

**UNIVERSIDADE DE LISBOA**

**FACULDADE DE CIÊNCIAS**

**DEPARTAMENTO DE BIOLOGIA ANIMAL**



**ESTUDO DE *LINKAGE* E PESQUISA DE  
MUTAÇÕES NUMA FAMÍLIA PORTUGUESA  
COM DOENÇA DE BEHÇET E ESPONDILITE  
ANQUILOSANTE**

**JOANA GONÇALVES DE GOUVEIA MAIA XAVIER**

**MESTRADO EM BIOLOGIA HUMANA E AMBIENTE**

**2008**

**UNIVERSIDADE DE LISBOA**

**FACULDADE DE CIÊNCIAS**

**DEPARTAMENTO DE BIOLOGIA ANIMAL**



**ESTUDO DE *LINKAGE* E PESQUISA DE  
MUTAÇÕES NUMA FAMÍLIA PORTUGUESA  
COM DOENÇA DE BEHÇET E ESPONDILITE  
ANQUILOSANTE**

**JOANA GONÇALVES DE GOUVEIA MAIA XAVIER**

DISSERTAÇÃO ORIENTADA POR:

DOUTORA SOFIA OLIVEIRA

PROFESSORA DOUTORA DEODÁLIA DIAS

**MESTRADO EM BIOLOGIA HUMANA E AMBIENTE**

**2008**

## **Agradecimentos**

Primeiro que tudo queria agradecer à minha orientadora, Sofia Oliveira, por me ter dado a oportunidade de trabalhar no grupo de Genética Humana do Instituto Gulbenkian de Ciência neste projecto tão interessante no qual aprendi imenso e pude desenvolver o meu mestrado.

Também gostava de agradecer ao Dr. José Vaz Patto e ao Dr. Filipe Barcelos, por terem “descoberto” este precioso objecto de estudo e pela sua disponibilidade e atenção para esclarecer todas as dúvidas clínicas.

Um grande obrigada a todos os meus colegas de grupo, actuais e passados, Alexandra, Benedita, Madalena, Sara e Tiago por todo o apoio que me deram durante este último ano, pelo companheirismo e boa-disposição. Sem vocês esta experiência não teria sido tão boa.

Ao João Garcia, o informático de serviço, que tornou possível a realização das tão complicadas análises de linkage.

A todos os meus amigos, mesmo os que estão no estrangeiro: Emília, Érica, Nádía, Lu, Rita, Romina, Sampas, Su e Vânia, entre outros, obrigado pelo apoio, amizade, pelos dias e noites bem passados e por terem tornado este último ano, assim como os da faculdade, espectacular e cheio de boas recordações.

Ao Pedro por tudo.

Por último à minha família que sempre me acompanhou e guiou, e que me apoia incondicionalmente em todas as minhas decisões.

# Índice Geral

Agradecimentos	i
Índice Geral	ii
Abreviaturas e Símbolos	v
Resumo	vii
Abstract	viii
<b>1. INTRODUÇÃO</b>	<b>1</b>
1.1 Perspectiva histórica	1
1.2 Patologia clínica	1
1.3 Epidemiologia	6
1.4 Factores histopatológicos e imunológicos	7
1.5 Factores genéticos	8
1.5.1 Agregação familiar	8
1.5.2 Complexo MHC, Espondilartropatias e Behçet	9
1.5.3 Estudos de <i>linkage</i>	12
1.6 Factores ambientais	13
1.7 Polimorfismos nucleotídicos simples (SNPs)	14
1.8 <i>N-Acetiltransferase 2 (NAT2)</i>	15
1.8.1 Função da proteína codificada	15
1.8.2 Gene <i>NAT2</i>	15
1.8.3 Haplótipos	16
1.8.4 Fenótipos	16
1.8.5 Estudos de associação	18
<b>2. OBJECTIVOS</b>	<b>19</b>

<b>3. METODOLOGIA</b>	<b>21</b>
3.1 Família estudada e exame clínico	21
3.2 Extração de DNA	23
3.3 Genotipagem do <i>locus</i> B do sistema HLA	23
3.4 Estudo de <i>linkage</i>	24
3.4.1 Genotipagem com o <i>GeneChip</i> de 250K da Affymetrix	25
3.4.2 Modelos fenotípicos	26
3.4.3. Análise estatística	26
3.5 Análise da perda de heterozigotia (LOH)	29
3.6 Análise dos genótipos na região 8p21.3-22	30
3.7 Pesquisa de mutações	30
3.7.1 Escolha de <i>primers</i>	30
3.7.2 Amplificação dos genes candidatos	31
3.7.3 Sequenciação e análise dos produtos de PCR obtidos	33
3.8 Construção dos haplótipos para o gene <i>NAT2</i>	34
<b>4. RESULTADOS</b>	<b>35</b>
4.1 Agregação familiar da DB, EA e EpAI	35
4.2 Associação dos alelos HLA-B com a DB, EA e EpAI	36
4.3 Estudo de <i>linkage</i>	37
4.3.1 Genotipagem com o <i>genechip</i> de 250K da Affymetrix	37
4.3.2 Modelos fenotípicos	38
4.3.3 Análise estatística	39
4.4 Análise da perda de heterozigotia (LOH)	42
4.5 Análise dos genótipos na região 8p21.3-22	42

4.6 Pesquisa de mutações	43
4.6.1 Genes candidatos	43
4.6.2 Sequenciação dos genes candidatos	44
4.7 Haplótipos do gene <i>NAT2</i>	46
4.8 Interação entre o gene <i>NAT2</i> e o sistema HLA-B	47
<b>5. DISCUSSÃO</b>	<b>49</b>
5.1 Agregação familiar da DB, EA e EpAI	49
5.2 Influência não linear dos alelos do sistema HLA-B	49
5.3 Estudo de <i>linkage</i>	50
5.3.1 Limitações do programa usado para a análise	50
5.3.2 Ausência de <i>linkage</i> com o sistema MHC	51
5.3.3 Evidência de <i>linkage</i> na região 8p21.3-22	51
5.3.4 Diferentes modelos estatísticos aplicados	52
5.4 Inexistência de LOH	53
5.5 Influência do <i>NAT2</i> e do HLA-B na susceptibilidade à DB e às EpA	54
<b>6. CONCLUSÃO</b>	<b>58</b>
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>60</b>
<b>Anexos</b>	
Anexo A – Declaração da comissão de ética do IPR	71
Anexo B – Informação adicional sobre a análise no Merlin	73
Anexo C – Árvores genealógicas dos modelos fenotípicos	75
Anexo D – Gráficos de <i>linkage</i>	79
Anexo E – Imagens da sequenciação	91

## Abreviaturas e Símbolos

AO – Aftas **o**rais

AG – Aftas **g**enitais

CNAT – Ferramenta de análise da variação do número de cópias (*Copy Number Analysis Tool*)

cM – Centimorgan

CNV – Variação no número de cópias (*Copy Number Variation*)

CHGN – Factor polimerizante da condroitina

DB – Doença de **B**ehçet

DNA – Ácido desoxirribonucleico (*Deoxyribonucleic Acid*)

dNTPs – Trifosfatos de desoxirribonucleótidos

EA – Espondilite **a**nquilosante

EpA – **E**spondilartropatias

EpAI – **E**spondilartropatias **i**ndiferenciadas

EDTA – Ácido Etilenodiaminotetracético

ESSG – Grupo Europeu para o Estudo das Espondilartropatias (*European Spondyloarthropathies Study Group*)

GTTYPE – Software de análise dos *chips* de genotipagem (*GeneChip Genotyping Analysis Software*)

HLA – Antígeno Leucocitário Humano (*Human Leucocyte Antigen*)

HSP – Proteínas de choque térmico (*Heat Shock Proteins*)

IBD – Idênticos por descendência (*Identical by Descendent*)

IgA – **I**munoglobulina **A**

IGC – Instituto **G**ulbenkian de **C**iência

IPR – Instituto **P**ortuguês de **R**eumatologia

Kb – Quilobase

LC – Lesões **c**utâneas

LO – Lesões **o**culares

LOD – Razão de probabilidades (*Logarithm of Odds*)

LOH – Perda de heterozigotia (*Loss of Heterozygosity*)

m- – mili ( $10^{-3}$ )  
Mb – Megabase  
MHC – Sistema Maior de Histocompatibilidade (*Major Histocompatibility Complex*)  
RNA – Ácido ribonucleico (*Ribonucleic Acid*)  
mRNA – RNA mensageiro  
MSN – **M**anifestações do **s**istema **n**ervoso  
n- – nano ( $10^{-9}$ )  
NAT1 – Gene codificante da **N**-acetiltransferase **1**  
NAT2 – Gene codificante da **N**-acetiltransferase **2**  
NAT1 – Isoenzima N-acetiltransferase 1  
NAT2 – Isoenzima N-acetiltransferase 2  
NPL – *Linkage* não paramétrico (*Non Parametric Linkage*)  
NK – Células *natural killer*  
NKT – Células *T natural killer*  
pb – **p**ares de **b**ases  
PCR – Reacção de polimerização em cadeia (*Polymerase Chain Reaction*)  
ROS – Espécies Reactivas de Oxigénio (*Reactive Oxygen Species*)  
rpm – **R**otações **p**or **m**inuto  
S<sub>all</sub> – Estatística *all* de Whittmore e Halpern  
SD – *Stepdown*  
SNP – Polimorfismos nucleotídicos simples (*Single Nucleotide Polymorphisms*)  
S<sub>pairs</sub> – Estatística *pairs* de Whittmore e Halpern  
SSP – Primers específicos para a sequência (**S**equência **S**pecific **P**rimers)  
TAE – Tampão Tris EDTA  
TGF-β – *Transforming Growth Factor beta*  
T<sup>o</sup>m – Temperatura de desnaturação (*melting*)  
TNF – *Tumor Necrosis Factor*  
μ- – micro ( $10^{-6}$ )  
°C – Graus Celsius

## Resumo

A doença de Behçet (DB) é uma desordem inflamatória multissistémica que pode afectar as membranas mucosas, a pele, as articulações e o sistema gastrointestinal e nervoso central. As Espondilartropatias (EpA) são um grupo de doenças reumáticas inflamatórias às quais pertencem a espondilite anquilosante (EA) e as Espondilartropatias indiferenciadas (EpAI). A EA afecta sobretudo o esqueleto axial, com maior severidade nas junções sacroilíacas e coluna, apresentando também manifestações extra-articulares.

Trata-se de doenças complexas, com contribuição genética e ambiental. Os únicos factores genéticos consistentemente associados a estas patologias são os alelos HLA-B\*51 e B\*27, que explicam 19% e 20-40% da susceptibilidade à DB e à EA, respectivamente. A frequente coexistência destas doenças, a sobreposição de características clínicas, a sua co-agregação na família Portuguesa em estudo e o facto de serem doenças inflamatórias crónicas, leva a considerar que a DB e a EA poderão partilhar factores genéticos comuns e a investigá-los.

Realizou-se um estudo de *linkage* (ligação genética) no genoma total, seguido de pesquisa de mutações, numa família com múltiplos casos de DB, EA e EpAI. Realizaram-se análises de *linkage* não paramétricas com o programa Merlin e genotipagem do HLA-B em todos os indivíduos. Uma análise conjunta da DB e EA mostrou uma forte evidência de *linkage* no cromossoma 8p21.3-22 (LOD=3.59,  $p=2.0 \times 10^{-5}$ ). Uma análise manual desta região, mapeou o *locus* da doença a uma região menor, albergando 24 genes, entre eles o *N-acetiltransferase 2 (NAT2)*. Verificou-se que nesta família os indivíduos com DB ou EpA são HLA-B\*27 ou B\*51 positivos e possuem um fenótipo de acetilação lenta promovido pelo *NAT2*. Este gene codifica uma importante enzima metabolizadora de xenobióticos, podendo os xenobióticos não acetilados induzir mecanismos auto-imunes. Este estudo sugere que o estatuto de acetilação lenta em combinação com os alelos B\*51 ou B\*27 pode ser determinante na susceptibilidade para a DB e EpA.

**Palavras-chave:** Doença de Behçet; Espondilartropatias; agregação familiar; estudo de *linkage* no genoma total; *N-acetyltransferase 2*; HLA-B

## Abstract

Behçet's disease (BD) is a multi-system inflammatory condition that may involve the skin, mucocutaneous membrane, eyes, joint, lung and gastrointestinal and central nervous system. Spondyloarthropathies (SpA) are a group of related inflammatory rheumatic diseases that includes Ankylosing spondylitis (AS) and undifferentiated spondyloarthropathies (USpA) among others. AS mainly affects the axial skeleton, with more severity in spinal and sacroiliac joints, and less frequently extra-articular systems. These are multifactorial disorders, with genetic and environmental factors implicated in their etiology. The only genetic factors that have been consistently associated worldwide with these diseases are the HLA-B\*51 that explains up to 19% of the BD risk, and HLA-B\*27 that has an attributable risk of 20-40% for AS.

The previously reported co-occurrence of BD and AS in some patients, the clinical overlap between these two diseases, the fact that both are chronic and inflammatory disorders, and their aggregation in the Portuguese family under study, led us to hypothesize that BD and AS may share common genetic factors and to search them.

We performed a whole genome linkage screen followed by mutation screening in a multiplex family with BD, AS and UspA patients. Non-parametric linkage analysis was performed using Merlin. HLA-B was typed in all individuals. A joint linkage analysis for BD and AS found strong evidence for linkage (LOD=3.59,  $p=2.0 \times 10^{-5}$ ) on chromosome 8p21.3-22. A fine manual analysis confined the disease *locus* to a small region containing 24 genes, among them the *N-Acetyltransferase 2* (NAT2). We found that in this family, individuals with BD or SpA are carriers of HLA-B\*27 or HLA-B\*51 alleles and have a NAT2 slow acetylator phenotype. NAT2 is an important xenobiotic-metabolizing enzyme and nonacetylated xenobiotics may induce autoimmune mechanisms. This study suggests that NAT2 acetylator status in combination with HLA-B\*51 or B\*27 alleles may be a determinant in susceptibility to BD and SpA.

**Keywords:** Behçet's disease; Spondyloarthropathies; familial aggregation; whole genome linkage screen; *N-acetyltransferase 2*, HLA-B.

# 1. INTRODUÇÃO

## 1.1 Perspectiva histórica

A doença de Behçet (DB) foi descrita pela primeira vez em 1931 por um oftalmologista grego, Adamantiades, que descreveu um caso de hipópion (acumulação de fluído no olho) acompanhado de ulcerações orogenitais, flebite e hidrartrose bilateral dos joelhos. Em 1937, Hulusi Behçet, um dermatologista turco, teve o mérito de insistir na individualidade clínica desta doença, postulando uma etiologia viral e salientando as suas manifestações mais frequentes: aftose oral, aftose genital e uveíte<sup>1</sup>.

