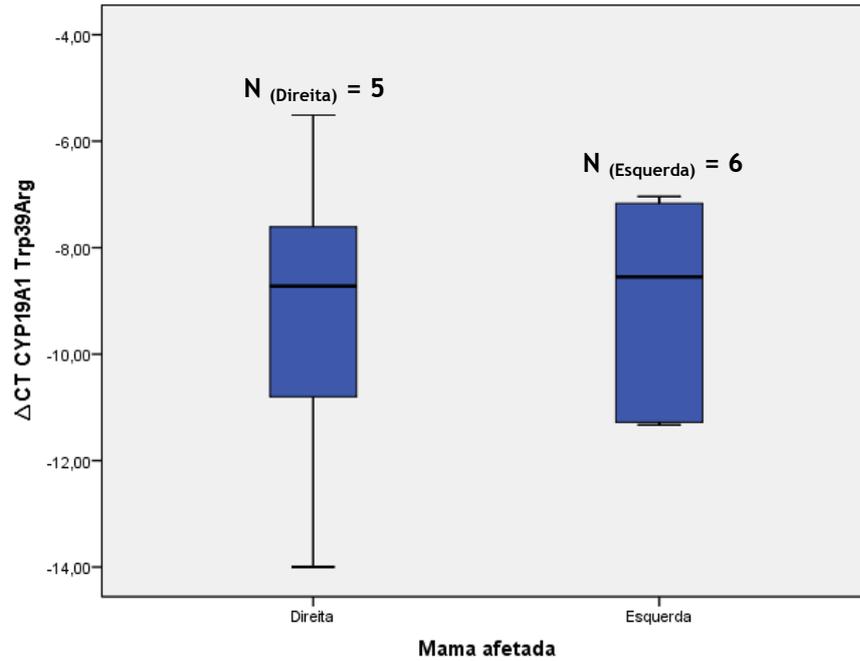


Gráfico 9 - Boxplot em que é ilustrada a distribuição de ΔC_T CYP19A1 Trp39Arg entre as categorias, “Direita” e “Esquerda”, da variável “Mama afetada”

A frequência absoluta em cada categoria é apresentada no gráfico.

Boxplot com distribuição de ΔCT de CYP19A1 Trp39Arg entre as categorias da variável Localização afetada

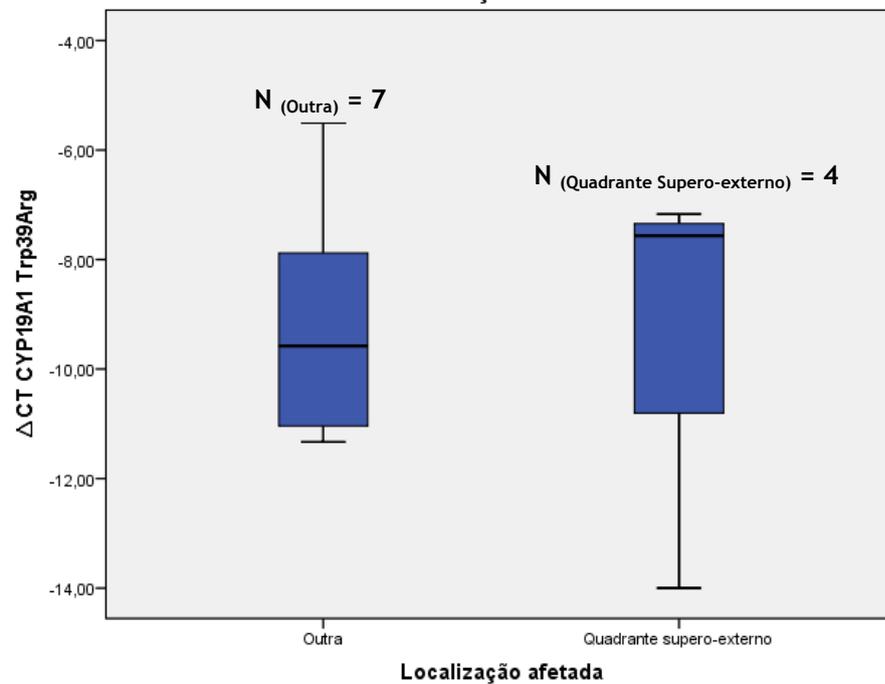


Gráfico 10 - Boxplot em que é ilustrada a distribuição de ΔC_T de CYP19A1 Trp39Arg entre as categorias, “Quadrante Supero-externo” e “Outra”, da variável “Localização afetada”

A frequência absoluta em cada categoria é apresentada no gráfico.

Como previsto em materiais e métodos, outras variáveis clínicas e histopatológicas foram também analisadas, sendo o principal objetivo verificar se alguma delas poderia exercer algum tipo de influência nas variáveis de interesse. Desta forma se tal ocorresse, essa variável seria tomada em consideração para o posterior cálculo do rácio de expressão de *CYP19A1* Trp39Arg entre as categorias das variáveis de interesse (“Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”).

Assim, o estudo da média/distribuição de ΔC_T do polimorfismo da aromatase nas categorias destas variáveis permitiu prever, recorrendo ao teste não paramétrico *Mann Whitney U* e visualmente pelos *boxplot*, se alguma variável pudesse apresentar uma tendência a influenciar as variáveis de interesse.

Deste modo a variável “RE (Recetores de estrogénio)” mereceu um olhar mais atento. Esta obteve um nível de significância estatística mais elevado, i.e., *p-value* mais baixo no teste de *Mann Whitney U* (figura 7), que as variáveis “Envolvimento do gânglio sentinela e/ou outros” e “Estadio clínico”. No respetivo *boxplot*, gráfico 11, constata-se uma clara distinção da distribuição de ΔC_T de *CYP19A1* Trp39Arg entre os grupos “recetores de estrogénio positivos” e “recetores de estrogénio negativos”.

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔC_T <i>CYP19A1</i> Trp39Arg is the same across categories of RE+ vs RE-.	Independent-Samples Mann-Whitney U Test	,073 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 7 - Teste Mann-Whitney U, em que a hipótese nula testada é que a média de ΔC_T de *CYP19A1* Trp39Arg é a mesma entre as categorias, “RE+ (Positivo) ” e “RE- (Negativo) ”, da variável “RE (Recetores de estrogénio) ”

Boxplot com distribuição de ΔCT de CYP19A1 Trp39Arg entre as categorias da variável RE (Recetores de estrogénio)

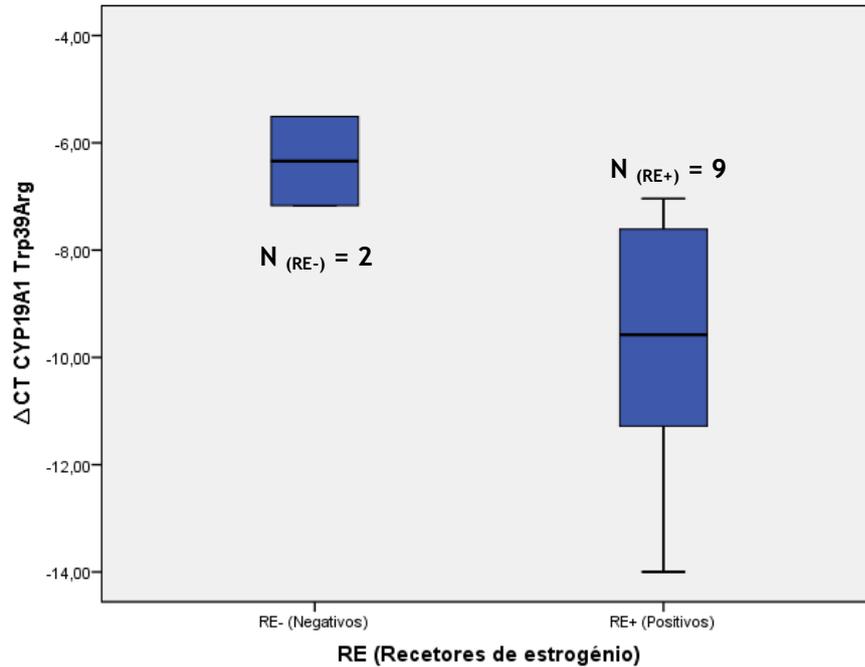


Gráfico 11 - *Boxplot* em que é ilustrada a distribuição de ΔC_T de CYP19A1 Trp39Arg entre as categorias, “RE+ (Positivo)” e “RE- (Negativos)”, da variável “RE (Recetores de estrogénio)”

A frequência absoluta em cada categoria é apresentada no gráfico.

Desta forma, pareceu-nos importante realizar uma nova análise das variáveis de interesse, tendo em consideração a variável “Recetores de estrogénio”.

Foram assim efetuados novos testes não paramétricos, representados a seguir, figura 8 e 9, para as variáveis “Envolvimento do gânglio sentinela e/ou outros” e “Estadio clínico”, traçando os respetivos *boxplot*, gráficos 12 e 13. Apenas se consideraram casos RE+ (Recetores de estrogénio tumorais positivos).

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔCT CYP19A1 Trp39Arg is the same across categories of Metastização vs Sem metastização.	Independent-Samples Mann-Whitney U Test	1,000 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 8 -Teste *Mann-Whitney U*, em que a hipótese nula testada é que a média de ΔC_T de CYP19A1 Trp39Arg é a mesma entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”, quando RE+

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔCT CYP19A1 Trp39Arg is the same across categories of Estadio Clínico (I e II vs III).	Independent-Samples Mann-Whitney U Test	,381 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 9 - Teste *Mann-Whitney U*, em que a hipótese nula testada é que a média de ΔC_T de *CYP19A1* Trp39Arg é a mesma entre as categorias, “I e II” e “III”, da variável “Estadio clínico”, quando RE+

Boxplot com distribuição de ΔCT de *CYP19A1* Trp39Arg entre as categorias da variável Envolvimento do gânglio sentinela e/ou outros, nos casos RE+

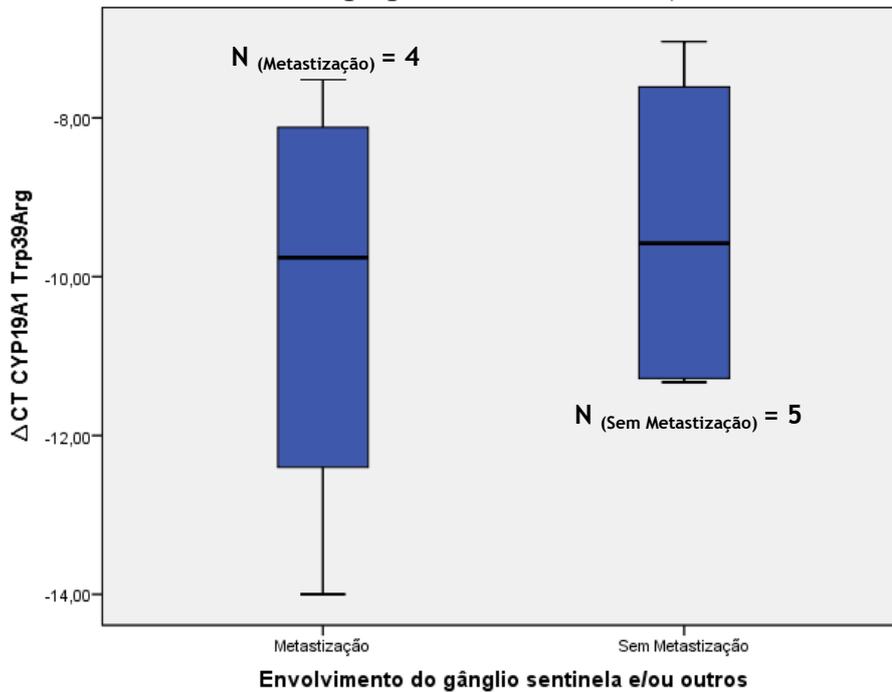


Gráfico 12 - *Boxplot* em que é ilustrada a distribuição de ΔC_T de *CYP19A1* Trp39Arg entre as categorias, “Metastização” e “Sem metastização”, da variável “Envolvimento do gânglio sentinela e/ou outros”, quando RE+

A frequência absoluta em cada categoria é apresentada no gráfico.

Boxplot com distribuição de ΔC_T de CYP19A1 Trp39Arg entre as categorias da variável Estadio clínico, nos casos RE+

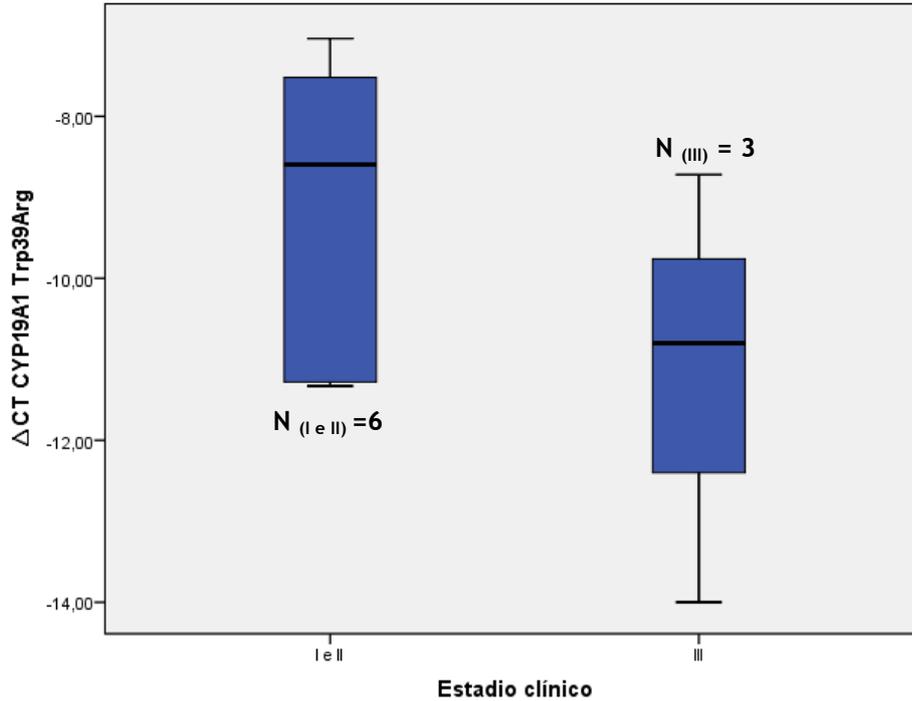


Gráfico 13 - *Boxplot* em que é ilustrada a distribuição de ΔC_T de CYP19A1 Trp39Arg entre as categorias, “I e II” e “III”, da variável “Estadio clínico”, quando RE+

A frequência absoluta em cada categoria é apresentada no gráfico.

Verifica-se que em ambas as variáveis de interesse, “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”, existe diminuição do nível de significância estatística no teste de *Mann Whitney U* (aumento de *p-value*) assim como alteração relevante na distribuição, verificada pelos *boxplot*. Estes dados sugerem que, de facto, a variável “Recetores de estrogénio” possa exercer uma influência no rácio de expressão entre as categorias das variáveis de interesse. Deve por isso ser tida em consideração no cálculo do rácio de diferenças de expressão entre as categorias das variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”.

Realizou-se ainda um outro teste de hipóteses, para o caso particular da categoria “Metastização” na variável “Envolvimento do gânglio sentinela e/ou outros”, tal como se apresenta a seguir na figura 10 e gráfico 14. Apesar de a amostra ser ainda mais reduzida, pretendemos comparar dois grupos de significado prognóstico claramente distinto, “Metastização ganglionar regional N1” vs “Metastização ganglionar regional N2 e N3”. O valor de *p-value* apresentado na figura 10 e visualmente pelo gráfico 14 sugerem uma tendência para que o rácio de expressão de CYP19A1 Trp39Arg entre estes grupos seja claramente distinto.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of ΔCT CYP19A1 Trp39Arg is the same across categories of Metastização ganglionar (N2-3) vs Metastização ganglionar (N1).	Independent-Samples Mann-Whitney U Test	,333 ¹	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

¹Exact significance is displayed for this test.

Figura 10 - Teste *Mann-Whitney U*, em que a hipótese nula testada é que a média de ΔC_T de *CYP19A1* Trp39Arg é a mesma nas subcategorias, “Metastização ganglionar, N2-3” e “Metastização ganglionar, N1”, da categoria “Metastização” da variável “Envolvimento do gânglio sentinela e/ou outros”, quando RE+

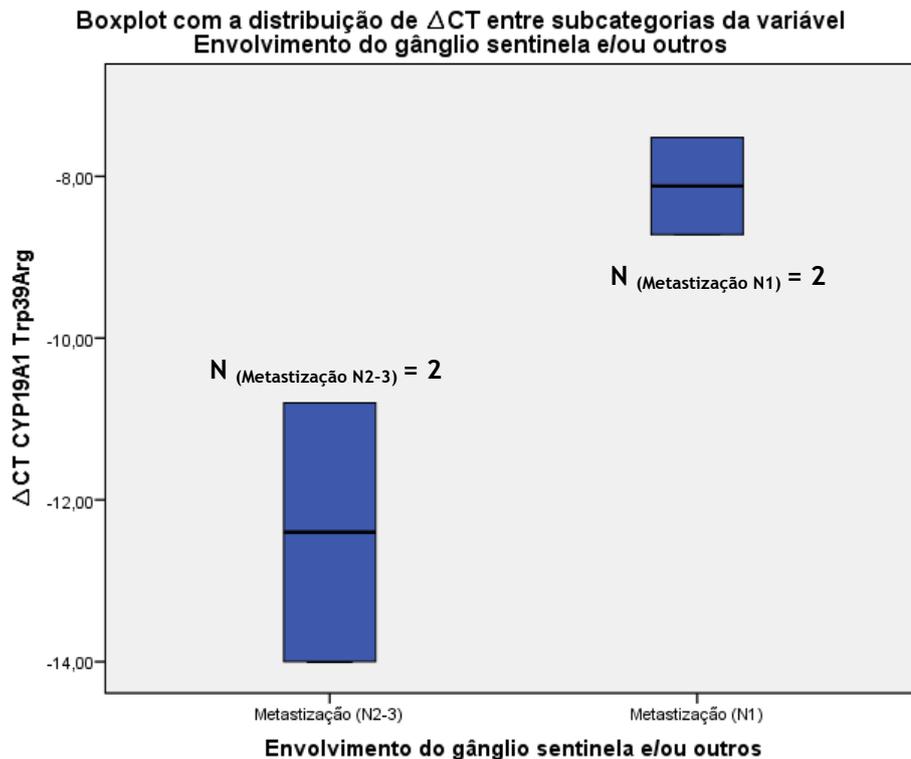


Gráfico 14 - *Boxplot* em que é ilustrada a distribuição de ΔC_T de *CYP19A1* Trp39Arg entre as subcategorias, “Metastização ganglionar, N2-3” e “Metastização ganglionar, N1”, da categoria “Metastização” da variável “Envolvimento do gânglio sentinela e/ou outros”, para casos RE+

A frequência absoluta de cada grupo é apresentada no gráfico.

3.4.2. Cálculo do rácio de expressão de *CYP19A1* Trp39Arg entre as categorias das variáveis

Neste subcapítulo calcularam-se as diferenças de expressão (rácio de expressão) de *CYP19A1* Trp39Arg (T/C) entre categorias de variáveis, segundo o método explicitado em 2.6.

A tabela 26, em baixo, evidencia os 11 casos em que ocorreu expressão da variante polimórfica da aromatase, por mRNA, com as respetivas variáveis que serão consideradas neste cálculo.

Tabela 26 - Descrição das variáveis, nos casos em que ocorreu expressão de *CYP19A1* Trp39Arg, consideradas para cálculo de diferenças na expressão da variante polimórfica da aromatase

Caso	RE	Envolvimento do gânglio sentinela e/ou outros	Estadio N	Estadio clínico	Expressão de <i>CYP19A1</i> Trp39Arg
1	Positivo	Com metastização	N2	III	Sim
13	Positivo	Sem metastização	N0	II	Sim
15	Positivo	Sem metastização	N0	II	Sim
16	Positivo	Sem metastização	N0	II	Sim
17	Positivo	Sem metastização	N0	I	Sim
19	Positivo	Com metastização	N3	III	Sim
22	Negativo	Sem metastização	N0	I	Sim
23	Positivo	Com metastização	N1	II	Sim
24	Negativo	Sem metastização	N0	II	Sim
36	Positivo	Sem metastização	N0	I	Sim
37	Positivo	Com metastização	N1	III	Sim

RE - Recetores de estrogénio

A análise previamente efetuada em 3.4.1, “Estudo de média/distribuição de ΔC_T de *CYP19A1* Trp39Arg (T/C) entre as categorias das variáveis”, sugeriu que a variável “Recetores de estrogénio (RE)” deve ser tomada em consideração neste cálculo, rácio de expressão de *CYP19A1* Trp39Arg, entre as categorias das variáveis de interesse: “Envolvimento do gânglio sentinela e/ou outros” e “Estadio clínico”. Assim para estudar estas diferenças de expressão da variante polimórfica da aromatase entre as categorias destas variáveis foram efetuados dois momentos de cálculo, não considerando a variável “RE” (tabela 27 em baixo) e considerando “RE” (tabela 29 em baixo). Intermediamente também a variável “RE” foi alvo de cálculo de diferenças de expressão de *CYP19A1* Trp39Arg entre as suas categorias (tabela 28 em baixo).