No que respeita à espondilite anquilosante (EA), estudos paleopatológicos de múmias Egípcias sugerem que esta doença tem afectado a humanidade desde a antiguidade. No entanto, a primeira descrição histórica só aparece na literatura em 1559 quando Realdo Colombo fez uma descrição anatómica de dois esqueletos com anomalias típicas da EA no seu livro *De Re Anatomica*<sup>2</sup>.

## 1.2 Patologia clínica

A doença de Behçet é uma condição inflamatória crónica e multissistémica, que pode afectar a maioria dos órgãos do corpo através de uma vasculite generalizada das veias e das artérias. Para além dos sintomas mais frequentes, que são as aftas orais e genitais, também podem existir manifestações oculares (uveíte, vasculite retiniana), cutâneas, articulares, gastrointestinais, pulmonares, do sistema nervoso central e vasculares, entre outras<sup>3</sup>. A doença é caracterizada por crises recorrentes de duração variável, com períodos de remissão ou exacerbação da doença, podendo ter várias recorrências durante um ano<sup>4</sup>. Sendo uma doença crónica, a DB pode levar a incapacidades funcionais temporárias ou permanentes, podendo também causar desordens psiquiátricas<sup>5</sup>. Nestes pacientes estima-se que haverá uma taxa de mortalidade de 9.8%, maioritariamente devido a doenças vasculares graves (com envolvimento de vasos de grande porte), envolvimento neurológico e úlceras intestinais perfurantes<sup>3</sup>. No entanto, a cegueira, que chega a afectar 25% dos doentes, é a principal causa de incapacidade permanente. Apesar da etiologia da DB permanecer desconhecida, o seu mecanismo mais provável é o de uma resposta inflamatória excessiva

desencadeada pela exposição a um agente infeccioso (bacteriano ou viral) ou ambiental (químicos, metais ou outros), num hospedeiro geneticamente susceptível<sup>4</sup>.

As Espondilartropatias (EpA) constituem um grupo heterogéneo de doenças reumáticas crónicas e inflamatórias que partilham características genéticas e clínicas comuns. As EpA incluem a espondilite anquilosante, a artrite reactiva, a artrite psoriática, a artrite associada com a doença inflamatória dos intestinos, e as espondilartropatias indiferenciadas (EpAI)<sup>6</sup>. Já foram introduzidas classificações ou critérios de diagnóstico para várias das doenças incluídas no grupo das EpAs, no entanto, nem sempre é possível diferenciar claramente as várias formas de EpA, especialmente nos seus estados iniciais<sup>7</sup> e porque podem ocorrer simultaneamente ou sequencialmente no mesmo paciente. A sua apresentação e evolução clínica parecem ser também (para além dos factores genéticos) influenciadas pela etnia, idade de início das manifestações e sexo<sup>8,9</sup>.

A espondilite anquilosante é o protótipo das EpAs, e o seu subtipo mais comum e mais severo<sup>2</sup>. Pode afectar todas as articulações, embora com maior incidência no esqueleto axial (coluna e articulações sacroilíacas) e nas inserções tendinosas (enteses), mas também pode apresentar manifestações extra-esqueléticas, tais como manifestações oculares (uveíte anterior), cardíacas (insuficiência aórtica), gastrointestinais e ainda lesões pulmonares<sup>10</sup>. É uma entesopatia inflamatória que leva à ossificação e anquilose. As características principais desta doença são a rigidez e a dor infamatória axial; de início geralmente numa forma insidiosa, tornando-se num quadro crónico, que pode também apresentar períodos de remissão e exacerbação dos sintomas<sup>11-13</sup>. Na EA uma das complicações mais graves é a fractura da coluna vertebral. No entanto, os problemas extra-esqueléticos tais como as lesões cardiovasculares, problemas gastrointestinais, circulatórios e cerebrovasculares, podem ocorrer em pacientes com doença de longo termo provocando um aumento de 50% no risco de mortalidade relativamente à população em geral<sup>14,15</sup>. Esta doença também tem um forte impacto na qualidade de vida e bem-estar psicológico dos pacientes que dela padecem, e, como surge numa idade jovem, possui um forte impacto económico para o paciente e a sociedade<sup>16</sup>. A etiologia da EA também permanece por elucidar, no entanto, sabe-se que está dependente de um contexto genético muito forte, assim como de um contexto ambiental<sup>13</sup>.

As EpAI são o segundo tipo mais comum de EpA e incluem as formas que não preenchem os critérios para as doenças inseridas neste complexo<sup>6,17</sup>. Delas fazem parte as

fases iniciais e incompletas de EpA definitivas e também formas que permanecem indiferenciadas por um longo período de tempo. Uma elevada percentagem das EpAI evoluem para outras formas de EpA num período de 10 anos após o início das manifestações, sendo um dos maiores riscos atribuídos a esta doença, a evolução para outro subtipo mais severo, como é o caso da EA<sup>16,18</sup>. O espectro clínico das EpAI é vasto e resulta da combinação de várias manifestações clínicas e radiológicas das EpAs, como a artrite periférica, entesite periférica, dactilite, dor torácica, raquialgia inflamatória, sacroileíte, conjuntivite, uveíte aguda anterior e insuficiência aórtica<sup>6</sup>.

Para a DB, EA e EpAI não existem testes laboratoriais específicos, sendo o diagnóstico baseado em critérios clínicos com recurso a meios complementares de diagnóstico<sup>4,19</sup>. Os critérios de diagnóstico geralmente aceites para estas patologias encontram-se descritos nas **Tabelas 1 – 4**. Embora o estabelecimento de critérios permita um melhor diagnóstico, manifestações incompletas das doenças ou manifestações pouco comuns, como frequentemente acontece em doentes com Behçet, muitas vezes representam um desafio no diagnóstico<sup>4</sup>. Além disso, como entre o aparecimento dos primeiros sintomas e o diagnóstico destas doenças podem passar vários anos, muitas vezes o tratamento é tardio, o que pode levar à progressão destas doenças e à diminuição da qualidade de vida dos pacientes<sup>14</sup>.

**Tabela 1.** Critérios do Grupo de Estudo Internacional para a Doença de Behçet (1990)<sup>20</sup>

Necessário:
<b>Ulceração oral recorrente</b> - Aftas <i>minor</i> , aftas <i>major</i> ou ulceração herpetiforme observada pelo médico ou paciente, ocorrendo pelo menos 3 vezes num ano.
Mais dois dos seguintes sintomas:
<b>Ulceração genital recorrente</b> - Úlceras aftosas ou em forma de cicatriz, observadas pelo médico ou paciente.
<b>Lesões oculares</b> - Uveíte anterior ou posterior, ou células no vítreo. - Vasculite da retina observada pelo oftalmologista.
<b>Lesões na pele</b> - Eritema nodoso, pseudo-foliculite ou lesões pápulo-pústulares observadas pelo paciente ou médico. - Nódulos acneiformes observados pelo médico em pacientes pós-adolescentes sem tratamento com corticósteroides.
<b>Teste de patergia positivo</b> - Observado pelo médico às 24-48h.

**Tabela 2.** Critérios de estudo internacionais para a doença de Behçet revistos<sup>21</sup>

<b>Características clínicas</b> (necessário perfazer 3 ou mais pontos):	<b>Pontos</b>
Aftose oral	1
Manifestações cutâneas	1
Lesões vasculares (trombose venosa e arterial, aneurisma)	1
Fenómeno patérgico (teste)	1
Aftose genital	2
Lesões oculares	2

**Tabela 3.** Critérios de diagnóstico modificados de Nova York (1984) para a Espondilite Anquilosante<sup>22</sup>

<b>Critérios Clínicos</b> (necessário pelo menos um): - Presença ou história de dor e rigidez na transição dorsolombar ou lombar, que melhora com o exercício mas não com o repouso. - Limitação da mobilidade lombar nos planos anterior, posterior e laterais. - Redução da expansibilidade torácica, medida no quarto espaço intercostal, relativamente aos valores normais para a idade e sexo.
<b>Critérios Radiológicos</b> (necessário pelo menos um): - Sacroileíte bilateral de grau 2 a 4. - Sacroileíte unilateral de grau 3 ou 4.

**Tabela 4.** Critérios do Grupo de Estudo Europeu para as Espondilartropatias (ESSG) para a classificação das Espondilartropatias<sup>23</sup>

**Necessário:**

- Raquialgia inflamatória, ou Sinovite assimétrica (predominando nos membros inferiores)

**Mais um dos seguintes achados:**

- Dor na nádega alternando de um lado para o outro
- Entesite
- Psoríase
- Sacroileíte radiológica
- Diarreia aguda, uretrite ou cervicite precedendo a artrite
- Doença inflamatória do intestino (IBD)
- História familiar positiva de EpA

O tratamento da DB continua predominantemente empírico e dirigido às manifestações clínicas de cada paciente, sendo baseado na administração de corticosteróides, colchicina e imunomoduladores. A abordagem terapêutica é, regra geral, multidisciplinar, necessitando de uma colaboração próxima entre especialistas de medicina oral, dermatologia, oftalmologia e outros<sup>3</sup>. Os pacientes masculinos, assim como aqueles em que a doença está no seu início, necessitam normalmente de um tratamento mais agressivo do que restantes indivíduos afectados. Embora o prognóstico tenha melhorado com tratamentos imunossuppressores e imunomoduladores precoces e agressivos, o tratamento permanece, no entanto, insatisfatório devido à raridade e heterogeneidade da doença, incertezas acerca da etiologia e patogénese, e à falta de dados de ensaios clínicos bem controlados<sup>3</sup>.

Na EA existem terapêuticas recentes mais eficazes. Todavia, o seu sucesso está dependente dum diagnóstico precoce da doença que muitas vezes não é realizado (atraso de 5-10 anos entre o aparecimento dos primeiros sintomas e o diagnóstico)<sup>2</sup>. Estas terapêuticas incluem tanto as abordagens mais tradicionais (terapia física, exercício, educação do paciente e prescrição de medicamentos anti-inflamatórios de origem não esteróide), como os eficazes medicamentos antagonistas do TNF (*tumor necrosis factor*) que suprimem a actividade da doença e melhoram a capacidade física dos pacientes<sup>19,24,25</sup>.

Embora tenha sido sugerido, em 1974, que a doença de Behçet pudesse ser incluída no complexo das EpA, este é um tópico bastante controverso. A favor desta proposta está o

facto da DB partilhar características clínicas com algumas doenças do complexo EpA, tais como, a EA, o síndrome de Reiter e a doença inflamatória dos intestinos (IBD)<sup>26</sup>. Também têm sido reportados em alguns estudos populacionais e individuais um aumento da prevalência da sacroileíte e de EA em pacientes com DB<sup>26-29</sup>, tendo tanto o alelo HLA-B\*27, como HLA-B\*51, sido reportados como mais frequentes nestes pacientes. Todavia, outros autores constataram que a prevalência de EA ou SI na DB não é diferente da que se verifica na população em geral<sup>30-32</sup>. Até hoje não foram, no entanto, reportados casos de agregação familiar da DB e EA.

### 1.3. Epidemiologia

A DB tem uma distribuição global, embora seja mais prevalente ao longo da antiga Rota da Seda, que se estende desde o Mediterrâneo até ao Japão, incluindo a área do Golfo<sup>3,4,33</sup>. Este padrão de distribuição sugere a possível existência de um factor genético que se difundiu via migração de tribos nómadas ancestrais<sup>3</sup>. Na Turquia, o país com maior incidência da doença, esta ocorre em 0,1-0,4% da população<sup>4</sup>, e em Portugal a sua incidência estimada é de 2.5 em 100,000 indivíduos, uma das mais elevadas da Europa Ocidental<sup>1</sup>. É, no entanto, uma doença muito rara nos países da Europa do Norte e da América do Norte. Notavelmente, a distribuição da DB sobrepõe-se com a prevalência do alelo HLA-B\*51 do complexo maior de histocompatibilidade (MHC). A DB surge normalmente na terceira ou quarta década de vida, afectando ambos os sexos, com uma maior incidência masculina na área mediterrânica e ocidental, e feminina nas populações Asiáticas<sup>34</sup>.

As EpA possuem uma prevalência global de 1.9%<sup>16</sup>, representando o segundo grupo mais comum de doenças reumáticas inflamatórias crónicas, juntamente com a artrite reumatóide<sup>35</sup>. São caracterizadas por uma forte associação com o alelo HLA-B\*27, que é no entanto variável entre populações. Afectam pessoas jovens, geralmente antes de atingirem os 30 anos, e ambos os sexos de igual modo, excepto para a EA que afecta três vezes mais homens do que mulheres<sup>7</sup>.

A EA ocorre em todo o mundo com uma taxa de incidência na população de 0.5 – 8.2 por 100 000 indivíduos e uma prevalência global de 0.2 – 1.2%. A sua prevalência é,

contudo, muito variável em diferentes populações e fortemente relacionada com a distribuição do alelo HLA-B\*27<sup>2</sup>. Em Portugal, a sua prevalência não está bem estudada mas estima-se que seja de 2 a 3 casos em cada 1000 indivíduos<sup>13</sup>. No Médio Oriente é pouco frequente (0,07%), assim como no Japão (0,007%)<sup>29</sup>, sendo nestes países também menor a associação com o alelo HLA-B\*27<sup>2</sup>. Os primeiros sintomas de EA surgem habitualmente na adolescência ou nos jovens adultos<sup>2</sup>.

A DB e a EA são pouco frequentes em crianças, ocorrendo apenas em 2-3% dos casos com DB e em 4% dos casos com EA. No entanto, nas crianças os sintomas evoluem normalmente para manifestações mais severas, tal como acontece nos indivíduos do sexo masculino. Nas crianças afecta ambos os sexos igualmente, e é mais frequente uma história familiar da doença do que em adultos<sup>3,4,16,34</sup>.

A prevalência das EpAI varia entre diferentes populações, embora esta não esteja tão bem estudada como outras doenças do complexo EpA. As EpAI podem manifestar-se em qualquer grupo etário, seja na infância, adolescência, idade adulta ou meia-idade, sendo o espectro clínico semelhante em todos eles. É tão comum em homens como mulheres, embora nos homens as manifestações possam ser mais severas. Tem-se referido que os traumas físicos podem despoletar o início das manifestações periféricas em pacientes com diferentes formas de EpA, incluindo as EpAI<sup>6</sup>.

#### **1.4. Factores histopatológicos e inflamatórios**

A lesão histopatológica comum, existente por detrás das manifestações clínicas da DB, é uma vasculite sistémica (inflamação dos vasos sanguíneos), que afecta vasos de diferentes tipos, tamanhos e localização, com infiltração de neutrófilos, dilatação das células endoteliais e necrose dos tecidos circundantes<sup>36</sup>. As principais características imunológicas desta doença consistem num aumento da actividade dos neutrófilos, com consequente aumento da quimiotaxia, fagocitose e produção de espécies reactivas de oxigénio (ROS), incluindo o anião superóxido (que pode ser responsável pela danificação dos tecidos). Também ocorre aumento da resposta dos linfócitos B e T às proteínas de choque térmico (HSP) e alterações dos níveis de citocinas. As interrelações entre estas respostas não são, no

entanto, claras<sup>37</sup>, não existindo, até à data, modelos *in vitro* ou *in vivo* da imunopatogénese da DB<sup>3</sup>.