A nomenclatura, rácio de expressão, usada nas tabelas deste subcapítulo pretende evidenciar a diferença de expressão relativa entre o grupo de interesse e o grupo controlo (quantas

vezes mais a variante polimórfica da aromatase, *CYP19A1* Trp39Arg, é expressa no grupo interesse relativamente ao grupo controlo). Estes dois grupos considerados correspondem a categorias e subcategorias das diversas variáveis.

Tabela 27 - Diferença de expressão de *CYP19A1* Trp39Arg nas categorias das variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”

Variável	Grupo de Interesse	Freq. Abs.	Grupo de Controlo	Freq. Abs.	Rácio de Expressão
Envolvimento do gânglio sentinela e/ou outros	Metastização	4	Sem metastização	7	3,38
	(N1 - 3) $\sigma = 2,83$		(N0) $\sigma = 2,26$		
Estadio clínico	Estadio III $\sigma = 2,66$	3	Estadio I e II $\sigma = 2,12$	8	6,93

σ - Desvio-padrão; Freq. Abs. - Frequência Absoluta de cada grupo

Tabela 28 - Diferença de expressão de *CYP19A1* Trp39Arg nas categorias da variável “RE (Recetores de estrogénio)”

Variável	Grupo de Interesse	Freq. Abs.	Grupo de Controlo	Freq. Abs.	Rácio de Expressão
RE	RE+ (Positivo) $\sigma = 2,29$	9	RE- (Negativo) $\sigma = 1,17$	2	10,74

σ - Desvio-padrão; Freq. Abs. - Frequência Absoluta de cada grupo

Tabela 29 - Diferença de expressão de *CYP19A1* Trp39Arg nas categorias das variáveis “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”, apenas considerando casos RE+

Variável	Grupo de Interesse	Freq. Abs.	Grupo de Controlo	Freq. Abs.	Rácio de Expressão
Envolvimento do gânglio sentinela e/ou outros	Metastização (N1 - 3)	4	Sem metastização (N0)	5	1,86
	$\sigma = 2,84$		$\sigma = 2,26$		
Estadio clínico	Estadio III $\sigma = 2,66$	3	Estadio I e II $\sigma = 1,94$	6	4,33

σ - Desvio-padrão; Freq. Abs. - Frequência Absoluta de cada grupo

Quando particularizada a variável “Envolvimento do gânglio sentinela e/ou outros” estudámos também diferenças de expressão do polimorfismo da aromatase entre as subcategorias (“N1” vs “N2 -N3”) relativas à categoria “Metastização” (tabela 30 em baixo).

Tabela 30 - Diferença de expressão de *CYP19A1* Trp39Arg em subcategorias da categoria “Metastização” da variável “Envolvimento do gânglio sentinela e/ou outros”

Categoria da Variável	Grupo de Interesse	Freq. Abs.	Grupo de Controlo	Freq. Abs.	Rácio de Expressão
“Metastização” da variável	Metastização		Metastização		
“Envolvimento gânglio sentinela e/ou outros”	(N2-3) $\sigma = 2,26$	2	(N1) $\sigma = 0,85$	2	19,43

σ - Desvio-padrão; Freq. Abs. - Frequência Absoluta de cada grupo

As variáveis “Mama afetada” e “Localização afetada”, utilizadas para controlo do acaso, foram igualmente alvo de cálculo de diferenças de expressão (tabela 31 em baixo).

Tabela 31 - Diferença de expressão de *CYP19A1* Trp39Arg nas categorias das variáveis “controlo do acaso”, “Mama afetada” e “Localização afetada”

Variável	Grupo de Interesse	Freq. Abs.	Grupo de Controlo	Freq. Abs.	Rácio de Expressão
Mama afetada	Direita $\sigma = 3,23$	5	Esquerda $\sigma = 2,02$	6	1,27
Localização afetada	Outra $\sigma = 2,24$	4	Quadrante Supero-externo $\sigma = 3,29$	7	1,09

σ - Desvio-padrão; Freq. Abs. - Frequência Absoluta de cada grupo

4. Discussão/Conclusão

Para alcançar os objetivos deste estudo recorreremos a amostras de tecido de mama fixadas e incluídas em parafina de mulheres com diagnóstico de carcinoma neste órgão. O recurso a amostras de tecido em parafina apresentou diversas vantagens mas também desvantagens. A principal vantagem foi a de recorrer a material de arquivo, sem necessidade de recorrer a um novo procedimento invasivo, favorecendo assim a sua utilização. Como principal desvantagem foi o difícil manuseamento destas amostras, nomeadamente no que concerne a extração de RNA (20). Foi nossa preocupação demonstrar a possibilidade da sua utilização, assim como a sua otimização a custo favorável, que permitam não só a continuação do estudo, como também a realização de outros, recorrendo a este vasto banco de dados.

A utilização de tecido de mama histologicamente “normal” foi realizada no sentido de estudar genes somáticos da mulher e não genes tumorais. É conhecido que os tecidos que envolvem os tumores participam no seu comportamento biológico, sendo assim nossa preocupação estudar a expressão destes genes em tecido “normal” da mesma mama em que o tumor estava localizado. Também com o assegurar de ausência de células tumorais, pretendemos controlar eventuais vieses na expressão de genes. Isto porque estas células apresentam uma grande instabilidade, alvo de inúmeras mutações, muitas vezes derivadas do processo oncogénico (30). Ainda assim, uma eventual extensão do estudo, que pretenda comparar a expressão de genes entre o tecido normal e tumoral, pode ser realizada.

Todas as amostras de parafina solicitadas foram da mesma espessura, no sentido de uniformizar o material a estudar para cada caso. Com isto, foi utilizada uma porção semelhante de tecido para todos os casos, de forma a verificar as diferenças de expressão de mRNA para uma mesma quantidade aproximada de células. Ambicionámos com isto controlar a variabilidade do fragmento de tecido estudado, que poderia assim influenciar a quantidade de ácidos nucleicos extraídos.

Como já foi mencionado, o desenho do estudo contemplou a recolha de dados com recurso a processos clínicos em suporte informático. Alguns dados, como idade de menarca, primeira gestação, menopausa, antecedentes pessoais e familiares, realização de amamentação, toma de estrogénios, entre outros, também poderiam demonstrar interesse em ser recolhidos. No entanto, estes não foram registados, uma vez que não estavam disponíveis em todos os processos clínicos.

A não inclusão de pacientes que efetuaram terapia neoadjuvante prendeu-se com o facto de este tipo de terapêutica poder alterar a evolução natural da doença e também as funções biológicas na mama normal, o que poderia acarretar influências nas características histopatológicas do tumor e acometimento ganglionar, assim como alterações na expressão dos polimorfismos de interesse.

Quando existiu suspeita ou comprovação de metastização à distância, optámos também pela exclusão do estudo. Esta decisão foi tomada pelo facto dos mecanismos de metastização regional e à distância poderem ser diferentes, não sendo esta última alvo do presente estudo. No entanto, não descartamos que tais casos possam ser incluídos em futuros trabalhos, com a perspectiva de alargar o conhecimento dos genes estudados.

A codificação das amostras de parafina pelo Serviço de Anatomia Patológica foi fundamental para que os dados obtidos em âmbito laboratorial fossem analisados de forma objetiva e imparcial. Esta encriptação, só decifrada no final da análise de todos os dados laboratoriais, permitiu que o investigador não exercesse influência nos procedimentos realizados. Dos 78 casos selecionados, escolhemos trabalhar com 37, respeitando os critérios de inclusão/exclusão estabelecidos, que se basearam no princípio de maior quantidade de amostra para um determinado caso. Isto permitiu assegurar a preservação de material para uma hipotética continuação do estudo, mais abrangente e otimizado. Esta decisão teve sobretudo em consideração o facto de estarmos a realizar um estudo piloto, em que no seu decorrer poderão ser detetados erros e formas de melhorar os métodos usados.

O RNA extraído não foi possível ser quantificado em termos absolutos, pois considerámos que as medições efetuadas não foram consistentes. No entanto, este passo não foi limitante, uma vez que o nosso objetivo foi a determinação de diferenças de expressão relativas de genes, entre grupos. Foi, deste modo, um aspeto essencial, certificar que as amostras estavam normalizadas, o que foi garantido pela utilização de um gene *housekeeping*.

A expressão dos polimorfismos de interesse foi determinada por RT-PCR e os protocolos utilizados tiveram que ser adaptados às pequenas quantidades de cDNA. Foi necessário utilizar 60 ciclos de amplificação devido à baixa quantidade de cDNA esperada, em consequência do também baixo mRNA obtido. Com este número de ciclos pretendemos aumentar a sensibilidade de deteção do fragmento de interesse. Ou seja, o aumento do número de ciclos permitiu maior amplificação de cDNA, por forma a ser detetado quando na presença do fragmento de interesse.

A realização de triplicados das amostras em RT-PCR é de importante significado, porém acarreta custos substancialmente mais avultados, assim como a utilização de três vezes mais quantidade de amostra. Estamos conscientes de que a sua não utilização também aumenta a suscetibilidade a erros, em que se inclui o ato de pipetar, e consequentemente pode influenciar o cálculo preciso do C_T e ΔC_T para cada caso. Porém, saliente-se que como o nosso objetivo foi a determinação de diferenças relativas de expressão entre grupos e não entre casos, ter-se-á em consideração uma medida de tendência central, a média de ΔC_T por grupo. Desta forma e tendo em consideração a diminuta quantidade de cDNA sintetizada, optámos pela não realização de triplicados das amostras.

A análise dos resultados obtidos para o gene *housekeeping*, β -actina, foi necessária para validar e definir qual a amostragem definitiva, posteriormente analisada para os genes de interesse e para estudo estatístico. Esta validação, mediante a aplicação de diversos critérios, mostrou-se decisiva porque permitiu eliminar os casos em que a amplificação do fragmento do gene da β -actina foi dúbia. Os gráficos 2 e 3 permitem constatar a presença inicial de casos em que os resultados deveriam ser obrigatoriamente confirmados devido ao facto de o seu valor estar muito afastado do valor de mediana. Uma vez que não foi possível a repetição do procedimento, a sua exclusão permitiu uma maior uniformidade da temperatura de *melting* na amostragem definitiva considerada (gráfico 4 e tabela 12). A medida de tendência central usada nestes critérios de validação considerou, tal como já referido, a existência de valores discrepantes. Estes podem influenciar a média, por isso optou-se pela utilização da mediana, que é uma medida de tendência central mais robusta, quando estão presentes.

A análise dos resultados obtidos para a expressão do gene *housekeeping* foi ainda fulcral para determinar a eficácia dos métodos laboratoriais realizados. Foi, para isso, determinante a utilização de *primers* de β -actina situados em regiões de fronteira de exões, sendo desta forma, altamente discriminativos para cDNA, em detrimento de DNA genómico. Salienta-se que a eficácia global dos métodos usados para desparafinação, extração de mRNA e respetiva síntese em cDNA foi de cerca de 81,1%. Este cálculo foi realizado tendo em conta o número de casos processados nos diversos métodos laboratoriais supracitados, 37, e o número de casos em que foi possível afirmar a presença de gene *housekeeping*, 30.

Em nenhum dos controlos de reação (NTC) relativos à β -actina foi detetada leitura de fluorescência. Este dado leva-nos a crer que os *primers* utilizados têm muito pouca tendência para a formação de dímeros e/ou sequências inespecíficas.

Como podemos observar, o valor de C_T da β -actina foi superior ao C_T dos genes em estudo, nomeadamente da variante polimórfica da aromatase, que foi o único em que ocorreu deteção do fragmento de interesse. Isto pode ser explicado por níveis aumentados de expressão da *CYP19A1 Trp39Arg*, mas não podemos descartar a hipótese de este facto se dever a, se ter otimizado as condições para esta última, estando o valor de C_T da β -actina aumentado. Esta particularidade é especialmente significativa na fase de *annealing*, sendo a temperatura específica para cada par de *primers*. Porém, este aspeto não representa nenhuma limitação, uma vez que a presença do gene *housekeeping* pretende sobretudo normalizar a expressão do gene de interesse, como já se explicitou.

Já a observação dos resultados da variante polimórfica da aromatase, *CYP19A1 Trp39Arg*, sugere, com um elevado grau de suspeição, a existência de dímeros de *primer*. Esta suspeita reside no facto de os controlos de reação (NTC), em que não foi colocado cDNA, terem obtido leitura de fluorescência, podendo estar amplificada em virtude das amostras serem muito diminutas.

Sugerimos que o *forward primer* seja reformulado ou alterado, em caso de prossecução do estudo desta variante polimórfica da aromatase. No caso particular do *reverse primer*, é conveniente a sua preservação, não apenas pela sua menor capacidade de formar dímeros, mas sobretudo porque é este que reconhece a variante polimórfica. Tal acontece pela sua última base nucleotídica, na região 3'. Como esta é a região do *primer* que apresenta maior especificidade (31, 32), é a que, por conseguinte, permite o reconhecimento da variável polimórfica da aromatase de interesse.

Importa também ressaltar que em todos os casos em que foi determinada a presença de expressão de *CYP19A1 Trp39Arg* foram detetadas duas temperaturas de *melting* distintas. Uma que por nós foi interpretada como a de formação de dímeros, dada a sua similitude com a temperatura dos controlos de reação (NTC), e uma segunda mais alta, por nós entendida como a de amplificação do fragmento de interesse.

Atendendo à análise dos resultados, verificamos que cerca de 36,7% da amostra estudada exhibe expressão da variante polimórfica da aromatase, *CYP19A1 Trp39Arg*.

Num dos estudos que serviu de partida ao nosso trabalho, determinou-se que a prevalência deste polimorfismo no genótipo de mulheres com diagnóstico histopatológico de cancro da mama é de 60,4%, sendo a frequência alélica de 32,2% (15). A amostra desse estudo baseou-se na mesma população que o nosso, i.e., população de mulheres inscritas no CHCB. No entanto, a semelhança desta frequência alélica com a expressão por nós determinada pode ser uma casualidade, tendo em conta que não foram estudadas as mesmas pacientes. Esta semelhança apenas será esclarecida se for realizada a determinação do genótipo nos casos por nós estudados. Sugerimos, portanto, a sua realização para que seja melhor esclarecida a relação entre presença no genótipo e sua expressão por mRNA.

Relativamente ao genótipo de *CYP19A1 Trp39Arg* salienta-se também, que a sua presença no património genético da população em estudo é muito superior comparativamente a outras populações estudadas no Japão (33), Índia (34) e *Han-Chinese-American* (35). No entanto, é semelhante a uma população europeia, Polónia (36).

No que concerne aos resultados obtidos para os genes *GSTM1* e *GSTT1*, estes são a favor de uma ausência de expressão em todos os casos estudados.

Desta forma, estes não foram considerados para posterior análise e estudo estatístico, uma vez que não ocorreu dicotomia de resultados, i.e., casos com e sem expressão destes genes. Dado que os resultados apoiam a inexistência de expressão em todos os casos, não faria qualquer sentido realizar uma análise comparativa de diferenças de expressão, entre diversos grupos.

Porém, não foi só este o motivo pelo qual não se realizou qualquer alusão a estes genes ao longo da análise descritiva e estatística. A principal causa do sucedido fundamentou-se na

estranheza dos resultados obtidos, uma vez que um dos estudos que fomentou o nosso trabalho deu conta que a presença de *GSTM1* foi detetada em 36,5% e a presença de *GSTT1* em 56,5%, numa amostra da mesma população alvo do nosso estudo (16). Atendendo a estes dados, seria de esperar que uma proporção considerável da nossa amostra apresentasse expressão destes genes. Ainda que numa remota hipótese, poderíamos estar apenas a estudar casos em que o genótipo foi *null* para ambos os genes.

No entanto, com recurso à ferramenta informática *DBATE (DataBase of Alternative Transcripts Expression)*, constatámos que os transcritos canónicos destes genes, que são os alvos dos nossos *primers*, são expressos em quantidades muito baixas no tecido mamário. Esta expressão situa-se na ordem de 10^{-2} *FPKM (fragments per kilobase)*, o que, associado às consideráveis baixas quantidades de RNA passíveis de serem extraídos em tecidos de parafina, leva à explicação mais viável para o sucedido. Como tal, desencoraja-se a continuação do estudo de expressão por mRNA de *GSTM1* e *GSTT1*, em tecido de mama, proveniente de amostras de parafina.

A determinação do genótipo dos genes *GSTT1* e *GSTM1*, dos casos por nós estudados, poderá ser realizada para uma melhor caracterização da amostra, contudo não representará uma explicação na ausência da expressão destes genes.

À semelhança do que vimos para a variante polimórfica da aromatase, também pareceu existir amplificação de dímeros de *primers* nas reações efetuadas para os genes *GSTT1* e *GSTM1*, só que, neste caso, não foram detetados os fragmentos de interesse. A presença de hipotéticos dímeros, tal como sugerido para *CYP19A1 Trp39Arg*, também poderá ser confirmada por sequenciação.

A presença de eventuais contaminações também poderia ser ponderada, uma vez que nos controlos negativos (NTC) de *GSTT1*, *GSTM1* e também *CYP19A1 Trp39Arg* ocorreu deteção de fluorescência. Porém, tendo em consideração que foram sempre utilizados os mesmos reagentes e o mesmo extrato de cDNA para cada caso, excetuando os *primers* que foram específicos de cada gene, a ausência de qualquer leitura de fluorescência detetada nos controlos negativos (NTC) do gene β -actina, levam à exclusão de contaminações nos reagentes comuns e no extrato de cDNA de cada caso, restando apenas os *primers* específicos dos genes de interesse como suspeitos. Como já foi sendo discutido, parece-nos que esta deteção de fluorescência seja atribuível à formação de dímeros de *primers*, tendo sido sugerida a sequenciação para deslindar esta incógnita. Com a realização deste procedimento também pretendemos verificar a possibilidade de se tratar de contaminações ou amplificações inespecíficas.

Tal como se constatou no capítulo referente aos resultados, apenas foi alvo de estudo estatístico a expressão de *CYP19A1 Trp39Arg*, que passaremos agora a discutir.

Na análise estatística, tanto nos testes de qui-quadrado de *Pearson* como nos testes de *Mann-Whitney U* não foi encontrada significância estatística em qualquer variável, quando considerado um intervalo de confiança de 95%. Estamos cientes de que o tamanho da amostra é determinante para este tipo de análise, podendo no nosso estudo, com diminuta amostragem, precipitar um erro tipo II. Este define-se como a probabilidade de a hipótese nula ser falsa, mas que não está a ser rejeitada (37).

No “Estudo da associação entre a expressão de *CYP19A1* Trp39Arg e as variáveis qualitativas clínicas e histopatológicas” (subcapítulo 3.3) não obtivemos diferenças estatisticamente significativas para as variáveis de interesse “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros” (tabela 20). Não parece, portanto, existir evidência de uma tendência para que a presença ou ausência de expressão seja distinta nas categorias das variáveis de interesse. Deste modo, não foi possível determinar uma relação entre estas variáveis consideradas e a presença ou ausência de expressão de *CYP19A1* Trp39Arg, razão pela qual não foi calculado o *odds ratio*.

As variáveis usadas para controlo do acaso, “Mama afetada” e “Localização afetada” também não obtiveram diferenças estatisticamente significativas entre as suas categorias e a ausência ou presença da expressão do polimorfismo da aromatase. Estes resultados estão em linha de conta com o que foi previsto - a inexistência de qualquer associação com a variável presença ou ausência de expressão de *CYP19A1* Trp39Arg. Os resultados obtidos para estas variáveis fortalecem que o acaso possa exercer menor influência do que o expectável.

Pode-se, desta forma, concluir que não existem evidências ou tendências que apontem para uma expressão de *CYP19A1* Trp39Arg, em termos qualitativos (expressa ou não expressa), distinta entre grupos de diferente prognóstico, correspondentes às categorias das variáveis de interesse.

Ao estudarmos as “Diferenças de expressão da variante polimórfica *CYP19A1* Trp39Arg (T/C), entre grupos” (subcapítulo 3.4), foi necessário ter atenção à casuística ainda mais reduzida, de 11 casos, que correspondem àqueles em que foi detetada expressão da variante polimórfica da aromatase. Desta forma, a sua análise consistiu em dois pontos fundamentais:

- 1 - Garantir/Prever que as variáveis de interesse não foram influenciadas por outras, nem pelo acaso (subcapítulo 3.4.1 - “Estudo da média/distribuição de ΔC_T de *CYP19A1* Trp39Arg entre as categorias das variáveis”).

- 2- Determinar o rácio de expressão entre as categorias das variáveis de interesse. (subcapítulo 3.4.2 - “Cálculo do rácio de expressão de *CYP19A1* Trp39Arg entre as categorias das variáveis”). Foi só mediante este cálculo que foi analisado o seu significado fisiológico.

O 1º ponto foi efetuado com o intuito de auxiliar a uma correta e validada interpretação dos resultados obtidos para o cálculo das diferenças de expressão relativa entre as categorias das

variáveis de interesse. Para isso recorreremos ao teste não paramétrico de *Mann Whitney U* e elaboração de *boxplots*.