Os locais mais comuns de inflamação na EA incluem as articulações sacroilíacas, as enteses, corpos vertebrais adjacentes aos discos intervertebrais, sinóvia das articulações periféricas, sistema gastrointestinal e olho. Muitas destas lesões são de difícil acesso pelo que existe pouca informação acerca da histopatologia. A infiltração de linfócitos B e T e de macrófagos é acompanhada de proliferação de fibroblastos e neovascularização e existe aumento da expressão de mRNA de TNF- $\alpha$  e TGF- $\beta$  (*transforming growth factor beta*). Também ocorre aumento dos níveis de IgA no soro e de proteínas de fase aguda, assim como aumento da expressão de moléculas de adesão e de factores de quimiotaxia. O osso destruído é eventualmente substituído e ocorre ossificação endocondral que resulta na anquilose (perda de mobilidade) do osso<sup>12</sup>.

## **1.5. Factores genéticos**

### **1.5.1. Agregação familiar**

Embora a maioria dos casos de DB e de doenças do complexo EpA sejam esporádicos e os progenitores não sejam afectados, a agregação familiar é uma característica epidemiológica de ambas as doenças<sup>24,38</sup>.

Na DB a agregação é pouco frequente nas populações caucasianas, mas uma história positiva pode ser observada em 12% dos pacientes não caucasianos<sup>4</sup>. Nestes casos a doença é caracterizada por uma antecipação genética, um fenómeno epidemiológico de base molecular, caracterizado pela manifestação mais precoce da doença ou pelo aumento da sua severidade, em gerações sucessivas<sup>39,40</sup>. O único mecanismo conhecido para a antecipação genética em certas doenças é uma classe de mutações que contêm sequências repetitivas e patogénicas de trinucleótidos, que se expandem e aumentam de tamanho após a transmissão<sup>41</sup>. O motivo da sequência que é repetido e o tamanho e localização das repetições em expansão dentro dos genes responsáveis pela doença, variam consideravelmente entre doenças e influenciam o modo como a repetição causa a doença<sup>42</sup>.

Uma elevada prevalência de manifestações isoladas da doença, tais como úlceras

orogenitais recorrentes ou um teste positivo para a patergia cutânea, também foram observados entre os parentes em primeiro grau dos pacientes com DB<sup>3,34,40</sup>. Estudos em famílias sugerem contudo que esta doença não é monogénica, pois não possui um padrão de hereditariedade Mendeliano simples<sup>3</sup>.

Relativamente à EA, existe um aumento significativo da sua ocorrência, assim como de outras EpA, nos parentes de indivíduos afectados pela doença em relação à sua ocorrência na população em geral<sup>13,38</sup>. De Blecourt *et al.*<sup>43</sup>, verificou que a espondilite era 22.6 vezes mais frequente em parentes dos pacientes espondilíacos do que em parentes de controlos. Este risco de recorrência associado às EpA diminui sucessivamente em parentes mais afastados dos pacientes e de um modo mais rápido do que o que acontece em doenças monogénicas. Este padrão é consistente com o de doenças poligenéticas, onde uma interacção génica multiplicativa determina a susceptibilidade<sup>7</sup>. O número de genes envolvidos na susceptibilidade à EA é desconhecido, no entanto, modelos de risco resultantes de estudos em famílias indicam que está envolvido um número limitado de genes<sup>44</sup>, possivelmente entre 3 e 9<sup>45</sup>, e que esta doença poderá seguir um padrão de hereditariedade autossómico dominante, com baixa penetrância<sup>46</sup>.

### **1.5.2. Complexo MHC, espondilartropatias e Behçet**

O *locus* genético mais amplamente estudado em ambas as doenças tem sido o sistema HLA (*Human Leucocyte Antigen*), pertencente ao complexo maior de histocompatibilidade (MHC), que se situa no cromossoma 6p21. Este complexo é composto por três regiões (classe I, II e III) que cobrem 4 megabases e codificam as moléculas HLA, altamente polimórficas, que são responsáveis pela apresentação de antígenos às células T. As moléculas classe I estão principalmente envolvidas na apresentação de antígenos virais ou endógenos às células T CD8<sup>+</sup>. Outra função destas moléculas foi descoberta mais recentemente, através da identificação de uma nova família de receptores expressos maioritariamente por células NKT (*natural killer T cells*), e também pelas células T CD8<sup>+</sup> e células T $\gamma\delta$ . Estes receptores, os KIR (killer immunoglobulin-like receptors) ligam-se a epitopos conservados, que são partilhados por diferentes grupos alélicos de moléculas HLA

classe I, e podem levar à inibição selectiva de células NK (natural killer) ou da citotoxicidade mediada por células T<sup>47</sup>.

O alelo HLA-B\*51 é o marcador genético mais consistentemente associado à DB, pois encontra-se presente em 60% dos indivíduos afectados (contra 12-20% na população em geral)<sup>33</sup> e explica 19% da susceptibilidade à doença<sup>34</sup>. Não se sabe se este alelo está implicado directamente na patogénese da DB, ou se está associado com a doença apenas porque se encontra em desequilíbrio de *linkage* com um gene próximo que confere susceptibilidade à DB<sup>4</sup>. O alelo HLA-B\*51 tem pelo menos 34 variantes genéticas, no entanto a associação tem sido restrita aos seus subtipos moleculares mais comuns, HLA-B\*5101 e B\*5108<sup>4</sup>. Têm sido sugeridos diferentes mecanismos de acção da molécula HLA-B\*51, incluindo mimetismo molecular com antigénios da retina ou então actuar como um factor que leva à activação dos neutrófilos<sup>47</sup>. A favor da associação directa com a doença está o facto de estudos em ratinhos transgénicos para o HLA-B\*51 terem mostrado um aumento da produção do ião superóxido, devido ao aumento da actividade dos neutrófilos. No entanto, estes animais não expressavam o fenótipo da DB. Tendo em conta a principal função das moléculas classe I, a falta das características clínicas da DB neste modelo pode ser explicada pela ausência de um péptido patogénico restrito ao HLA-B\*51 que activaria as células T CD8<sup>+</sup> relevantes para a doença<sup>48</sup>.

De um modo semelhante, também as EpA se encontram notoriamente associadas a um alelo deste sistema, o HLA-B\*27. No entanto, o grau de associação varia consideravelmente entre as várias formas de EpA, assim como em diferentes grupos étnicos. As EpAI parecem estar bastante associadas ao B\*27, havendo estudos que indicam que 70-84% dos doentes são portadores deste alelo<sup>49-51</sup> (contra 7-8% na população em geral)<sup>12</sup>. No entanto, também existem estudos a reportar prevalências muito baixas deste alelo nas EpAI<sup>7</sup>.

A EA é a doença que dentro do grupo das EpA se encontra mais fortemente associada, sendo mais de 95% dos pacientes portadores deste alelo. A contribuição global do HLA-B\*27 para a susceptibilidade à EA é substancial (20-40%) e este parece estar directamente implicado na sua patogénese<sup>6</sup>. Na EA também existem diferenças na associação das variantes alélicas do HLA-B\*27: os subtipos HLA-B\*2705, HLA-B\*2704 e B\*2702 parecem estar mais fortemente associados à doença, enquanto os subtipos HLA-B\*2706 e HLA-B\*2703 não estão associados<sup>52</sup>.

A favor da acção directa do HLA-B\*27 na susceptibilidade às EpA está a ocorrência de artrite periférica e axial, inflamação dos intestinos e lesões genitais e na pele de ratos HLA-B\*27 transgénicos. No entanto, o mecanismo exacto pelo qual a molécula B\*27 leva à indução de uma inflamação crónica é ainda alvo de debate<sup>53</sup>. O mimetismo molecular entre antigénios de bactérias e a molécula HLA-B\*27 tem sido invocado em alguns estudos como o meio pelo qual as respostas imunitárias patológicas e auto-reactivas se iniciam, enquanto outros defendem a apresentação de péptidos artritogénicos que se ligam especificamente a esta molécula como o mecanismo pelo qual esta resposta é iniciada. Existem várias substituições de aminoácidos concentradas em torno do local de ligação ao antigénio que geram as várias variantes alélicas da molécula HLA-B\*27. Deste modo, estas podem influenciar a susceptibilidade para a EA, quer através da alteração dos epitopos dos antigénios da molécula HLA-B\*27, quer pela alteração da gama de péptidos ao qual essas variantes da molécula HLA-B\*27 se podem ligar<sup>52</sup>.

Embora a maioria dos estudos epidemiológicos apenas refira a associação dos alelos B\*51 e B\*27 com a DB e EpA, respectivamente, também têm sido encontradas associações mais fracas de outros alelos do sistema HLA-B com estas doenças. Na DB já foi referida uma associação em certas populações com o alelo B\*2702<sup>47</sup>, com o B\*57<sup>54</sup> e com o B\*15, tendo este último também sido associado em estudos familiares<sup>55,56</sup>. Estes resultados são interessantes pois todos estes alelos expressam o motivo Bw4 (epitopo serológico existente nas moléculas HLA-B e HLA-A que é especificado por cinco aminoácidos variáveis entre as posições 77 e 83)<sup>57</sup> que poderá ser o responsável pela susceptibilidade à doença<sup>36</sup>. A EA, também já foi associada ao alelo HLA-B\*60 em estudos populacionais<sup>52,58</sup> e ao HLA-B7 em estudos familiares. As EpA já foram associadas ao alelo HLA-B\*15<sup>59</sup>.

Apesar da forte associação do HLA-B\*51 e do HLA-B\*27 com a DB e a EA, respectivamente, o risco relativo para as doenças associado à presença destes alelos varia consideravelmente em diferentes populações étnicas e verifica-se que estes alelos estão presentes com elevada frequência em algumas populações nas quais as doenças são virtualmente inexistentes. Estas patologias devem, por isso, estar relacionadas com o efeito de mais genes, ainda por identificar, que definem a expressão fenotípica específica em conjunto com os factores ambientais. Estudos familiares na EA indicam que os indivíduos positivos para o alelo HLA-B\*27, que possuem parentes em primeiro grau com esta doença,

têm um risco 6 a 16 vezes maior de virem a desenvolver a patologia do que os indivíduos HLA-B\*27 positivos sem história familiar. Estes dados reforçam a ideia de que, nestas doenças, genes adicionais, dentro e/ou fora do sistema MHC<sup>32</sup>, modificam a expressão dos indivíduos B\*27 e B\*51 positivos<sup>7</sup>.

### 1.5.3. Estudos de *linkage*

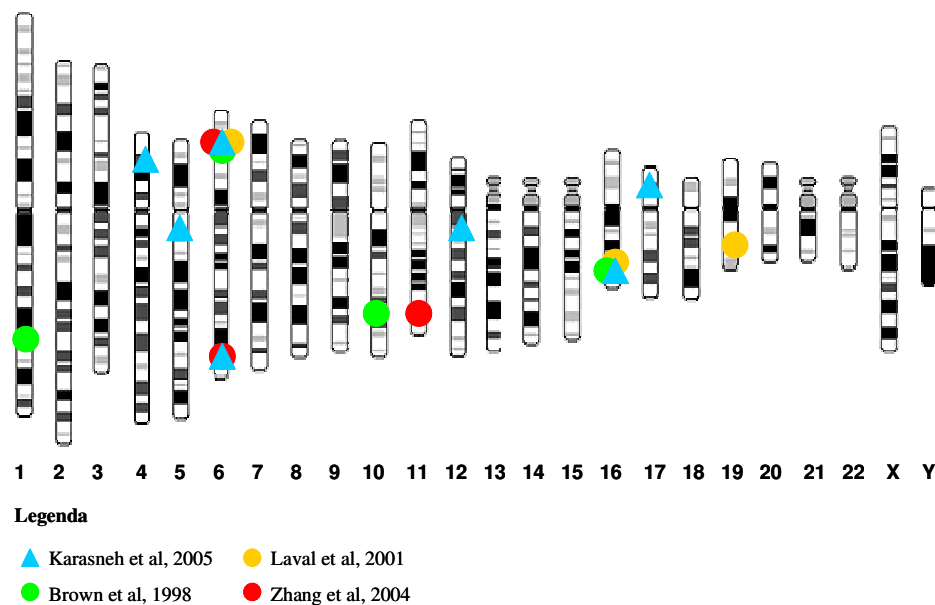
O *linkage* genético é um método pelo qual se estimam distâncias genéticas entre *loci*<sup>60</sup>. Este método é muitas vezes a primeira etapa na investigação genética de uma doença, pois uma das suas principais aplicações é a identificação de vastas regiões genómicas que contenham um ou mais genes responsáveis pela doença, mesmo na ausência de hipóteses biológicas previamente definidas<sup>61</sup>. São procuradas as áreas do genoma que indivíduos afectados por uma doença, dentro da mesma família, partilham mais frequentemente do que o que seria esperado se a segregação fosse independente, utilizando-se marcadores distribuídos pelo genoma. Esta é uma indicação que nessas regiões existem genes que causam, ou aumentam, o risco de se desenvolver uma doença<sup>62</sup>.

Os resultados dos estudos de *linkage* são normalmente reportados como um logaritmo de razões (odds), o LOD *score*, determinado para cada marcador estudado. *Scores* positivos elevados são uma evidência a favor de *linkage* (ou cosegregação do marcador) com a doença, enquanto *scores* negativos são uma evidência contra. Para calcular os valores de LOD *score* podem-se realizar análises de *linkage* paramétricas, em que se especifica um modelo de hereditariedade da doença, indicando-se parâmetros como a frequência do alelo responsável pela mesma, o seu modo de hereditariedade (dominante ou recessivo) e a sua penetrância, ou pode realizar-se uma análise não paramétrica, em que nenhum destes parâmetros é indicado<sup>61</sup>.

Em pedigrees raros, com hereditariedade mendeliana, a identificação dos genes envolvidos e de mutações raras com elevada penetrância, através de estudos de *linkage*, tem conduzido a um melhor conhecimento dos mecanismos patogénicos das formas mais comuns de certas doenças<sup>63</sup>.

Para a EA, já foram publicados três estudos de *linkage* no genoma inteiro por Brown *et al.*<sup>64</sup>, Laval *et al.*<sup>65</sup> e Zhang *et al.*<sup>66</sup> e feitas duas meta-análises desses estudos<sup>67,68</sup>. Para a

DB, apenas foi publicado um estudo de *linkage* no genoma total por Karasneh<sup>69</sup>. Em todos estes estudos, os marcadores utilizados foram microsátélites. Os resultados obtidos nestes estudos, para as análises *multipoint* (ver ponto 3.4.3 dos métodos), encontram-se sumarizados na **Fig 1**. Embora tenham sido obtidos resultados bastante diferentes em diferentes estudos, obtiveram-se picos de *linkage* para a DB e EA nos cromossomas 6p22-24 (onde se situa o MHC), 6q25-26 e 16q21-23, reforçando a ideia de que devem existir outros factores genéticos comuns a estas patologias, para além do HLA-B.