Pela utilização do teste não paramétrico verificamos que em nenhuma variável houve significância estatística (*p-value* inferior a 0,05). Contudo, a visualização dos *boxplot* permite prever que poderão existir diferenças de expressão do polimorfismo da aromatase entre as categorias das variáveis “Envolvimento do gânglio sentinela e/ou outros”, “Estadio clínico” e “Recetores de estrogénio” (gráficos 7, 8 e 11, respetivamente).

Após um olhar atento sobre estas três variáveis, verificámos que aquela que apresentou maior relevância estatística foi a variável “Recetores de estrogénio”, com *p-value* de 0,073. Apesar de apenas existirem 9 casos RE+ tumorais, com expressão de *CYP19A1* Trp39Arg e 2 casos RE-, há tendência a uma evidente diferença nas médias de ΔC_T do polimorfismo da aromatase entre as duas categorias (gráfico 11). Estes dados sugerem que esta variável (Recetores de estrogénio) deverá ser considerada para o posterior cálculo das diferenças de expressão, em termos relativos, nas variáveis de interesse.

Tendo em consideração a íntima relação dos estrogénios com a patogenia do cancro da mama (9, 10) e o facto de a aromatase participar na biossíntese de estrogénios, determinando o seu passo final de síntese, conversão de androgénios a estrogénios (18), justifica-se a relação encontrada entre expressão de *CYP19A1* Trp39Arg e recetores de estrogénio.

Assim, realizámos uma reanálise das variáveis de interesse, tendo em conta a variável “Recetores de estrogénio” (gráfico 12 e 13) que, como referido no subcapítulo 3.4.1, sugere a possível influência nas primeiras.

Pareceu-nos desta forma importante considerar cautelosamente a variável “Recetores de estrogénio”, quando efetuado o cálculo do rácio de expressão relativa de *CYP19A1* Trp39Arg entre as categorias das variáveis de interesse, “Estadio clínico” e “Envolvimento do gânglio sentinela e/ou outros”.

Uma vez que a amostragem é reduzida (11 casos), o acaso poderá exercer influência, devendo todos os potenciais influenciadores dos resultados serem tomados em consideração. Desta forma, a análise preliminar da média/distribuição de ΔC_T de *CYP19A1* Trp39Arg entre as categorias das variáveis, realizada previamente ao cálculo do rácio de expressão em termos relativos, foi importante, pois permitiu constatar, hipoteticamente, a existência de uma variável que devesse ser tomada em atenção neste cálculo. Devemos reforçar que esta análise não pretendeu inferir ou atribuir qualquer valor fisiológico na comparação de ΔC_T de *CYP19A1* Trp39Arg (T/C), mas apenas complementar o posterior cálculo do rácio de expressão entre as categorias das variáveis de interesse, em termos relativos.

No 2º ponto, quando determinámos o rácio de expressão entre as categorias das variáveis, verificámos que a variante polimórfica da aromatase em estudo é cerca de 10,74 vezes mais

expressa no grupo RE+ comparativamente a RE- (tabela 28). Este resultado exprime a significativa diferença da expressão do *CYP19A1* Trp39Arg entre estes dois grupos e está em linha com a análise realizada no primeiro ponto, que sugeria que esta variável (RE) deveria ser considerada quando comparadas as variáveis de interesse.

Na variável “Envolvimento do gânglio sentinela e/ou outros”, dos 11 casos em que se detetou expressão de *CYP19A1* Trp39Arg, em 4 houve metastização ganglionar regional e em 7 casos não houve. O rácio de expressão relativa de *CYP19A1* Trp39Arg entre os grupos “Metastização/Sem metastização ganglionar”, de 3,38 (tabela 27), é indicativo de que o grupo com envolvimento ganglionar por metastização tem uma expressão consideravelmente maior que o grupo sem envolvimento ganglionar. Porém, constatámos que todos os casos em que ocorreu envolvimento ganglionar, com presença de expressão da variante polimórfica da aromatase, são RE+ tumorais. Os únicos dois casos RE- tumorais, com expressão deste polimorfismo, pertencem ao grupo sem metastização. Desta forma e dado as diferenças encontradas para a expressão entre as categorias da variável “RE”, pareceu-nos importante considerar apenas os casos “RE+” para o cálculo das diferenças de expressão das categorias da variável “Envolvimento do gânglio sentinela e/ou outros”. Nesta reanálise (tabela 29) observámos que o rácio da expressão do polimorfismo da aromatase, entre as categorias da variável “Envolvimento do gânglio sentinela e/ou outros”, continua a ser maior no grupo “Metastização” comparativamente a “Sem metastização”, mas mais próximo de 1. Assim, a descida do rácio de expressão de 3,38 para 1,86, entre os grupos “com metastização” vs “sem metastização”, confirma a possível influência que os casos “RE-“ poderiam exercer no cálculo inicialmente efetuado. Parece não existir diferença de expressão significativa de *CYP19A1* Trp39Arg, nos grupos com e sem metastização ganglionar regional. Contudo, com estes resultados não podemos ser conclusivos, sobretudo porque a casuística é ainda mais diminuta, após reanálise considerando a variável “RE”.

Na variável “Estadio Clínico”, realizámos primeiramente o cálculo do rácio de expressão entre categorias, com os 11 casos em que foi detetada expressão de *CYP19A1* Trp39Arg. O número de casos na categoria “estadio I e II” é de 8 e na categoria “estadio III” é de 3, sendo o rácio de expressão entre “estadio III/estadio I e II” de 6,93 (tabela 27). Estes dados também apontariam para uma diferença considerável de expressão entre estas categorias, porém, à semelhança do que ocorreu com a variável “Envolvimento do gânglio sentinela e/ou outros”, os dois casos RE- pertencem ao grupo “estadio I e II”, estadios estes associados a um melhor prognóstico que o “estadio III” (38). Do mesmo modo, efetuámos uma análise subsequente desta variável, tendo em consideração a ausência ou presença de “Recetores de estrogénio”, i.e., apenas considerando os casos RE+. O rácio de expressão calculado é assim de 4,33 (tabela 29), inferior ao inicialmente determinado (6,93), mas continuando a ser expressivo no que concerne às diferenças de expressão entre as categorias desta variável. Parece existir uma tendência de expressão distinta de *CYP19A1* Trp39Arg, entre estes dois grupos considerados (“Estadio III” vs “Estadio I e II”).

Importa também salientar que a variável “Estadio clínico”, representa um fator de prognóstico (38), que tem em consideração o tamanho do tumor, o envolvimento ganglionar regional e a metastização à distância (28), sendo mais abrangente que a variável “Envolvimento do gânglio sentinela e/ou outros”, ao englobá-la. Assim, a variável “Estadio clínico” é um melhor preditor de prognóstico do que a variável “Envolvimento do gânglio sentinela e/ou outros”.

Constatámos, portanto, uma expressão aumentada de *CYP19A1* Trp39Arg, 4,33 vezes maior, no grupo de pior prognóstico, “Estadio III”, comparativamente ao grupo de melhor prognóstico, “Estadio I e II”. Especulamos que a expressão desta variante polimórfica da aromatase poderá ter algum tipo de influência na determinação do prognóstico do cancro da mama, nesta população estudada.

Foram ainda apreciados dois grupos de distinto prognóstico, tendo-se agora em consideração o número de gânglios envolvidos em casos com metastização regional (39). Desta forma, consideraram-se dois subgrupos (“N1” e “N2-3”), na categoria “Metastização” da variável “Envolvimento do gânglio sentinela e/ou outros”. Em todos estes casos os recetores de estrogénio tumorais foram positivos. Por coincidência, no grupo de pior prognóstico (“N2-3”) (39), todos os casos apresentaram disrupção da cápsula ganglionar. Esta última propicia a metastização dos gânglios axilares por invasão do tecido adiposo, em volta dos gânglios linfáticos que sofreram rutura (40). Foram assim definidas as subcategorias “Metastização a um gânglio linfático axilar ipsilateral (N1) sem disrupção da cápsula” e “Metastização a mais que um gânglio linfático (N2-3) com disrupção da cápsula”. Como já referimos, estamos perfeitamente cientes da ainda mais reduzida casuística aqui considerada, contudo, pretendemos realçar o rácio de expressão entre estas subcategorias (tabela 30), de 19,43. A expressão de *CYP19A1* Trp39Arg é quase 20 vezes mais expressa no grupo de pior prognóstico, “Metastização a mais que um gânglio linfático (N2-3) com disrupção da cápsula”.

As ilações que considerámos ao longo do 2º ponto da determinação de “Diferenças de expressão da variante polimórfica *CYP19A1* Trp39Arg (T/C), entre grupos” podem ser claramente influenciadas pela diminuta amostragem em estudo, colocando por isso em causa a robustez dos resultados obtidos e conseqüentemente da análise efetuada. Existe, desta forma, uma grande vulnerabilidade da amostra, para que o acaso exerça uma proeminente influência nas conclusões retiradas.

Para confirmar que estes resultados obtidos não foram ao acaso, foi decisiva a análise das duas variáveis “controlo do acaso” - “Mama afetada” e “Localização afetada”. Como já referido, é expectável que estas não apresentem quaisquer diferenças de expressão de *CYP19A1* Trp39Arg entre as suas categorias. Quando efetuámos o cálculo do rácio de expressão entre as suas categorias, constatámos uma grande semelhança (tabela 31). Sendo que a “Mama direita” expressa 1,27 vezes mais que a “Mama esquerda” e uma “Outra localização” expressa 1,09 vezes mais que o “Quadrante supero-externo”. Os resultados

obtidos para estas variáveis controlo permitem aumentar a validade e fortalecer as inferências referentes à análise dos resultados obtidos, para as variáveis de interesse, apoiando que estes não são somente dependentes do acaso, em função da diminuta casuística.

As evidências já elucidadas, parecem favorecer a hipótese de: “a expressão de *CYP19A1* Trp39Arg em tecido de mama «normal», envolvente ao tumor, pode auxiliar na determinação do prognóstico do cancro da mama”, sendo que existe uma clara tendência a que esta variante seja mais expressa em grupos de pacientes com pior prognóstico.

Apesar de reconhecermos a reduzida amostragem, parece que, nesta população considerada, a variante polimórfica da aromatase em estudo poderá exibir algum potencial na determinação do prognóstico de cancro da mama. Desta forma, será necessário um estudo mais abrangente, de maior casuística e com maior robustez estatística, para melhor apreciar as inferências que os nossos resultados sugerem.

Com este estudo preliminar alcançámos os objetivos do nosso trabalho relativamente ao polimorfismo da aromatase, possibilitando determinar que o polimorfismo *CYP19A1* Trp39Arg (T/C) poderá ter valor na definição do prognóstico do cancro da mama, na população estudada. Desenvolvemos e otimizámos uma técnica que permitiu a determinação da expressão através de mRNA de *CYP19A1* Trp39Arg (T/C) em amostras de tecido parafinado. No entanto, não foi possível determinar o valor prognóstico dos polimorfismos da *GSTM1* e *GSTT1* pela ausência de expressão através de mRNA dos referidos genes.

Ao nosso conhecimento, este trabalho foi o único que analisou a expressão de *CYP19A1* Trp39Arg com indicadores de prognóstico do cancro da mama. Em apenas um estudo há alusão a que os portadores deste polimorfismo, no seu genótipo, poderão estar associados à metastização de gânglios linfáticos (17), contudo, não foi estudada a expressão do mesmo. Os resultados obtidos com o nosso trabalho favorecem a primeira hipótese por nós postulada, “H₁ - O polimorfismo da aromatase, *CYP19A1* Trp39Arg (T/C), é mais expresso em grupos de pior prognóstico de cancro da mama”. Já relativamente à segunda e terceira hipóteses, “H₂- O polimorfismo *GSTM1 null* é mais expresso em grupos de pior prognóstico de cancro da mama” e “H₃- O polimorfismo *GSTT1 null* é mais expresso em grupos de pior prognóstico de cancro da mama”, nada poderemos concluir.

Decorrente deste trabalho é possível estabelecer as seguintes linhas futuras.

Perspetivas futuras:

O nosso grupo de trabalho primeiramente demonstrou que a *CYP19A1* Trp39Arg está associada a um risco aumentado de cancro da mama. O presente trabalho veio demonstrar que a expressão deste polimorfismo pode estar relacionada com o estadio clínico. Seria interessante avaliar o comportamento biológico da expressão de *CYP19A1* Trp39Arg, assim como a sua possível participação nos processos de metastização à distância, incluindo a disseminação com o tipo de órgão.

Outro ponto de interesse clínico relaciona-se com a morbilidade associada à linfadectomia axilar, para o diagnóstico de metastização ganglionar regional. O linfedema do membro superior pós cirurgia do cancro da mama é uma complicação relativamente frequente (entre 9 a 40%) (41, 42), particularmente após o esvaziamento da cadeia ganglionar axilar (41, 43). Esta morbilidade representa efeitos significativos na qualidade da vida dos pacientes acometidos (43), sobretudo pelo seu difícil tratamento (44). Deste modo, possíveis preditores que auxiliem a melhorar as indicações para a realização deste procedimento têm um relevante interesse. Assim, tal como proposto com a extensão do estudo, poderão ainda ser avaliados e melhor esclarecidos os interessantes indícios da grande diferença de expressão documentada, da variante *CYP19A1* Trp39Arg, entre casos com envolvimento de mais que um gânglio regional e disrupção da cápsula, comparativamente aos casos com apenas um gânglio axilar ipsilateral envolvido. A determinação da *CYP19A1* Trp39Arg poderá ser avaliada no sentido de melhorar a seleção de doentes que terão de fazer linfadectomia por gânglio sentinela positivo.

Referências

As referências bibliográficas foram escritas de acordo com o sistema de referência de *Vancouver*.

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA: a cancer journal for clinicians*. 2011 Mar-Apr;61(2):69-90.
2. Filippini SE, Vega A. Breast cancer genes: beyond BRCA1 and BRCA2. *Frontiers in bioscience*. 2013;18:1358-72.
3. Di Pietro G, Magno LA, Rios-Santos F. Glutathione S-transferases: an overview in cancer research. *Expert opinion on drug metabolism & toxicology*. 2010 Feb;6(2):153-70.
4. Ripperger T, Gadzicki D, Meindl A, Schlegelberger B. Breast cancer susceptibility: current knowledge and implications for genetic counselling. *European journal of human genetics : EJHG*. 2009 Jun;17(6):722-31.
5. Weber BL, Nathanson KL. Low penetrance genes associated with increased risk for breast cancer. *European journal of cancer*. 2000 Jun;36(10):1193-9.
6. Mitrunen K, Hirvonen A. Molecular epidemiology of sporadic breast cancer. The role of polymorphic genes involved in oestrogen biosynthesis and metabolism. *Mutation research*. 2003 Sep;544(1):9-41.
7. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA : the journal of the American Medical Association*. 2002 Jul 17;288(3):321-33.
8. Prentice RL, Anderson GL. The women's health initiative: lessons learned. *Annual review of public health*. 2008;29:131-50.
9. Liang J, Shang Y. Estrogen and cancer. *Annual review of physiology*. 2013;75:225-40.
10. Cavalieri EL, Rogan EG. Unbalanced metabolism of endogenous estrogens in the etiology and prevention of human cancer. *The Journal of steroid biochemistry and molecular biology*. 2011 Jul;125(3-5):169-80.
11. Bhat HK. Critical role of oxidative stress in estrogen-induced carcinogenesis. *Proceedings of the National Academy of Sciences*. 2003;100(7):3913-8.
12. Klaunig JE, Kamendulis LM, Hocevar BA. Oxidative stress and oxidative damage in carcinogenesis. *Toxicologic pathology*. 2010 Jan;38(1):96-109.

13. Tas F, Hansel H, Belce A, Ilvan S, Argon A, Camlica H, et al. Oxidative stress in breast cancer. *Medical oncology*. 2005;22(1):11-5.
14. Yadav A, Mishra PC. Modeling the activity of glutathione as a hydroxyl radical scavenger considering its neutral non-zwitterionic form. *Journal of molecular modeling*. 2013 Feb;19(2):767-77.
15. Ramalhinho AC, Fonseca-Moutinho JA, Breitenfeld Granadeiro LA. Positive association of polymorphisms in estrogen biosynthesis gene, CYP19A1, and metabolism, GST, in breast cancer susceptibility. *DNA and cell biology*. 2012 Jun;31(6):1100-6.
16. Ramalhinho AC, Fonseca-Moutinho JA, Breitenfeld L. Glutathione S-transferase M1, T1, and P1 genotypes and breast cancer risk: a study in a Portuguese population. *Molecular and cellular biochemistry*. 2011 Sep;355(1-2):265-71.
17. Miyoshi Y, Iwao K, Ikeda N, Egawa C, Noguchi S. Breast cancer risk associated with polymorphism in CYP19 in Japanese women. *International journal of cancer Journal international du cancer*. 2000 Jul 20;89(4):325-8.
18. Auchus ML, Auchus RJ. Human steroid biosynthesis for the oncologist. *Journal of investigative medicine : the official publication of the American Federation for Clinical Research*. 2012 Feb;60(2):495-503.
19. Korbler T, Grskovic M, Dominis M, Antica M. A simple method for RNA isolation from formalin-fixed and paraffin-embedded lymphatic tissues. *Experimental and molecular pathology*. 2003 Jun;74(3):336-40.
20. Evers DL, He J, Kim YH, Mason JT, O'Leary TJ. Paraffin embedding contributes to RNA aggregation, reduced RNA yield, and low RNA quality. *The Journal of molecular diagnostics : JMD*. 2011 Nov;13(6):687-94.
21. Crisóstomo LDM. Pilot-model for oxidative post-competition recovery in swimmers [MSc thesis]. Covilhã: Universidade da Beira Interior; 2013.
22. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols*. 2008;3(6):1101-8.
23. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 2001 Dec;25(4):402-8.
24. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research*. 2001 May 1;29(9):e45.
25. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012 Oct 4;490(7418):61-70.

26. Bansal C, Singh US, Misra S, Sharma KL, Tiwari V, Srivastava AN. Comparative evaluation of the modified Scarff-Bloom-Richardson grading system on breast carcinoma aspirates and histopathology. *CytoJournal*. 2012;9:4.
27. Manikandan S. Measures of dispersion. *Journal of pharmacology & pharmacotherapeutics*. 2011 Oct;2(4):315-6.
28. Edge S, Byrd, DR, Compton, CC, Fritz, AG, Greene, FL, Trotti, A (Eds.). *AJCC Cancer Staging Manual*. 7 ed: New York: Springer; 2010.
29. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Annals of surgical oncology*. 2010 Jun;17(6):1471-4.
30. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability--an evolving hallmark of cancer. *Nature reviews Molecular cell biology*. 2010 Mar;11(3):220-8.
31. Lefever S, Pattyn F, Hellemans J, Vandesompele J. Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clinical chemistry*. 2013 Oct;59(10):1470-80.
32. Wu JH, Hong PY, Liu WT. Quantitative effects of position and type of single mismatch on single base primer extension. *Journal of microbiological methods*. 2009 Jun;77(3):267-75.
33. Hirose K, Matsuo K, Toyama T, Iwata H, Hamajima N, Tajima K. The CYP19 gene codon 39 Trp/Arg polymorphism increases breast cancer risk in subsets of premenopausal Japanese. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2004 Aug;13(8):1407-11.
34. Samson M, Rama R, Swaminathan R, Sridevi V, Nancy KN, Rajkumar T. CYP17 (T34C), CYP19 (Trp39Arg), and FGFR2 (C906T) polymorphisms and the risk of breast cancer in south Indian women. *Asian Pacific journal of cancer prevention : APJCP*. 2009 Jan-Mar;10(1):111-4.
35. Ma CX, Adjei AA, Salavaggione OE, Coronel J, Pelleymounter L, Wang L, et al. Human aromatase: gene resequencing and functional genomics. *Cancer research*. 2005 Dec 1;65(23):11071-82.
36. Sobczuk A, Romanowicz H, Fiks T, Polac I, Smolarz B. The CYP17 and CYP19 gene single nucleotide polymorphism in women with sporadic breast cancer. *Polish journal of pathology : official journal of the Polish Society of Pathologists*. 2009;60(4):163-7.
37. Rothman KJ. Curbing type I and type II errors. *European journal of epidemiology*. 2010 Apr;25(4):223-4.