**Figura 1.** Representação esquemática no cariótipo humano dos picos de *linkage* obtidos num estudo de *linkage* para a DB (representados por triângulos) e em três estudos para a EA (representados por círculos).

## 1.6. Factores ambientais

Embora seja claro que existe uma componente genética significativa e forte nestas doenças, os factores ambientais também desempenham um papel importante<sup>3,12,13</sup>. Estudos em populações migrantes assim como em irmãos gémeos têm permitido descobertas epidemiológicas importantes e apoiado esta ideia. Na DB observou-se que indivíduos turcos emigrados na Alemanha possuíam um risco de doença significativamente menor do que indivíduos turcos a viver na Turquia, embora o seu risco se mantivesse mais elevado do que o da população alemã nativa<sup>4</sup>. O facto de não haver concordância total no desenvolvimento

da DB, assim como na EA, em gémeos monozigóticos (taxa de concordância de 75% para ambas as doenças) também aponta para uma influência ambiental<sup>16,70</sup>. Na EA foi mesmo calculado um risco genético total de apenas 16-50%<sup>7,71</sup>. O agente ambiental responsável pela reacção inflamatória destas doenças pode ser um antigénio externo, um antigénio com relacionamento genético, ou mesmo um auto-induzido. Este último não exclui uma infecção activadora, que pode operar através de mimetismo molecular, ou de outro mecanismo, mas implica que a doença seja perpetuada por uma resposta imune anormal a um autoantigénio, na ausência de uma infecção decorrente<sup>4,12</sup>.

Na DB, existem vários dados que suportam a hipótese de que possa ter um agente activador infeccioso e têm sido investigadas evidências de infecções prévias ou actuais com uma grande variedade de agentes virais e bacterianos. Contudo, o agente despoletador, assim como o modo como esta activação induz a doença não é ainda conhecido<sup>3,4</sup>. Não existe, no entanto, qualquer dúvida de que existe uma resposta inflamatória a vários auto-antigénios na DB<sup>4</sup>.

O agente exterior que activa a susceptibilidade está provado ser uma bactéria para as espondilites que ocorrem integradas numa artrite reactiva (Espondilite do Síndrome de Reiter, por ex.). Para as EA primárias e para outras EpA, a causa que activa o processo permanece por determinar<sup>13</sup>. Infecções entéricas com *Klebsiella pneumoniae* e *Escherichia coli* têm sido implicadas na patogénese da EA, em hospedeiros geneticamente susceptíveis. Adicionalmente, uma associação próxima entre a EA e a doença inflamatória dos intestinos sugere que a resposta imunitária gastrointestinal à flora normal com consequente inflamação pode ser o responsável pelo despoletar desta doença<sup>12,16</sup>.

### **1.7 Polimorfismos nucleotídicos simples (SNPs)**

Os polimorfismos nucleotídicos simples (SNPs) são o tipo mais comum de polimorfismos nas moléculas de DNA e ocorrem quando é alterado um único nucleótido na sua sequência. Como os SNPs ocorrem abundantemente e espaçadamente pelo genoma humano, oferecem um potencial muito elevado para serem usados como biomarcadores para o diagnóstico de doenças genéticas (como em estudos de *linkage*), comparativamente com outros polimorfismos menos comuns como os microsátélites<sup>72,73</sup>.

## 1.8 *N-Acetiltransferase 2 (NAT2)*

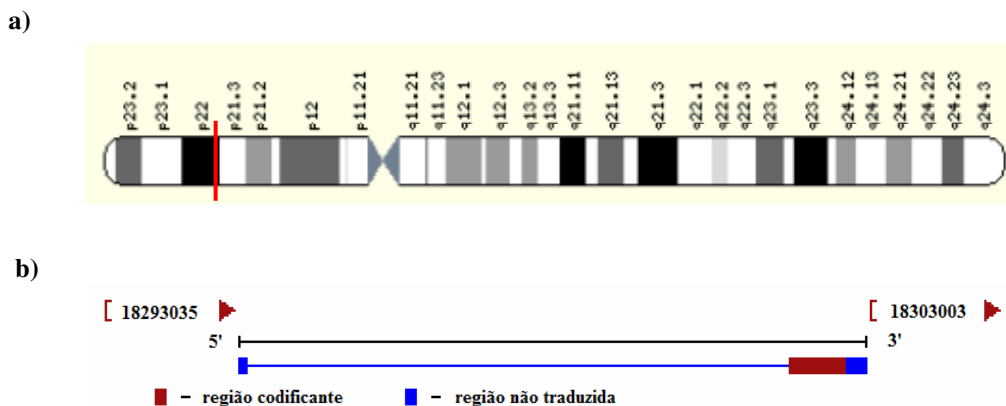
### 1.8.1 Função da proteína codificada

Muitas enzimas metabolizadoras de drogas catalizam a ativação e desintoxicação de xenobióticos (compostos químicos estranhos a um organismo ou sistema biológico), sendo classificadas de enzimas de fase I e II, respectivamente. Na categoria de fase I estão incluídas enzimas relacionadas com o citocromo P450 assim como também hidrosilases epóxidas. As enzimas de fase II são as N-acetiltransferases (NATs), glutathione S-transferases, uridino difosfato-glucoronosiltransferases e sulfotransferases<sup>37</sup>.

As NAT de arilaminas humanas existem sob a forma de duas isoenzimas, NAT1 e NAT2, com especificidades diferentes e complementares<sup>74</sup>. Estas enzimas catalisam a transferência de um grupo acetil da coenzima A para as posições N- e O- de uma grande variedade de xenobióticos que servem como substratos de arilaminas<sup>75</sup>. Para além deste papel no metabolismo xenobiótico, pensa-se que as NATs possuem um papel no metabolismo endógeno<sup>76</sup>. A enzima NAT2 possui maior expressão no fígado, intestino delgado e cólon mas também está presente na maioria dos restantes tecidos<sup>77</sup>.

### 1.8.2 Gene NAT2

A NAT2 é codificada pelo gene *NAT2* que está localizado no braço pequeno do cromossoma 8 (8p22), e contém uma região codificadora, sem intrões, de 870 pares de bases (pb)<sup>78</sup> (**Fig 2**). Têm sido descritos polimorfismos genéticos no *NAT2* que causam variações na capacidade de acetilação da isoenzima NAT2. Estes polimorfismos são causados por alterações num único nucleótido no haplótipo *wild type NAT2\*4*<sup>79</sup>. Entre os sete polimorfismos que são mais frequentemente encontrados na população humana, quatro resultam na substituição dum aminoácido (substituições num único par de bases nas posições G191A, T341C, G590A e G857A), que conduz a uma diminuição significativa na capacidade de acetilação da enzima. Os outros três, ou são substituições silenciosas (C282T, C481T) ou substituições não sinónimas mas que não alteram o fenótipo (A803G)<sup>80</sup>.



**Figura 2.** Gene *NAT2*. a) Representação da localização cromossômica (a vermelho) no cromossoma 8. b) Estrutura dos exões e representação das regiões codificantes e não codificantes.

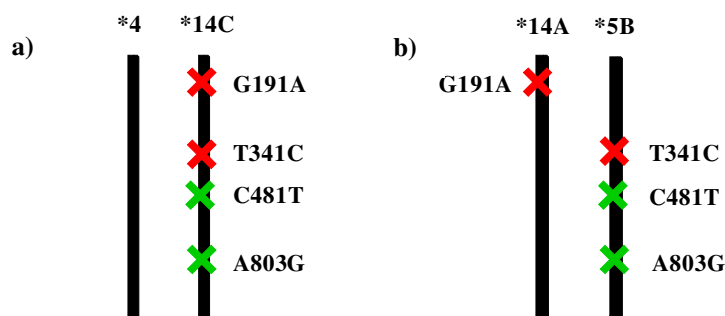
### 1.8.3 Haplótipos

O conjunto dos SNPs localizados ao longo da região codificante do *NAT2* forma haplótipos, pois toda esta região é segregada em conjunto para codificar a enzima *NAT2*. Cada combinação de SNPs identificada até hoje constitui um haplótipo distinto, tendo sido já registados 52 haplótipos diferentes<sup>81</sup>. Existem evidências crescentes de que para genes que contêm múltiplos SNPs em desequilíbrio de *linkage*, tal como o *NAT2*, a estrutura do haplótipo, e não os SNPs individualmente, pode ser o principal determinante das consequências fenotípicas. Deste modo, é muito importante a consideração e estudo dos haplótipos, já que um polipéptido funcional é o resultado de um haplótipo que cobre totalmente a região codificante e que é codificado por um único cromossoma<sup>80</sup>. Devido ao desequilíbrio de *linkage* que existe entre os SNPs do gene *NAT2* na população, a identificação dos polimorfismos nas posições 481, 590, 803 e 857 é suficiente para determinar os haplótipos, já que os restantes polimorfismos estão ligados aos mencionados em 98% dos casos<sup>79</sup>.

### 1.8.4 Fenótipos

O haplótipo *wild type NAT2\*4* (sem substituições) e os haplótipos que possuem substituições silenciosas (sem alteração de aminoácido) ou substituições não sinónimas mas que não levam a alteração da função, tais como o *NAT2\*12* e *NAT2\*13*, codificam proteínas

plenamente funcionais, que promovem acetilação rápida do substrato. Por oposição, os haplótipos com substituições não sinónimas levam à diminuição da função da proteína, que promove uma acetilação lenta. Nestes incluem-se o *NAT2*\*5, *NAT2*\*6 e *NAT*\*7<sup>79</sup>. Os indivíduos que possuem dois haplótipos *NAT2* de acetilação lenta são fenotipicamente acetiladores lentos, os que têm dois haplótipos de acetilação rápida têm um fenótipo de acetilação rápida e os que têm um haplótipo de acetilação rápida e um de acetilação lenta são fenotipicamente intermédios para a velocidade de acetilação<sup>80</sup>. Antigamente, não se fazia a distinção entre os acetiladores rápidos e intermédios, classificando-se ambos os tipos como rápidos, havendo alguns estudos actuais que ainda optam por esta denominação<sup>79,82</sup>. Como ilustrado na **Fig 3**, indivíduos com o mesmo genótipo podem ser acetiladores lentos ou intermédios dependendo se essas substituições estão localizadas no mesmo ou em cromossomas diferentes. É necessário existirem substituições que levem à diminuição na capacidade de acetilação, em ambos os cromossomas, para um indivíduo ser acetilador lento. Estas substituições nucleotídicas levam a reduções na afinidade para o substrato, e da actividade catalítica, e/ou a uma instabilidade proteica da aloenzima N-acetiltransferase recombinante<sup>83</sup>, podendo um metabolismo de acetilação alterado levar a sérias reacções adversas às drogas, como neurite periférica, febre e toxicidade hepática. A frequência do fenótipo de acetilação lenta é baixa na população Asiática (5-30%), mas elevada nas populações Africana e Caucasiana (40-90%)<sup>85</sup>. As relações fenótipo/genótipo no *NAT2* têm sido intensivamente estudadas, no entanto ainda não são completamente percebidas<sup>84</sup>.



**Figura 3.** Representação esquemática dos haplótipos no gene *NAT2*. a) Indivíduo com um fenótipo de acetilação intermédio *NAT2*\*4/*NAT2*\*14C b) Indivíduo com um fenótipo de acetilação lento *NAT2*\*14A/*NAT2*\*5B. O símbolo (X) representa uma substituição que leva a uma diminuição da actividade da enzima *NAT2* enquanto o símbolo (X) indica substituições sem impacto no fenótipo de acetilação.

### 1.8.5 Estudos de associação

O fenótipo de N-acetilação rápida parece estar associado a várias doenças, tais como, o cancro do cólon, carcinoma gástrico, cancro dos pulmões e dermatite de contacto, enquanto o fenótipo de acetilação lenta já foi associado em diferentes estudos a doenças alérgicas e asma extrínseca em pacientes com características atópicas<sup>82</sup>. Os polimorfismos na acetilação do *NAT2* também têm sido associados à susceptibilidade a algumas doenças auto-imunes e inflamatórias, tais como lúpus eritematosus sistémico (SLE)<sup>85-87</sup>, artrite reumatóide<sup>88</sup> e escleroderma<sup>89</sup>, e com a resposta à medicação e tratamento em pacientes com artrite reumatóide<sup>90</sup> ou com doença inflamatória dos intestinos<sup>91</sup>. Em Behçet também já foram realizados três estudos em que se procurou relacionar polimorfismos do *NAT2* com susceptibilidade para esta doença. Tamer *et al.*<sup>37</sup> concluiu que o fenótipo de acetilação lenta poderia ser determinante na susceptibilidade para a DB, Aynacioglu *et al.*<sup>92</sup> não encontrou uma associação significativa embora o alelo *NAT2\*5B* estivesse sobre-representado nos pacientes com DB e Najim *et al.*<sup>93</sup> verificou que, para além dos pacientes com DB terem tendência para ter acetilação lenta, também possuíam um fenótipo de acetilação único (nulo) derivado de uma interacção com o HLA-B\*51.

O gene *NAT2* também já foi associado com a formação de cataratas relacionadas com a idade, tendo-se observado que os acetiladores lentos possuíam maior risco de desenvolver esta complicação do que os que possuíam fenótipo de acetilação rápida<sup>94,95</sup>. Estudos em ratinhos demonstraram defeitos oculares relacionados com uma mutação nula no gene *NAT2*, pois este é expresso durante a embriogénese, desde a fase de blastocisto e no tubo neural em desenvolvimento e olho<sup>96</sup>.

## 2. OBJECTIVOS

Encontram-se descritos na literatura científica vários casos de coexistência no mesmo indivíduo das doenças de Behçet e espondilite anquilosante. As duas doenças são caracterizadas por uma inflamação crónica, derivada de uma resposta imunitária anormal. Também se verifica que sintomas como a uveíte e a artrite, entre outros, podem fazer parte do quadro clínico de ambas. Estas observações, juntamente com o facto destas doenças pouco frequentes aparecerem agregadas numa família portuguesa com múltiplos casos de DB, EA, EpAI e ainda de manifestações isoladas características da DB, leva-nos a colocar a hipótese, de que estas doenças, possam partilhar factores genéticos nesta família.

O objectivo deste estudo é a identificação de factores genéticos comuns entre a doença de Behçet e a espondilite anquilosante, através da realização de um estudo de *linkage* seguido de pesquisa de mutações, na família portuguesa em estudo. Pretendeu-se também estudar a associação nesta família dos alelos HLA-B\*51 e HLA-B\*27 à DB e às EpA, respectivamente. A descoberta de um factor genético comum entre DB e EA poderá contribuir para um melhor conhecimento acerca destas doenças, ao revelar vias celulares envolvidas na patogénese de ambas. Este conhecimento também poderá ajudar a elucidar a relação que existe entre a DB e a EA, nomeadamente a sua agregação familiar e coexistência individual.