38. Ording AG, Nielsson MS, Froslev T, Friis S, Garne JP, Sogaard M. Completeness of breast cancer staging in the Danish Cancer Registry, 2004-2009. *Clinical epidemiology*. 2012;4 Suppl 2:11-6.
39. Ajithkumar TV, Barrett A, Hatcher H, Cook N. *Oxford Desk Reference: Oncology*. Oxford: Oxford University Press; 2011.
40. Tille JC, Egger JF, Devillaz MC, Vlastos G, Pelte MF. Frozen section in axillary sentinel lymph nodes for diagnosis of breast cancer micrometastasis. *Anticancer research*. 2009 Nov;29(11):4711-6.
41. Degnim AC, Miller J, Hoskin TL, Boughey JC, Loprinzi M, Thomsen K, et al. A prospective study of breast lymphedema: frequency, symptoms, and quality of life. *Breast cancer research and treatment*. 2012 Aug;134(3):915-22.
42. Suami H, Chang DW. Overview of surgical treatments for breast cancer-related lymphedema. *Plastic and reconstructive surgery*. 2010 Dec;126(6):1853-63.
43. Morrell RM, Halyard MY, Schild SE, Ali MS, Gunderson LL, Pockaj BA. Breast cancer-related lymphedema. *Mayo Clinic proceedings*. 2005 Nov;80(11):1480-4.
44. Warren AG, Brorson H, Borud LJ, Slavin SA. Lymphedema: a comprehensive review. *Annals of plastic surgery*. 2007 Oct;59(4):464-72.

Glossário

Adutos de DNA - nome dado a formas de DNA alteradas, que desencadeiam carcinogénese, por se encontrarem ligadas a substâncias com potencial carcinogénico.

Água DEPC - água que contém o composto dicarbonato de dietila, um inativador das *RNases*.

Boxplot - diagrama que permite observar a distribuição de uma variável.

Cartão FTA® - estrutura que permite proteger os ácidos nucleicos, mediante a desnaturação de proteínas, inativação de microrganismos, protegendo ainda dos efeitos da nuclease e danos provocados por UV.

Ciclo limiar (C_T) - indica o ciclo em que é determinado o limiar de deteção de fluorescência, ou seja em que a quantidade de amostra começa a aumentar exponencialmente. Quanto menor o C_T maior a quantidade de amostra inicial.

Controlo pareado (*matched*) - indivíduo semelhante a um caso no que concerne a uma panóplia de parâmetros, excetuando a variável em estudo.

Curva de *melting* - permite realizar uma interpretação das características de dissociação do DNA de cadeia dupla, durante um aquecimento progressivo. Permite estimar a presença de dímeros de *primers* ou sequências inespecíficas.

Fase de *annealing* - hibridização dos *primers* à respetiva cadeia simples de DNA.

Gene *housekeeping* - gene necessário para manter as funções básicas de uma célula, expresso em todas as células do organismo, em condições normais ou com presença de patologia.

Genótipo *null* - define-se como a ausência de ambos os alelos do gene, não sendo possível realizar a função do mesmo.

Linfadectomia - procedimento cirúrgico que consiste na remoção completa de gânglio linfático.

Mastectomia - procedimento cirúrgico que consiste na remoção completa da mama.

Missense SNP - polimorfismo, que representa uma variação na sequência de DNA de um só nucleótido, resultando na alteração de um aminoácido distinto.

Polimorfismo genético - variação genética que ocorre em frequência igual ou superior a 1% na população.

Primer - sequência de nucleótidos que serve como ponto de início para síntese de DNA.

Ressuspender - realizar suspensões repetidas com o uso de pipeta.

SNP - polimorfismo, que representa uma variação na sequência de DNA de um só nucleótido.

Temperatura de *melting* - é a temperatura em que a hélice dupla de DNA se dissocia em cadeias simples.

Terapêutica neoadjuvante - administração de agentes terapêuticos anteriormente à terapêutica, dita, principal, com o objetivo de facilitar e aumentar o sucesso desta última.

Termociclador - aparelho utilizado para realizar reações em cadeia de polimerase.

Tubo de *Eppendorf* - também conhecido como tubo de microcentrifugadora, tem fundo cónico e uma tampa unida ao corpo.

Tumorectomia - procedimento cirúrgico que consiste na remoção completa do tumor/carcinoma.

Xenobiótico - substância química presente num organismo, que lhe é estranha, pois não é normalmente produzida por este. Ou quando é produzida pelo organismo, está presente em concentrações muito superiores ao habitual.

Anexos

Anexo 1. Autorização do projeto de investigação



Centro Hospitalar Cova da Beira, E.P.E.

Pract
11/03/2013
[Signature]

G.A. Invest
15 MAR 2013

<p>Parecer:</p>	<p>Despacho:</p> <p style="text-align: center;"><i>CA</i></p> <p style="text-align: center;"><i>11/3/2013</i></p> <p style="text-align: center; font-size: small;">Centro Hospitalar Cova da Beira Presidente do Conselho de Administração Prof. Doutor Miguel Castelo Branco</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="font-size: x-small;">Centro Hospitalar Cova da Beira Presente em reunião de C.A. Em 231.3.2013</td> </tr> <tr> <td style="font-size: x-small;">Despacho <i>[Signature]</i></td> </tr> <tr> <td style="font-size: x-small;">Presidente do C.A. Prof. Doutor Miguel Castelo Branco</td> </tr> <tr> <td style="font-size: x-small;">Vogal do C.A. Prof.ª Dra. Anabela Almeida</td> </tr> <tr> <td style="font-size: x-small;">Vogal do C.A. Técnica Superior Omiinda Sucena</td> </tr> <tr> <td style="font-size: x-small;">Directora Clínica Dra. Rosa Maria Ballesteros</td> </tr> <tr> <td style="font-size: x-small;">Enfermeiro Director Enf.ª António João Rodrigues</td> </tr> </table>	Centro Hospitalar Cova da Beira Presente em reunião de C.A. Em 231.3.2013	Despacho <i>[Signature]</i>	Presidente do C.A. Prof. Doutor Miguel Castelo Branco	Vogal do C.A. Prof.ª Dra. Anabela Almeida	Vogal do C.A. Técnica Superior Omiinda Sucena	Directora Clínica Dra. Rosa Maria Ballesteros	Enfermeiro Director Enf.ª António João Rodrigues
Centro Hospitalar Cova da Beira Presente em reunião de C.A. Em 231.3.2013									
Despacho <i>[Signature]</i>									
Presidente do C.A. Prof. Doutor Miguel Castelo Branco									
Vogal do C.A. Prof.ª Dra. Anabela Almeida									
Vogal do C.A. Técnica Superior Omiinda Sucena									
Directora Clínica Dra. Rosa Maria Ballesteros									
Enfermeiro Director Enf.ª António João Rodrigues									
<p>ASSUNTO: Projecto de Investigação nº22/2013 - "Valor preditivo do polimorfismo aromatase Trp39Arg para metastização ganglionar no cancro da mama"</p>									
<p>PARA: Exmo. Sr. Presidente do Conselho de Administração</p> <p>DE: Gabinete de Investigação e Inovação</p>		<p>N.º 21/GII</p> <p>Data 11/03/2013</p>							
<p>Em relação ao assunto em epígrafe, junto envio o pedido de autorização de João Paulo Baptista Godinho, aluno do Mestrado Integrado em Medicina da Faculdade de Ciências da Saúde da Universidade da Beira Interior, para a realização de um estudo subordinado ao tema "Valor preditivo do polimorfismo aromatase Trp39Arg para metastização ganglionar no cancro da mama", através da consulta de processos clínicos no Departamento de Saúde da Criança e da Mulher – Serviço de Obstetria e Ginecologia e Serviço de Anatomia Patológica.</p> <p>O estudo supra citado está incluído nos trabalhos de investigação "Associação entre genes polimórficos e o risco de cancro da mama na população da Beira Interior" com o parecer favorável nº 24/2008 da Comissão de Ética e "Análise da variação de expressão de proteínas associadas com as vias de síntese e metabolismo dos estrogénios no carcinoma da mama" com o parecer favorável nº 321/2011 da Comissão de Ética.</p> <p>Informo que se encontram reunidos todos os requisitos necessários de acordo com o Regulamento e Normas do Núcleo de Investigação.</p> <p>Com os melhores cumprimentos, <i>psr</i></p> <p style="text-align: center;">? O Gabinete de Investigação e Inovação</p> <p style="text-align: center;"><i>Rosa Saraiva</i></p> <p style="text-align: center;">(D.ª Rosa Saraiva)</p>									

Anexo 2. Localização de *primers* na sequência de cDNA do transcrito ACTB-001 da β -actina

5'...ATGGATGATGATATCGCCGCGCTCGTCGTCGACAACGGCTCCGGCATGTGCAAGGCCGGCTTC
 GCGGGCGACGATGCCCCCGGGCCGTCTTCCCCTCCATCGTGGGGCGCCCCAGGCACCAGGGCGTGA
 TGGTGGGCATGGGTCAGAAGGATTCTATGTGGGCGACGAGGCCAGAGCAAGAGAGGCATCCTCACC
 CTGAAGTACCCCATCGAGCACGGCATCGTCACCAACTGGGACGACATGGAGAAAATCTGGCACCACACC
 TTCTACAATGAGCTGCGTGTGGCTCCCGAGGAGCACCCCGTGTGCTGACCGAGGCCCCCTGAACCCC
 AAGGCCAACCGCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCCAGCCATGTACGTTGC
 TATCCAGGCTGTGCTATCCCTGTACGCTCTGGCCGTACCACTGGCATCGTGTGACTCCGGTGAC
 GGGGTCACCCACACTGTGCCATCTACGAGGGGTATGCCCTCCCCATGCCATCCTGCGTCTGGACC
 TGGCTGGCCGGACCTGACTGACTACCTCATGAAGATCCTCACCGAGCGCGGCTACAGCTTACCAC
 CACGGCCGAGCGGGAATCGTGCGTGACATTAAGGAGAAGCTGTGCTACGTCGCCCTGGACTTCGA
 GCAAGAGATGGCCACGGCTGCTTCCAGCTCCTCCCTGGAGAAGAGCTACGAGCTGCCTGACGGCCA
 GGTCATCACCATTGGCAATGAGCGGTTCCGCTGCCCTGAGGCACTCTCCAGCCTTCTTCTGGGC
ATGGAGTCCTGTGGCATCCACGAACTACCTTCAACTCCATCATGAAGTGTGACGTGGACATCCGCAA
 GACCTGTACGCCAACACAGTGCTGTCTGGCGGCACCACCATGTACCCTGGCATTGCCGACAGGATGCAG
 AAGGAGATCACTGCCCTGGCACCCAGCACAATGAA**GATCAAGATCATTGCTCCTCCT**GAGCGCAAGTA
 CTCCGTGTGGATCGGCGGCTCCATCCTGGCCTCGCTGTCCACCTTCCAGCAGATGTGGATCAGCAAG
 CAGGAGTATGACGAGTCCGGCCCCCTCCATCGTCCACCGCAAATGCTTCTAG...-3'

Apenas se apresenta a sequência codificante. O cDNA deste transcrito tem 6 exões, sendo apenas 5 codificantes. O primeiro não é codificante.

A negrito alternando com normal, estão representados os 5 exões codificantes.

A azul situa-se o *forward primer*. A vermelho situa-se a região complementar inversa do *reverse primer*.