Para concretizar este objectivo, subdividiu-se o trabalho nas seguintes etapas:

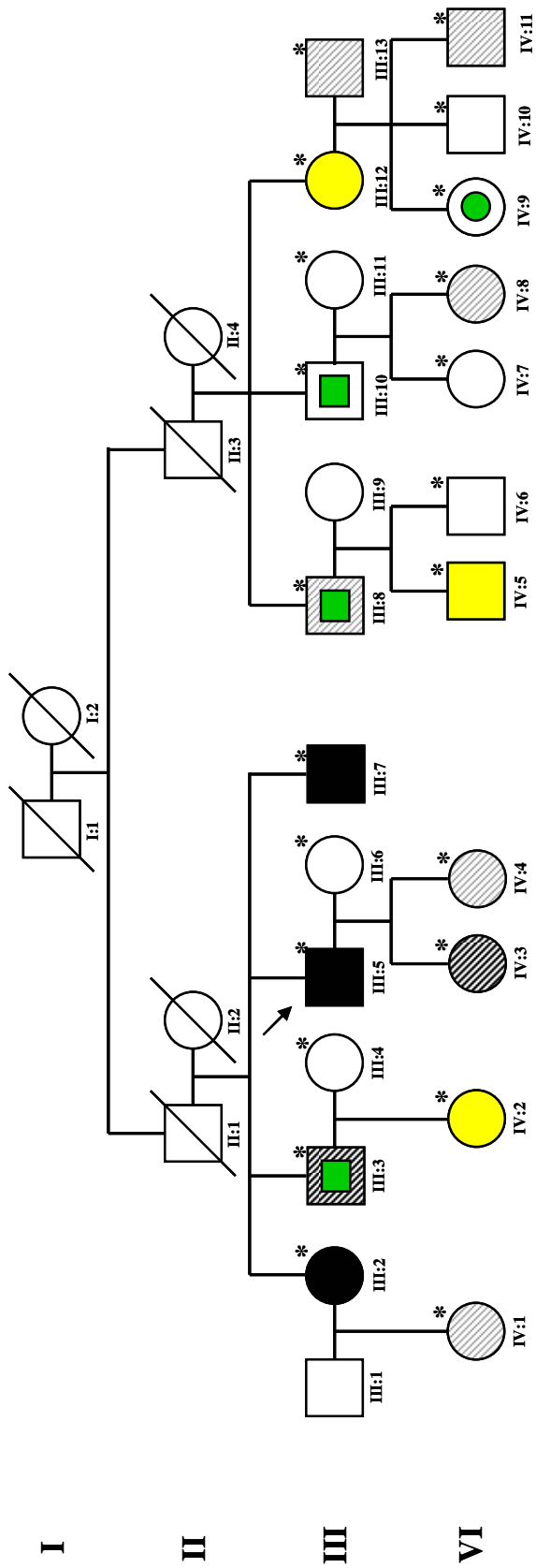
1. Colheita de amostras e dados, no Instituto Português de Reumatologia, de todos os indivíduos da família que aceitaram participar neste estudo (afectados e não afectados);
2. Genotipagem do *locus* B do sistema HLA, em todos os indivíduos participantes, da família;
3. Realização do estudo de *linkage* (ligação genética) para a DB e EA e também características isoladas de DB, com microarrays de 238,000 polimorfismos e da respectiva análise estatística;

4. Selecção de genes candidatos, situados nos picos com maior evidência de *linkage*, de acordo com as suas funções biológicas;
5. Sequenciação de todos os exões dos genes candidatos em indivíduos representativos da família (afectados e não afectados), de modo a encontrar possíveis mutações;
6. Confirmação para todos os elementos da família participantes da existência de segregação do factor genético com as doenças.

## 3.0 METODOLOGIA

### 3.1. Família estudada e examinação clínica

O estudo foi realizado numa família portuguesa, com múltiplos casos de DB, EA e EpAI, que é seguida no Instituto Português de Reumatologia (IPR) pelo Dr. José Vaz Patto. A examinação clínica foi efectuada em todos os indivíduos da família por médicos especialistas em reumatologia, o Dr. José Vaz Patto e o Dr. Filipe Barcelos. Este estudo foi aprovado pela comissão de ética do IPR (**Anexo A**), tendo todos os participantes recebido uma folha de informação sobre o mesmo e assinado uma declaração de consentimento informado. Foram recolhidas amostras e dados de 22 elementos da família, assim como alguma informação sobre os seus antepassados. A informação clínica dos participantes foi obtida na altura da colheita de amostras, através do preenchimento de um formulário desenvolvido especificamente para este estudo, e guardada numa base de dados BCI|Gene segura. A árvore genealógica da família encontra-se representada na **Fig 4**.



**Legenda:**

- Não afectados
- Doença de Behçet
- ◐ Um sintoma de Behçet
- ◑ Dois sintomas de Behçet
- ◓ Espondilite Anquilosante
- ◔ Espondiloartropatia indiferenciada
- \* Indivíduos em que se colheu amostras de sangue

**Figura 4.** Árvore genealógica da família com respectivos diagnósticos clínicos

### 3.2 Extração de DNA

Foram colhidos dois tubos de sangue (EDTA 7.5ml, Sarsted) por participante. Um dos tubos foi usado para extração de DNA utilizando o *kit* NucleoSpin Blood XL (Macherey-Nagel). Através deste método, o DNA genómico é preparado a partir do sangue total e a lise celular adquirida por incubação do mesmo, numa solução contendo elevadas quantidades de iões destabilizantes na presença de proteínase K. São criadas condições apropriadas para a ligação do DNA à membrana de sílica das colunas do Kit, através da adição de etanol ao lisado. A ligação à membrana é reversível e específica para ácidos nucleicos, e são realizados diversos passos de lavagem para remoção eficiente de contaminações. O DNA genómico puro é depois eluído sob condições de baixa força iónica num tampão de eluição ligeiramente alcalino. Todas as amostras de DNA obtidas foram armazenadas a 4°C e conservado com clorofórmio. O outro tubo de sangue colhido foi armazenado a -80°C, para o caso de ser necessária a sua posterior utilização. Nos procedimentos de manuseamento e de armazenamento das amostras foram utilizados códigos de barras específicos para cada amostra, a fim de se reduzirem possíveis erros de registo.

O DNA genómico foi quantificado usando o Nanodrop. A sua pureza foi inferida através da razão A260/230 e a sua integridade verificada através da razão A260/A280 e confirmada num gel de agarose.

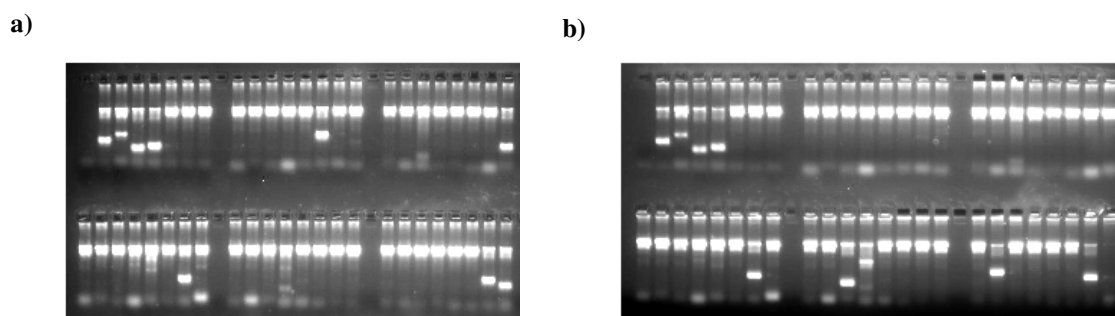
### 3.3 Genotipagem do *locus* B do sistema HLA

Procedeu-se à genotipagem do *locus* HLA-B nesta família de modo a verificar-se se os alelos HLA-B\*51 e HLA-B\*27 se encontram associados à DB e às EpA, respectivamente, e para estudar as suas interacções com outros *loci* identificados no estudo de *linkage*.

A genotipagem foi realizada com o *kit* Micro SSP™ HLA Class I B Locus Specific (One lambda), de acordo com o protocolo descrito pelo fabricante. A mistura de reacção foi adicionada a uma placa de 48 poços com *primers* liofilizados no interior. O primeiro poço de cada placa serve como controlo negativo, pois toda a mistura de reacção é adicionada, excluindo o DNA. Segue-se depois a metodologia de PCR-SSP (Reacção de polimerase em

cadeia – PCR, utilizando primers específicos da sequência – SSP) que se baseia no princípio de que um par de *primers* de oligonucleótidos, com correspondência completa, são utilizados de forma mais eficaz pela *taq polimerase* recombinante para amplificar uma sequência alvo, do que um par de *primers* sem correspondência total. Os pares de *primers* são concebidos para apresentar apenas correspondências perfeitas com um único alelo ou grupos de alelos.

Depois do processo de PCR, os fragmentos de DNA amplificados são separados por electroforese em gel de agarose e visualizados por coloração com brometo de etídio e exposição a luz ultravioleta. A interpretação dos resultados de PCR-SSP baseou-se na presença ou ausência de um fragmento específico de DNA amplificado e os resultados foram analisados com recurso ao software One Lambda DNA/LMT. Como durante a reacção de PCR a amplificação pode ser influenciada adversamente por vários factores (erros de pipetagem, DNA de qualidade deficiente, presença de inibidores, etc.), é incluído um par de *primers* de controlo interno em todas as reacções de PCR. Este par de *primers* de controlo amplifica uma região conservada do gene da  $\beta$ -globulina humana, que está presente em todas as amostras de DNA e que é usada para confirmar a integridade da reacção de PCR. A **Fig 5** exemplifica o tipo de resultados obtidos com este protocolo.



**Figura 5.** Géis de agarose com bandas indicativas dos alelos HLA-B. O primeiro poço de cada gel corresponde ao controlo negativo, em que não ocorre amplificação, e observa-se amplificação do gene da  $\beta$ -globulina humana em todos os restantes poços (controlo positivo). A amplificação específica observa-se apenas em alguns poços. **a)** Indivíduo HLA-B\*14,51; **b)** Indivíduo HLA-B\*27,51

### 3.4 Estudo de *linkage*

### 3.4.1 Genotipagem com o *GeneChip* de 250K da Affymetrix

O *GeneChip Mapping 500K Array Set* humano da Affymetrix permite uma cobertura muito densa do genoma humano para estudos de genotipagem. É constituído por dois *chips*, cada um capaz de genotipar uma média de 250,000 SNPs (aproximadamente 262,000 para o *chip* da enzima NspI e 238,000 para o da enzima StyI). Cada *chip* contém polimorfismos bialélicos arbitrariamente distribuídos pelo genoma, excepto para o cromossoma Y. Esta técnica permite analisar marcadores genéticos com uma distância média de 5.9Kb e uma heterozigotia média de 0,30<sup>98</sup>.

O DNA genómico total dos participantes foi concentrado por precipitação e posterior resuspensão do *pellet* formado. Cada amostra foi hibridada apenas com um dos *chips* de 250K, que se optou por ser o StyI. Para tal, as amostras de DNA foram processadas de acordo com o manual de técnicas de mapeamento do *GeneChip* de 500K (*GeneChip Mapping Assay*) da Affymetrix e com os procedimentos em vigor na Unidade de Genómica do Instituto Gulbenkian de Ciência. Cada amostra foi estudada em apenas um dos *chips* com 250K, neste caso o StyI.

Digeriu-se 250ng de DNA genómico com a enzima de restrição StyI, que reconhece uma sequência específica de nucleótidos e cliva o genoma. Os fragmentos obtidos foram ligados a um adaptador e amplificados por PCR usando um *primer* genérico que reconhece a sequência do adaptador e amplifica preferencialmente os fragmentos de tamanhos compreendidos entre 200 e 1100 pb. Os produtos de PCR foram de seguida purificados, fragmentados e hibridados ao *GeneChip Mapping 250K StyI*. Finalmente, os *microarrays* são lavados, corados e lidos com um scanner Agilent GeneArray 2500. Os genótipos obtidos para cada uma das sondas que representam os SNPs no *array* foram gerados através do *software GeneChip* de análise de genotipagem (GTYPE 4.1). Este programa determina também um valor de confiança para cada genótipo individual, usando um algoritmo automático de análise dos sinais obtidos (BRLMM)<sup>98</sup>. Foram seguidos numerosos e rigorosos passos de controlo de qualidade ao longo das fases de preparação das amostras e de processamento dos *microarrays*.

Excluíram-se de análises posteriores os SNPs com erros mendelianos ou SNPs com ausência de genótipo em pelo menos um indivíduo.

### 3.4.2 Modelos fenotípicos

Para o presente estudo foram criados diferentes modelos fenotípicos de modo a analisar-se, não só a transmissão dos fenótipos – DB e EA – em conjunto, mas também destas doenças em separado e de alguns manifestações isoladas características da DB. Deste modo foram avaliados os seguintes modelos nas análises de *linkage*:

1. **Modelo DB** – Consideraram-se como afectados apenas os indivíduos DB;
2. **Modelo EA** – Consideraram-se como afectados apenas os indivíduos com EA;
3. **Modelo BD e EA** – Considerou-se como afectados os indivíduos com DB, assim como os indivíduos com EA;
4. **Modelo Lesões Oculares** – Consideraram-se como afectados os indivíduos com problemas oculares, tais como, uveítes anteriores e posteriores, edemas na papila, escotomas e amauroses;
5. **Modelo Aftas orais** – Consideraram-se como afectados os indivíduos com aftas orais recorrentes;
6. **Modelo Lesões cutâneas** – Consideraram-se como afectados os indivíduos com problemas de pele típicos da DB, tais como pseudo-foliculite e eritema nodoso.

Como o programa usado para a análise de *linkage*, o Merlin<sup>99</sup>, possui limitações técnicas relativamente ao número máximo de indivíduos por família que pode analisar (referida no ponto 3.4.3), não se puderam englobar todos os indivíduos não afectados em cada um dos modelos referidos, tendo-se seleccionado para cada uma das análises, os indivíduos considerados mais informativos.