Forward primer 5' -TTCCTGGGCATGGAGTCCT- 3'

Reverse primer 5' -AGGAGGAGCAATGATCTTGATC- 3'

Anexo 3. Tabela com os dados recolhidos no SAM e obtidos em meio laboratorial

# Caso	Idade ao diagnóstico (anos)	Mama afetada	Localização afetada	Tipo Histológico	RP
1	77	Direita	União Quadrantes Externos	Carcinoma Ductal Invasivo	Positivo
3	69	Direita	União Quadrantes Internos	Carcinoma Ductal Invasivo	Positivo
4	41	Direita	Quadrante Supero-externo	Carcinoma Ductal Invasivo	Positivo
7	78	Direita	Quadrante Supero-interno	Carcinoma Ductal Invasivo	Negativo
9	77	Direita	Quadrante Infero-externo	Carcinoma Ductal Invasivo	Positivo
10	72	Esquerda	União Quadrantes Internos	Carcinoma Ductal Invasivo	Positivo
11	62	Direita	Quadrante Supero-externo	Carcinoma Lobular Invasivo	Positivo
12	34	Esquerda	União quadrantes Internos	Carcinoma Ductal Invasivo	Positivo
13	80	Esquerda	Quadrante Supero-interno	Carcinoma Ductal Invasivo	Positivo
14	70	Esquerda	União quadrantes Internos	Carcinoma Ductal Invasivo	Negativo
15	78	Esquerda	Quadrante Infero-externo	Carcinoma Ductal Invasivo	Negativo
16	73	Esquerda	União Quadrantes Internos	Carcinoma Mucinoso Invasivo	Negativo
17	36	Direita	União Quadrantes externos	Carcinoma Ductal Invasivo	Negativo
18	83	Esquerda	Região retro-areolar	Carcinoma Ductal Invasivo	Positivo
19	83	Direita	União Quadrantes inferiores	Carcinoma Lobular Invasivo	Positivo
22	57	Direita	União quadrantes internos	Carcinoma Ductal Invasivo	Negativo
23	50	Esquerda	Quadrante Supero-externo	Carcinoma Ductal Invasivo	Positivo
24	69	Esquerda	Quadrante Supero-externo	Carcinoma Ductal Invasivo	Negativo
25	57	Esquerda	Região peri-areolar/União Quadrantes externos	Carcinoma Ductal Invasivo	Positivo
26	86	Esquerda	Toda mama	Carcinoma Lobular Invasivo	Positivo
27	39	Direita	Quadrante Supero-externo	Carcinoma Ductal Invasivo	Positivo
28	46	Direita	Quadrantes internos	Carcinoma Ductal Invasivo	Negativo
29	56	Direita	Quadrante Supero-externo	Carcinoma Ductal Invasivo	Negativo
30	59	Esquerda	Quadrante Supero-externo	Carcinoma Ductal Invasivo	Negativo
31	78	Direita	Região Peri-areolar	Carcinoma Ductal Invasivo	Positivo
33	72	Esquerda	Quadrante Infero-externo	Carcinoma Ductal Invasivo	Negativo
34	66	Direita	União Quadrantes inferiores	Carcinoma Ductal Invasivo	Positivo
35	58	Esquerda	Quadrante Supero-externo	Carcinoma Ductal Invasivo	Positivo
36	70	Esquerda	Quadrante Infero-externo	Carcinoma Ductal Invasivo	Negativo
37	69	Direita	União Quadrantes Internos	Carcinoma Ductal Invasivo	Positivo

RP- Recetores de progesterona

# Caso	RE	C-er-B2+	Triplo negativo	Luminal A	Focalidade	Tamanho (mm)	Grau de diferenciação (Classificação de Bloom)	Invasão linfocascular
1	Positivo	Negativo	Não	Sim	Unifocal	40	G2	Não foi observado
3	Positivo	Negativo	Não	Sim	Unifocal	23	G2	Não foi observado
4	Positivo	Negativo	Não	Sim	Unifocal	22	G2	Não foi observado
7	Positivo	Negativo	Não	Sim	Unifocal	9	G1	Não foi observado
9	Positivo	Negativo	Não	Sim	Multifocal	13	G1	Não foi observado
10	Positivo	Negativo	Não	Sim	Unifocal	6	G1	Não foi observado
11	Positivo	Negativo	Não	Sim	Unifocal	15	G2	Não foi observado
12	Positivo	Negativo	Não	Sim	Unifocal	12	G2	Não foi observado
13	Positivo	Negativo	Não	Sim	Unifocal	25	G2	Não foi observado
14	Negativo	Negativo	Sim	Não	Unifocal	90	G3	Sim
15	Positivo	Negativo	Não	Sim	Unifocal	35	G2	Não foi observado
16	Positivo	Negativo	Não	Sim	Unifocal	28	G1	Não foi observado
17	Positivo	Negativo	Não	Sim	Unifocal	8	G1	Não foi observado
18	Positivo	Negativo	Não	Sim	Unifocal	42	G2	Não foi observado
19	Positivo	Negativo	Não	Sim	Multifocal	32	G1	Não foi observado
22	Negativo	Negativo	Sim	Não	Unifocal	15	G2	Não foi observado
23	Positivo	Negativo	Não	Sim	Unifocal	11	G1	Não foi observado
24	Negativo	Positivo	Não	Não	Unifocal	35	G2	Não foi observado
25	Positivo	Negativo	Não	Sim	Unifocal	1,5	G1	Não foi observado
26	Positivo	Negativo	Não	Sim	Unifocal	70	G1	Não foi observado
27	Positivo	Negativo	Não	Sim	Multifocal	25	G2	Não foi observado
28	Negativo	Negativo	Sim	Não	Unifocal	18	G2	Sim
29	Positivo	Negativo	Não	Sim	Unifocal	15	G3	Não foi observado
30	Negativo	Positivo	Não	Não	Unifocal	15	G3	Não foi observado
31	Positivo	Positivo	Não	Não	Unifocal	16	G1	Não foi observado
33	Negativo	Positivo	Não	Não	Unifocal	14	G2	Não foi observado
34	Negativo	Negativo	Não	Sim	Unifocal	19	G3	Não foi observado
35	Positivo	Negativo	Não	Sim	Unifocal	38	G2	Não foi observado
36	Positivo	Negativo	Não	Sim	Unifocal	19	G2	Não foi observado
37	Positivo	Negativo	Não	Sim	Unifocal	23	G2	Não foi observado

RE- Recetores de estrogénio

# Caso	Envolvimento da pele e/ou musculo	Carcinoma <i>in-situ</i>	Calcificação	Nº gânglios positivos	Axila/Gânglio Sentinela	Estadio	Estadio T
1	Não	Não	Não	2	Com metastização	T2N2a	T2
3	Sim	Não	Não	0	Sem metastização	pT2N0	T4
4	Não	Sim	Não	1	Com metastização	pT2N1a	T2
7	Não	Não	Não	0	Sem metastização	pT1bN0	T1
9	Não	Sim	Não	0	Sem metastização	pT1cN0	T1
10	Não	Sim	Não	0	Sem metastização	T1bN0	T1
11	Não	Não	Não	0	Sem metastização	T1cN0	T1
12	Não	Não	Não	0	Sem metastização	T1cN0	T1
13	Não	Não	Não	0	Sem metastização	pT2N0	T2
14	Sim	Não	Não	2	Com metastização	pT3N1a	T4
15	Não	Não	Não	0	Sem metastização	pT2N0	T2
16	Não	Não	Não	0	Sem metastização	T2N0	T2
17	Não	Não	Sim	0	Sem metastização	T1bN0	T1
18	Não	Sim	Não	2	Com metastização	T3N1a	T3
19	Não	Sim	Não	12	Com metastização	pT2N3a	T2
22	Não	Não	Não	0	Sem metastização	pT1cN0	T1
23	Não	Sim	Sim	1	Com metastização	pT1cN1a	T1
24	Não	Não	Não	0	Sem metastização	pT2N0	T2
25	Não	Sim	Sim	0	Sem metastização	pT1aN0	T1
26	Sim	Não	Não	0	Sem metastização	pT3N0	T4
27	Não	Sim	Sim	0	Sem metastização	pT2N0	T2
28	Não	Sim	Não	0	Sem metastização	pT1cN0	T1
29	Não	Sim	Não	2	Com metastização	T1N1a	T1
30	Não	Sim	Não	0	Sem metastização	T1cN0	T1
31	Sim	Não	Não	0	Sem metastização	T1cN0	T4
33	Não	Sim	Não	3	Com metastização	pT1cN1a	T1
34	Não	Sim	Sim	0	Sem metastização	pT1cN0	T1
35	Não	Sim	Não	0	Sem metastização	T2N0	T2
36	Não	Sim	Não	0	Sem metastização	pT1cN0	T1
37	Sim	Sim	Não	1	Com metastização	pT4bN1a	T4

# Caso	Estadio N	Estadio clínico	Expressão de <i>CYP19A1</i> Trp39Arg	ΔC_T de <i>CYP19A1</i> Trp39Arg	Expressão de <i>GSTM1</i>	Expressão de <i>GSTT1</i>
1	N2	IIIA	Sim	-14	Não	Não
3	N0	IIIB	Não	N/A	Não	Não
4	N1	IIB	Não	N/A	Não	Não
7	N0	I	Não	N/A	Não	Não
9	N0	I	Não	N/A	Não	Não
10	N0	I	Não	N/A	Não	Não
11	N0	I	Não	N/A	Não	Não
12	N0	I	Não	N/A	Não	Não
13	N0	IIA	Sim	-9,58	Não	Não
14	N1	IIIB	Não	N/A	Não	Não
15	N0	IIA	Sim	-11,28	Não	Não
16	N0	IIA	Sim	-7,04	Não	Não
17	N0	I	Sim	-7,61	Não	Não
18	N1	IIIA	Não	N/A	Não	Não
19	N3	IIIC	Sim	-10,8	Não	Não
22	N0	I	Sim	-5,51	Não	Não
23	N1	IIA	Sim	-7,52	Não	Não
24	N0	IIA	Sim	-7,17	Não	Não
25	N0	I	Não	N/A	Não	Não
26	N0	IIIB	Não	N/A	Não	Não
27	N0	IIA	Não	N/A	Não	Não
28	N0	I	Não	N/A	Não	Não
29	N1	IIA	Não	N/A	Não	Não
30	N0	I	Não	N/A	Não	Não
31	N0	IIIB	Não	N/A	Não	Não
33	N1	IIA	Não	N/A	Não	Não
34	N0	I	Não	N/A	Não	Não
35	N0	IIA	Não	N/A	Não	Não
36	N0	I	Sim	-11,33	Não	Não
37	N1	IIIB	Sim	-8,72	Não	Não