### 3.4.3 Análise estatística

Para cada um dos seis modelos fenotípicos considerados foram realizadas análises de *linkage multipoint* não paramétricas (NPL). Considerando a hipótese nula ( $H_0$ ) de um *locus*

não estar em *linkage* (ligado) com um gene de susceptibilidade à doença, o comportamento estatístico do número de alelos IBD (idênticos por descendência) entre familiares, depende apenas das relações de parentesco entre indivíduos, que são determinadas pela estrutura do pedigree e não pelo fenótipo dos indivíduos. Para um *locus* que está em *linkage* com um gene de susceptibilidade à doença, espera-se que haja um aumento no número de alelos IBD entre os indivíduos afectados, relativamente à hipótese nula. Deste modo, testar a existência de *linkage* é testar o excesso de partilha de alelos IBD entre indivíduos afectados. A magnitude deste aumento na partilha dependerá, de um modo geral, do modo de hereditariedade da doença e da distância entre o *locus* ligado (marcador) e o *locus* responsável pela doença. No entanto, para doenças complexas, o modo de hereditariedade é desconhecido, tornando-se mais fácil modelar o grau e direcção do excesso de partilha de alelos IBD directamente através de uma análise de *linkage* não paramétrica (NPL)<sup>100</sup>. Comparativamente aos métodos tradicionais paramétricos, esta abordagem tem a vantagem de não necessitar da especificação de um modelo de hereditariedade<sup>100</sup> pois em doenças multifactoriais, onde vários genes (e factores ambientais) podem contribuir para o risco para a doença, não existe um modelo claro de hereditariedade<sup>101</sup>. Foram realizadas análises *multipoint* de *linkage* de forma a poder ser avaliado o *linkage* de múltiplos marcadores na mesma região com o gene responsável pelas doenças. Este tipo de análise é particularmente poderoso quando a relação dos alelos IBD, no *locus* em *linkage* com a doença, não é clara entre os elementos da família. Esta análise torna-se melhor porque, ao contrário das análises que consideram cada marcador isoladamente (*two-point*), esta usa a informação dos haplótipos, constituídos por vários marcadores, para inferir as relações dos alelos IBD<sup>102</sup>.

As análises de *linkage* foram realizadas usando o software Merlin<sup>99</sup> versão 1.1-alpha3. Este programa foi seleccionado porque, além de realizar análises NPL, consegue lidar com informação muita extensa no que respeita ao número de marcadores usados, tornando-se útil quando se realiza um *whole genome linkage screen*. Possui, no entanto, limitações em relação ao número máximo de indivíduos por família que pode analisar, regendo-se pela fórmula  $2N-3F \leq 23$ , em que N é o número total de indivíduos e F é o número de fundadores<sup>99</sup>.

Para atribuição dos valores de LOD *score* foram usadas ambas as estatísticas  $S_{all}$  e  $S_{pairs}$  de Whittemore e Halpern<sup>103</sup>. Considerando-se um *locus* que não tem necessariamente

de ser um marcador, mas que tem um ou mais marcadores na sua vizinhança, para o *pedigree*  $i$ ,  $S_i$  é a função que é definida com base na partilha de alelos IBD nesse *locus* entre os indivíduos afectados. Whittemore e Halpern criaram estatísticas para medir esta partilha de alelos, representando  $S_{\text{pairs}}$  o número de alelos IBD partilhados por pares de indivíduos afectados no *pedigree*, enquanto  $S_{\text{all}}$  calcula a probabilidade dos alelos IBD observados sob todas as configurações possíveis, atribuindo maior peso quando há três ou mais indivíduos afectados que partilham o mesmo alelo IBD<sup>100</sup>. A estatística  $S_{\text{all}}$  pode oferecer maior poder estatístico pois considera grupos maiores de parentes afectados, enquanto a  $S_{\text{pairs}}$  considera apenas pares. Sucintamente, é mais “convicente” descobrir que cinco indivíduos afectados partilham o mesmo alelo IBD do que descobrir que cada par deles partilha um alelo IBD<sup>105</sup>. A estatística  $S_{\text{all}}$  também é melhor para analisar *loci* dominantes, uma vez que é mais provável que um número maior de indivíduos afectados por família partilhe o mesmo alelo<sup>106</sup>. Para famílias onde existem dois ou três irmãos afectados, com os pais classificados como não afectados ou com fenótipo desconhecido, as estatísticas  $S_{\text{pairs}}$  e  $S_{\text{all}}$  serão equivalentes após normalização<sup>105</sup>.

Realizaram-se análises NPL *multipoint* em que se calcularam os NPL *z-score* com *p-value* assumindo aproximação normal, de Kruglyak *et al.*<sup>104</sup>, e o LOD *score* não paramétrico segundo os modelos linear e exponencial de Kong and Cox<sup>100</sup>, tendo sido em cada um destes modelos utilizadas as estatísticas  $S_{\text{all}}$  e  $S_{\text{pairs}}$  para contabilizar o *linkage*. Kruglyak *et al.* propuseram uma estatística referida como NPL *z-score*, também designada por aproximação aos dados perfeitos, evidenciando que sob a hipótese nula, e quando são analisados muitos *pedigrees* semelhantes, um *score* NPL é uma estatística com uma distribuição normal – média de zero e variância de um<sup>105</sup>. Como a distribuição normal facilmente desaparece quando a amostra é pequena, esta estatística usa uma simulação para determinar a significância específica de cada amostra (um *p-value* exacto), através da enumeração/contabilização de todos os vectores de hereditariedade possíveis, e obtendo a distribuição do NPL *z-score*. No entanto, normalmente, a informação sobre a descendência é incompleta, não podendo ser contabilizados todos os vectores de hereditariedade possíveis. Nestes casos, as propriedades assintóticas da estatística NPL não são de confiança, pois embora a média se mantenha zero, a variância de um modo geral irá ser inferior a um. Mesmo tendo em conta esta excepção, Kruglyak *et al.* aconselham o uso de uma distribuição

normal (exacta), como referência para produzir um *p-value*, tornando o teste conservativo e inaceitável para muitos exemplos reais<sup>100,105</sup>. Para eliminar este conservadorismo, Kong e Cox propuseram uma abordagem de probabilidades, que calcula as probabilidades exactas e produz *LOD scores* baseados numa estatística de partilha de alelos de um só parâmetro,  $Z_{lr}$ , em vez de usar uma aproximação normal. Neste modelo linear de Kong e Cox, os *LOD scores* não paramétricos são produzidos com base na maximização de um único parâmetro variável ( $\delta$ ) no numerador, com base nos genótipos observados. O  $\delta$  representa o grau de partilha de alelos no *pedigree*, onde sob a hipótese nula ( $H_0$ ) o  $\delta=0$  e onde  $\delta>0$  corresponderá a um excesso na partilha<sup>105</sup>. No entanto, devido ao limite superior do  $\delta$ , este não permite que o modelo linear represente um elevado aumento na partilha de alelos IBD. Deste modo, quando o LOD é calculado a partir de uma amostra populacional pequena, e/ou *pedigrees* que diferem no tamanho, como nos casos em que existe um pequeno número de famílias e uma elevada partilha de alelos IBD, Kong e Cox sugeriram a utilização de outro modelo, que designaram de exponencial e que se baseia na mesma estatística que o linear, mas com algumas propriedades adicionais, nomeadamente o facto de  $\delta$  não ter limite superior. Este modelo não possui, no entanto, a capacidade do modelo linear de lidar com informação incompleta<sup>100</sup>.

Os *LOD scores* foram calculados para uma densidade de marcadores de 0.1cM e o conteúdo informativo dos marcadores estimado através da informação da entropia.

Para declarar *linkage* sugestivo e significativo foram usados os tradicionais limiares de *LOD scores* de 2.0 e de 3.0, respectivamente. Os comandos usados para realizar a análise no Merlin, assim como o tipo de ficheiro de input apresentam-se no **Anexo B**.

### **3.5 Análise de perda de heterozigotia (LOH)**

Analisou-se a perda de heterozigotia (LOH) com a ferramenta *Chromosome Copy Number Analysis Tool 4.0* (CNAT4), que faz parte do pacote do *GeneChip Genotyping Analysis Software* (GTYPE) da Affymetrix. Usou-se como referência um grupo de 48 Hap Maps, com genótipos disponibilizados pela Affymetrix. Os parâmetros usados para calcular o LOH foram: restrição da análise a fragmentos compreendidos entre 150 – 2000 pb, que é o tamanho de fragmentos que se espera obter após digestão com a StyI; normalização dos

dados com uma escala mediana, que faz uma normalização linear baseada na mediana das medianas de todos os ficheiros incluídos na análise e *transition decay* igual a 10Mb. Este último parâmetro controla a correlação esperada entre SNPs adjacentes<sup>106</sup>.

### 3.6 Análise dos genótipos na região 8p21.3-22

Para melhor visualização e confirmação dos resultados obtidos nos estudos de *linkage* (que não puderam incluir simultaneamente todos os indivíduos genotipados, pelas limitações técnicas já referidas), procedeu-se a uma análise “manual” dos genótipos dos marcadores situados no pico de *linkage* obtido (região 8p21.3-22) e nas regiões flanqueantes, em todos os indivíduos genotipados. Esta análise foi realizada de modo a averiguar-se se existia uma diferenciação desta região entre os indivíduos com DB e com EA e os restantes indivíduos, e se esta diferenciação era limitada à região 8p21.3-22.

### 3.7 Pesquisa de mutações na família

De modo a identificarem-se os possíveis factores genéticos partilhado entre a DB e a EA, procedeu-se à sequenciação dos genes candidatos, situados no pico de *linkage* obtido na análise do modelo fenotípico “Lesões oculares. A sequenciação possibilita a identificação de novas variantes genéticas responsáveis por determinadas doenças<sup>107</sup>. Os genes foram seleccionados de acordo com a sua função biológica.

#### 3.7.1 Escolha dos *primers*

Foram desenhados *primers* (iniciadores) para os três genes seleccionados, de modo a amplificar as regiões codificantes de todos os exões e cerca de 20 pb de ambas as regiões flanqueadoras não codificantes. Para os genes *NAT1* e *NAT2*, foram desenhados dois pares de *primers*, para se conseguir amplificar a totalidade do exão codificante. Foi necessário apenas um par de *primers* por exão para o gene *CHGN*, (Condroitina  $\beta$ -1,4-N-acetilgalactosaminiltransferase 1) com a excepção do exão 5, que possui dois transcritos diferentes. Os *primers* foram escolhidos com o auxílio de um programa disponível on-line, denominado de Primer 3 ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)). Os

*primers* escolhidos para cada gene e as suas características encontram-se descritos na **Tabela 5**.

**Tabela 5.** Características dos *primers* desenhados para amplificação dos genes candidatos

Gene	Exão	Primers (for-forward; rev-reverse)	T° m	Dimensão do produto de PCR amplificado
NAT1	3	AGCCATAATTAGCCTACTCAAATCC 3' (for)	60.24°C	685 pb
		AATCTTCAATTGTTTCGAGGCTTA 3' (rev)	59.3°C	
		ACAGCACTGGCATGATTCAC 3' (for)	59.71°C	
		CAGAAATACATGATAGGTCGTCA 3' (rev)	58.29°C	
NAT2	2	CACACGAGGAAATCAAATGC 3' (for)	59.13°C	705 pb
		GGAGACGTCTGCAGGTATGT 3' (rev)	57.74°C	
		ACCATTGACGGCAGGAATTA 3' (for)	60.33°C	
CGTGAGGGTAGAGAGGATATCTG 3' (rev)	59.27°C			
	3	CTCCAAGAGGCAGAACTCGT 3' (for)	59.6	950 pb
		TAAAAAGCCATGCCGTACCT 3' (rev)	59.61	
	4	TTCTGTGCTAGATCTCGGTTGA 3' (for)	60.02	411 pb
		GAGAAAGGAGGCTTTGTACTCG 3' (rev)	59.54	
CHGN	5 T1	TTCTAGTGGTGGCTGCAATG 3' (for)	59.86	262 pb
		GCTGGTTATCCTCAGCCAAT 3' (rev)	59.15	
	5 T2	TCGGCTCTCCTGGAAGTTTA 3' (for)	59.95	364 pb
		GCTGGTTATCCTCAGCCAAT 3' (rev)	60.49	
	6	ACTTCTTGGGGAGTCCCAGT 3' (for)	59.97	391 pb
		ACCCATCAAGAGGACGACAC 3' (ver)	59.97	
	7	GTTTCAGAGGGCCAAATGTC 3' (for)	59.53	315 pb
		GCACATCTGCCAAACGTAAA 3' (rev)	59.74	
	8	CAAATCCATCTGCGTTTCCT 3' (for)	60.07	398 pb
		CACCGCACCTAGTCTGGAAT 3' (rev)	60.13	
9	TTCTGACCCAGGTGGAACT 3' (for)	60.56	600 pb	
	AGCCCAACAGAGAGAAATCG 3' (rev)	59.43		

T1 – transcrito 1; T2 – transcrito 2, T°m – temperatura de *melting* (desnaturação) dos *primers*

### 3.7.2 Amplificação dos genes candidatos

Inicialmente realizaram-se reacções de PCR para cada um dos pares de *primers*, em dois indivíduos afectados com EA (indivíduos III:3 e III:10), em dois afectados com DB (III:5 e III:2), num afectado com EpAI (IV:2) e num indivíduo não afectado (IV:6) (**Fig 4**). Nos exões em que se observou a presença de polimorfismos nos indivíduos afectados, inexistentes nos não afectados, realizaram-se posteriormente reacções de PCR e de sequenciação em todos os elementos da família, de modo a poder confirmar-se a existência de uma possível mutação.

Relativamente à metodologia utilizada, o DNA foi primeiramente adicionado a placas de PCR e, seguidamente, adicionou-se a mistura de reacção, obtendo-se um volume final de 12.5µl. Os volumes e as concentrações dos componentes presentes na mistura de reacção encontram-se na **Tabela 6**. As condições utilizadas no termociclador My-Cycler da Bio-Rad estão descritas na **Tabela 7**.

**Tabela 6.** Mistura de reacção utilizada para a reacção de PCR

Componentes (concentração inicial)	Volume (µl) (concentração final)
DNA (25ng/µl)	1.0
Tampão (10x)	1.3 (1x)
dNTP MIX (25mM)	0.1 (0.2mM)
Primer Forward (0.1µg/µl)	0.5 (0.04µg/µl)
Primer Reverse	0.5 (0.04µg/µl)
MgCl <sub>2</sub> (25mM)	0.8 (1.5mM)
Taq Polimerase (2U/µl)	0.3 (0.5U)
Água HPLC	8.0
<b>Total</b>	<b>12.5</b>

**Tabela 7.** Condições usadas no termociclador durante a reacção de PCR *stepdown* para amplificação dos exões dos genes candidatos

Etapas do Ciclo de PCR	Temperatura	Tempo	Número de Ciclos
Desnaturação inicial	94°C	3 min	1
Desnaturação	94°C	30 seg	15
Annealing	70°C - 1°C/ciclo	30 seg	
Extensão	72°C	1 min	
Desnaturação	94°C	10 seg	30
Annealing	55°C	45 seg	
Extensão	72°C	1 min	
Extensão final	72°C	7 min	1

Usou-se um programa de PCR *Stepdown* (SD) para amplificar os genes candidatos, pois este é um método de PCR robusto que evita que os *primers* amplifiquem sequências não específicas. Ligações não específicas do *primer* perturbam os resultados da reacção de PCR, pois as sequências não específicas, às quais os *primers* se ligam durante os primeiros passos da amplificação, vão depois sobrepor-se às sequências específicas, devido à natureza exponencial da amplificação pela polimerase. Num programa de PCR SD, os primeiros

ciclos possuem temperaturas de *annealing* mais elevadas, que vão sendo diminuídas sequencialmente em cada ciclo. Deste modo, nos primeiros ciclos os *primers* vão hibridar apenas com a sequência de interesse, que possui uma elevada especificidade, e apenas estes fragmentos vão ser amplificados. Assim, os fragmentos específicos ganham vantagem, podendo-se nos ciclos seguintes baixar a temperatura, com conseqüente diminuição da especificidade, para aumentar a eficácia de amplificação.

A confirmação da amplificação resultante das reacções de PCR foi realizada por electroforese em géis de agarose TAE a 2%, corados com brometo de etídeo. Os produtos foram visualizados por exposição a luz ultravioleta.

### 3.7.3 Sequenciação e análise dos produtos de PCR obtidos

Após a amplificação realizou-se a sequenciação dos produtos de PCR amplificados. Esta foi realizada com o *kit BigDye® Terminator v1.1 Cycle Sequencing* da Applied Biosystems segundo as recomendações do fabricante. Realizaram-se amplificações em cadeia simples, para cada um dos produtos de PCR, fazendo-se reacções separadas para os *primers forward* e *reverse*, de modo a confirmar-se os resultados. Os reagentes utilizados, assim como a quantidade usada de cada um, encontram-se representados na **Tabela 8**. As condições utilizadas no termociclador *9800 Fast* da Applied Biosystems estão descritas na **Tabela 9**.

**Tabela 8.** Mistura de reacção utilizada na reacção de sequenciação.

Componentes (concentração inicial)	Volume (µl) (concentração final)
DNA- Produto de PCR (70ng/µl)	1.0
Tampão de Sequenciação (5x)	2.0 (1x)
<i>Big Dye Terminator v1.1</i>	2.0
Primer (0.1µg/ul)	0.32 (0.32pmol/µl)
Água	4,68
<b>Total</b>	<b>10.0</b>

Após a reacção de sequenciação, precipitou-se o produto de amplificação com etanol e acetato de sódio, de modo a purificar as amostras. Esta purificação consistiu na adição

consecutiva de 10µl de H<sub>2</sub>O, 2.0µl de acetato de sódio (3M) e 50µl de etanol a 95% a cada um dos produtos de amplificação, tendo-se deixado as misturas a precipitar durante a noite a 4°C. Posteriormente, realizou-se uma centrifugação a 4°C, 30min a 14000rpm, ao fim da qual se retirou cuidadosamente o sobrenadante e se adicionaram 250µl de etanol a 70%. Esta solução foi vortexada e centrifugada a 4°C, a 14000 rpm durante 15 min e por fim retirou-se e deixou-se o *pellet* a secar ao ar. O restante protocolo de sequenciação foi realizado de acordo com o protocolo da Applied Biosystems.

Os produtos de amplificação, purificados e tratados, foram sequenciados num analisador genético automático ABI Prism 3130 XL [Applied Biosystems] e as sequências obtidas analisadas usando o *software Staden Package*. Dentro deste pacote usou-se o programa Pregap, para confirmar a qualidade das sequências e transformá-las num formato compatível com o programa Gap4, do mesmo pacote, onde se fez a posterior análise. Para cada exão analisado, recorreu-se a uma sequência padrão, retirada de uma base de dados oficial (ensembl.org), para alinhar as respectivas sequências.

**Tabela 9.** Condições usadas no termociclador durante a reacção de sequenciação.

<b>Etapas do Ciclo de PCR</b>	<b>Temperatura</b>	<b>Tempo</b>	<b>Número de Ciclos</b>
Desnaturação inicial	96°C	1 min	1
Desnaturação	96°C	10 seg	
<i>Annealing</i>	50°C	5 seg	25
Extensão	60°C	4 min	

### 3.8 Construção dos haplótipos para o gene *NAT2*

Foram reconstruídos haplótipos para o gene *NAT2*, para todos os elementos da família, inclusivamente os não genotipados, recorrendo ao programa Simwalk2 v2.91<sup>108</sup> e ao Merlin v1.1-alpha3<sup>99</sup>. Os ficheiros de *input*, para ambos os programas, foram construídos a partir do *software* Solar v4.1.5<sup>109</sup>. A frequência dos alelos de cada SNP foi indicada de acordo com a informação existente para os trios de Hap-Maps caucasianos e retirada duma base de dados oficial (ensembl.org), assim como a informação acerca da distância entre os cinco SNPs que constituem os haplótipos.

## 4.0 RESULTADOS

### 4.1 Agregação familiar da DB, EA e EpAI

A árvore genealógica da família com os respectivos diagnósticos clínicos encontra-se representada na **Fig 4**. As características clínicas para todos os elementos da família encontram-se na **tabela 10**.

**Tabela 10:** Sumário de alguns dados clínicos dos indivíduos da família incluídos no estudo.

ID	Género (M/F)	Idade na observação	DB	AO	AG	LO	LC	EA	EpAI	Artrite	Artralgia	MSN
I:1	M	—								X	?	
I:2	F	—				X				X	?	
II:1	M	—				X						
II:2	F	—								X	?	
II:3	M	—				X						
II:4	F	—										
III:1	M	?										
III:2	F	64	47	45	47	48	42			45		62
III:3	M	63		15		X	15	17				
III:4	F	60										
III:5	M	57	X	10		37	20				X	
III:6	F	51										
III:7	M	60	30	8		30						
III:8	M	66		50				26		26	26	
III:9	F											
III:10	M	61				50		21				
III:11	F	62										
III:12	F	57		20				X		40	40	
III:13	M	62		54							46	28
IV:1	F	36					33				34	32
IV:2	F	35				X			X		X	
IV:3	F	23		8			19					
IV:4	F	18					13					
IV:5	M	40							X	13	13	
IV:6	M	31										
IV:7	F	29										
IV:8	F	25		10								
IV:9	F	34						18				
IV:10	M	32										
IV:11	M	30					16					

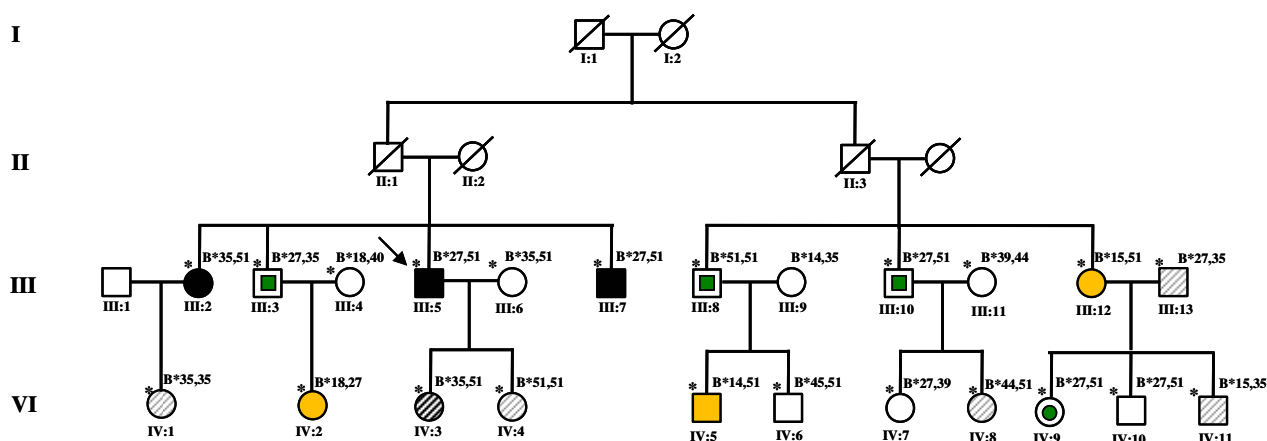
**Legenda:** Para cada sintoma/doença presente num indivíduo encontra-se a idade de início ou então um **X** nos casos em que a idade é desconhecida. **DB** – Doença de Behçet; **AO** – aftas orais; **AG** – aftas genitais; **LO** – Lesões oculares; **LC** – lesões cutâneas típicas da DB; **EA** – Espondilite anquilosante, **EpAI** – Espondilartropatia indiferenciada; **MSN** – manifestações do sistema nervoso.

Na família em estudo, três membros (III:2, III:5 e III:7) satisfazem os Critérios de Estudo Internacionais para a doença de Behçet Revistos (2006)<sup>21</sup> (**Tabela 2**), que são critérios com uma elevada sensibilidade, dois deles (III:2 e III:5) preenchendo também os critérios do Grupo de Estudo Internacional para a Doença de Behçet (1990)<sup>20</sup> (**Tabela 1**), que são critérios com grande especificidade mas menor sensibilidade<sup>21</sup>. Quatro indivíduos (III:3, III:8, III:10 e IV:9) satisfazem os critérios de classificação de Nova York modificados (1984) para a espondilite anquilosante<sup>22</sup> (**Tabela 3**), sendo um deles (III:10) classificado como uma forma espondilítica da Artrite Psoriásica (indistinto da EA). Três indivíduos (III:12, IV:2 e IV:5) preenchem os critérios de classificação do Grupo Europeu para o Estudo das Espondilartropatias (ESSG) (1991)<sup>23</sup> para as EpAI (**Tabela 4**). Dos quatro indivíduos com EA, dois (III:3 e III:8) apresentam também sintomas de DB, embora não possuam toda a sintomatologia necessária para serem considerados doentes de Behçet. Adicionalmente, existem nesta família seis indivíduos (III:3, IV:1, IV:3, IV:4, IV:8 e IV:11) que possuem manifestações isoladas da DB, tais como, aftas orais recorrentes e lesões na pele características, alguns deles exibindo também problemas articulares ligeiros (III:13 e IV:1), não tendo sido no entanto diagnosticados com qualquer doença reumática. Os restantes membros da família não possuem manifestações relacionadas com a DB ou com a EA.

#### **4.2 Associação dos alelos HLA-B com a DB, EA e EpAI**

Os resultados da genotipagem para o HLA-B encontram-se representados na **Fig 6**. Observa-se que existe uma elevada frequência dos alelos HLA-B\*27 e B\*51, tanto nos pacientes com DB e EA e EpAI, como nos indivíduos saudáveis. Relativamente à doença de Behçet, pode-se observar que todos os indivíduos afectados possuem pelo menos um alelo HLA-B\*51. No que respeita ao complexo das EpA observa-se que três dos quatro pacientes com EA (III:3, III:10 e IV:9) possuem pelo menos um alelo HLA-B\*27, possuindo o outro paciente (III:8) dois alelos B\*51. Observa-se ainda que um dos três indivíduos com EpAI (IV:2) é HLA-B\*27 positivo, sendo os outros dois indivíduos (III:12 e IV:5) positivos apenas para o alelo B\*51. Verificou-se, também, a existência de dois indivíduos afectados por DB (III:5 e III:7) e dois por EA (III:10 e IV:9) com o mesmo genótipo HLA-B\*27,51. O

facto de não se ter verificado nesta família uma perfeita segregação do B\*51 com DB e do B\*27 com EA que pudesse por si só justificar estas patologias nos indivíduos afectados, suportou o prosseguimento do estudo para uma análise de *linkage* de modo a identificar-se o(s) gene(s) responsáveis pelas patologias existentes nesta família.



**Figura 6.** Árvore genealógica com representação dos genótipos obtidos para a tipagem do HLA-B. Os indivíduos que foram genotipados encontram-se marcados por (\*), tendo sido inferidos os genótipos para os restantes indivíduos. Os diagnósticos clínicos encontram-se representados de modo semelhante à **Figura 4** (ver legenda da **Figura 4**)

### 4.3 Estudo de *linkage*

#### 4.3.1 Genotipagem com o *GeneChip* de 250K da Affymetrix

As amostras indicadas com (\*), na **Fig 2, 6 e 14**, foram genotipadas com o *chip* 250K StyI da Affymetrix. A tabela 11 apresenta os valores obtidos no controlo de qualidade para todos os *chips* efectuados. Foi necessário a repetição de um *chip* para o indivíduo IV:9, pois o número de calls obtido (91.80%) foi bastante inferior à média obtida (97.10%) e à média esperada (>96%)<sup>97</sup>, tendo-se com o novo *chip* (IV:9\_2) obtido um número superior de sinais positivos (93.07%), mas ainda assim inferior aos restantes. Obteve-se uma informação média de cerca de 232,000 genótipos por indivíduo, dos quais cerca de 182,000 foram seleccionados para serem analisados após a aplicação do filtro. Os géneros foram correctamente atribuídos para todos os indivíduos e a heterozigotia média observada para os SNP foi de 28.61%, aproximando-se da esperada (30%)<sup>97</sup>. Estes resultados indicam uma boa qualidade do DNA e sucesso no protocolo de realização dos *chips*.

**Tabela 11.** Valores obtidos no controlo de qualidade para os *chips* de 250K da Affymetrix

ID do <i>chip</i>	Género		SNP Call (%)	AA Call (%)	AB Call (%)	BB Call (%)
	atribuído (M/F)					
III:2	F		98.31	37.11	27.67	35.22
III:3	M		99.21	37.44	26.67	35.89
III:4	F		98.53	37.49	26.84	35.67
III:5	M		99.06	37.36	26.88	35.76
III:6	F		99.02	37.12	27.42	35.46
III:7	M		98.61	37.37	27.01	35.62
III:8	M		98.39	37.56	26.80	35.64
III:10	M		94.31	36.35	30.21	33.45
III:11	F		98.60	37.23	27.19	35.57
III:12	F		97.11	36.72	28.41	34.87
III:13	M		97.84	37.60	26.74	35.66
IV:1	F		96.82	34.13	33.49	32.38
IV:2	F		96.49	36.74	28.89	34.37
IV:3	F		97.95	36.97	28.05	34.99
IV:4	F		96.89	36.67	29.08	34.25
IV:5	M		95.95	36.57	29.46	33.97
IV:6	M		98.53	37.40	26.93	35.68
IV:7	F		96.82	36.81	28.81	34.37
IV:8	F		97.37	36.82	28.53	34.65
IV:9_1	F		91.80	34.58	33.63	31.79
IV:9_2	F		93.07	35.74	31.25	33.01
IV:10	M		98.23	37.16	27.58	35.26
IV:11	M		94.47	36.19	30.57	33.24
<b>Valor médio</b>			97.10	36.74	28.61	34.64

### 4.3.2 Modelos fenotípicos

Os indivíduos considerados como afectados e não afectados para cada modelo fenotípico nas análises de *linkage* encontram-se descritos na **Tabela 12** e no **Anexo C**.

**Tabela 12.** Indivíduos seleccionados para cada uma das análises de *linkage*.

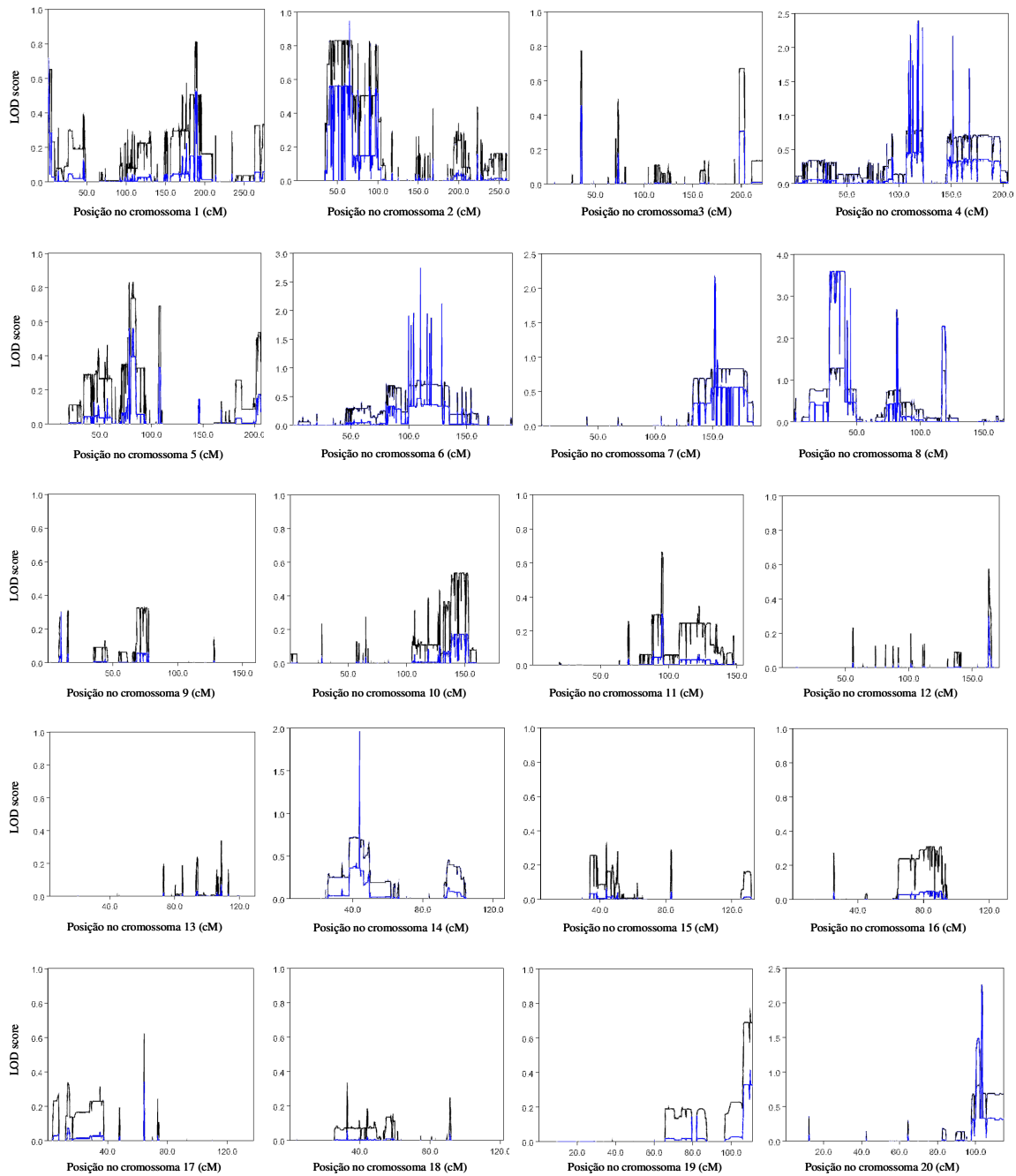
Modelo Fenotípico	Indivíduos afectados	Indivíduos não afectados
BD	III:2, III:5 e III:7	I:1, I:2, II:1, II:2, II:3, II:4, III:1, III:3, III:4, III:6, III:8, III:10, III:13, IV:1, IV:2 e IV:4
EA	III:3, III:8, III:10 e IV:9	I:1, I:2, II:1, II:2, II:3, II:4, III:2, III:4, III:5, III:7, III:9, III:12, III:13, IV:2, IV:5, IV:6 e IV:10
BD e EA	III:2, III:3, III:5, III:7, III:8, III:10 e IV:9	I:1, I:2, II:1, II:2, II:3, II:4, III:1, III:4, III:9, III:12, III:13, IV:1, IV:2, IV:5, IV:6, IV:10 e IV:11
Lesões oculares	I:2, II:1, II:3; III:2; III:3, III:5, III:7, III:10 e IV:2	I:1, II:2, II:4; III:1; III:4, III:13 e IV:1
Manifestações cutâneas	III:2, III:3, III:5; IV:1, IV:3, IV:4 e IV:11	I:1, I:2, II:1, II:2, II:3, II:4, III:1, III:4, III:6, III:7, III:8, III:10, III:12, III:13, IV:2, IV:9 e IV:10
Aftas orais	III:2, III:3, III:5, III:7, III:8, III:12; III:13, IV:3 e IV:8	I:1, I:2, II:1, II:2, II:3, II:4, III:1, III:4, III:6, III:9, III:10, III:11; IV:1, IV:2, IV:5, IV:7, IV:9 e IV:11

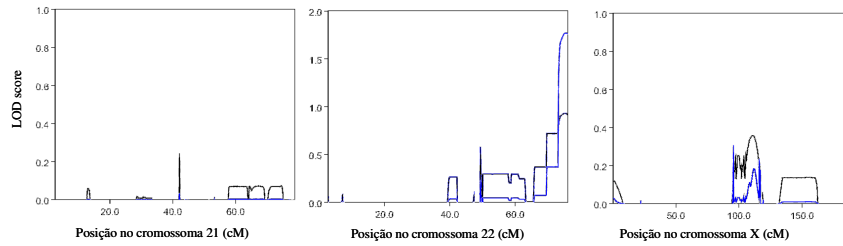
Na análise de *linkage* considerando o modelo fenotípico BD e EA, não se incluiu os indivíduos da geração IV, com idade inferior a 30 anos. Embora, em primeira análise, se pudesse considerar mais informativo incluir os descendentes de todos os indivíduos afectados, estes ainda não possuem toda, ou qualquer, sintomatologia de DB e EA, mas estão em risco de vir a desenvolver estas doenças. Sabe-se que a DB afecta predominantemente pessoas na terceira e quarta década de vida, deste modo, ao incluir estes indivíduos mais jovens na análise, correr-se-ia o risco de lhes atribuir um fenótipo errado. Já para as análises dos modelos “lesões oculares”, “problemas de pele” ou “aftas orais”, incluiu-se os indivíduos de idade inferior a 30 anos, que já possuíam esta sintomatologia.

### 4.3.3 Análise estatística

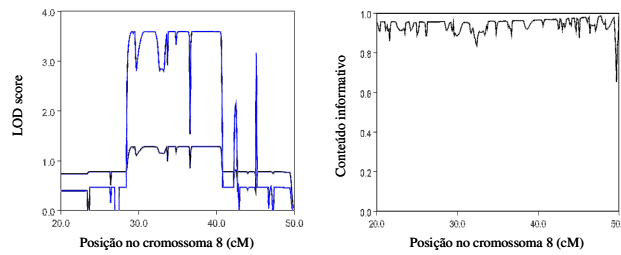
Os resultados das análises de *linkage multi-point* não paramétrico, para os diversos modelos fenotípicos, como a excepção da análise  $S_{all}$  para o modelo fenotípico “BD e EA” (**Fig 7**), encontram-se no **Anexo D**. Na **Tabela 13**, apresentam-se apenas os resultados que encontraram significância estatística. Observando de um modo geral os resultados, pode-se verificar que, embora as estatísticas *all* e *pairs* de Whittemore e Halpern tenham sido bastante consistentes para cada análise realizada em cada cromossoma, nos modelos “DB”, “DB e EA” e “EA” obtiveram-se LOD *scores* mais semelhantes entre as duas estatísticas do que nas análises “lesões oculares”, lesões cutâneas e “aftas orais”, em que  $S_{pairs}$  atingiu valores mais elevados. Analisando em pormenor os diferentes modelos fenotípicos, pode-se observar que para o modelo linear de Kong and Cox, não foram obtidos LOD *scores* significativos para nenhum deles. Já para o modelo exponencial de Kong and Cox, obteve-se um LOD *score* significativo (3.59) para o modelo fenotípico “BD e EA”, na região 8p21.3-22 entre os 28.5 e 40.7 cM (**Fig 8**). Nas análises considerando os modelos fenotípicos “Lesões oculares” e “aftas orais” (**Fig D6-D9**), obtiveram-se picos de *linkage* sugestivos (LOD=2.47 e LOD=2.64, respectivamente), também no cromossoma 8 e em regiões coincidentes com a obtida no modelo “BD e EA”, mas com extensões menores (**Fig 9**). Para além deste pico, também foram observados outros LOD *scores* elevados, no cromossoma 8 na região 117 – 119.8 cM para o modelo “aftas orais”, no cromossoma 4 e 6 para o modelo “BD e EA” ou no cromossoma 16 para o modelo “Lesões oculares”. No entanto, estes

reflectiram resultados de marcadores isolados ou de uma região muito pequena, não sendo valores de confiança. Para as restantes análises não se observou picos de *linkage*. Os valores do NPL *z-score* correlacionaram-se muito bem com os valores de LOD score para o modelo exponencial de Kong e Cox (**Tabela 13**), obtendo-se picos nas mesmas regiões.

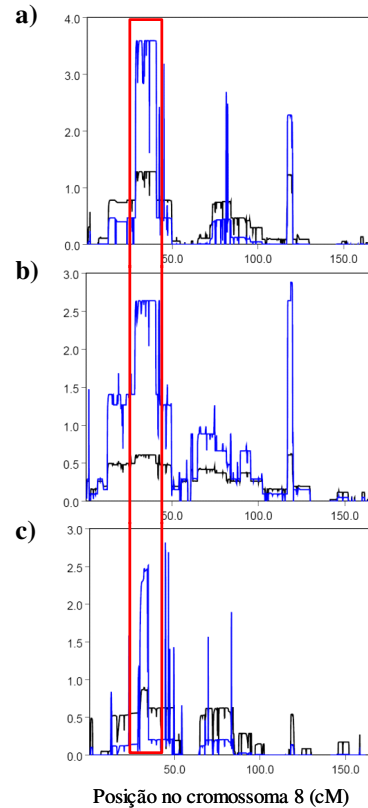




**Figura 7.** Análise NPL *multi-point* no genoma total para o modelo fenotípico “BD e EA”. Encontra-se representada apenas a análise  $S_{all}$ . A linha preta representa o modelo linear, a linha azul o modelo exponencial de Kong e Cox.



**Figura 8.** Pormenor da zona de *linkage* no cromossoma 8 para a análise NPL *multi-point* para o modelo fenotípico “BD e EA” e conteúdo informativo dos marcadores usados.



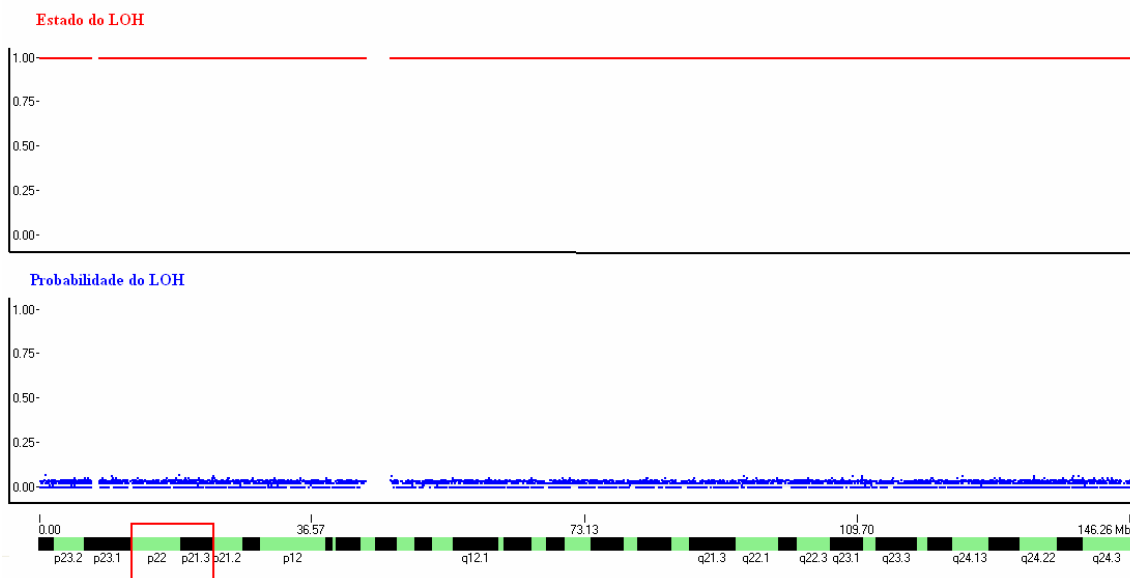
**Figura 9.** (À direita) Análises NPL *multi-point* no cromossoma 8.  
**a)** Estatística  $S_{all}$  para o modelo fenotípico “BD e EA”.  
**b)** Estatística  $S_{pairs}$  para o modelo fenotípico “aftas orais”.  
**c)** Estatística  $S_{all}$  para o modelo fenotípico lesões oculares

**Tabela 13.** Resumo das análises NPL *multi-point* com LOD score atribuído pela estatística  $S_{all}$  de Whittemore e Halpern para os diferentes modelos fenotípicos testados, com valores de *z-score* NPL, LOD score para o modelo exponencial de Kong a Cox e o *p-value* deste.

Cromossoma	Posição		SNPs flanqueantes	NPL Z-score	LOD exp	P value	Modelo fenotípico
	cM	Mb					
8	28.5 – 40.7	16.8 - 23	rs1389373 a rs4872049	12.74	3.59	$2.0 \times 10^{-5}$	DB e EA
8	30.1 – 34.6	17.8 - 20.2	rs3850746 a rs2410635	3.52	2.47	$4.0 \times 10^{-4}$	Lesões oculares
8	28.6 – 40.6	17 – 22.8	rs3844074 a rs1543019	5.12	2.64	$2.0 \times 10^{-4}$	Aftas orais
8	117–119.8	115.8 – 120.2	rs16886356 a rs11993915	7.16	3.02	$1.0 \times 10^{-4}$	Aftas orais

#### 4.4 Análise da perda de heterozigotica (LOH)

Não foram observados desvios do valor de 1 no que respeita à perda de heterozigotia no cromossoma 9, como se pode observar na **Fig 10**.



**Figura 10.** Representação dos valores de LOH e da probabilidade associada a este para o indivíduo III:2 da família em estudo. A escala de quantificação da perda de LOH vai de 0-1, onde 1 representa diploidia. Resultados semelhantes foram obtidos para todos os indivíduos.

#### 4.5 Análise dos genótipos na região 8p21.3-22

A análise dos genótipos suportou os resultados obtidos no estudo de *linkage* no genoma total, mostrando co-segregação dos alelos dos marcadores com as doenças de Behçet e espondilite anquilosante, na região 8p21.3-22. A região observada na análise “manual” era um pouco menor que a obtida na análise para o modelo fenotípico BD e EA e englobava a região obtida na análise de *linkage* com o modelo fenotípico “lesões oculares” (**Fig 11**). Também se observou que o indivíduo III:12 (paciente com EpAI) possuía esta região em comum com os indivíduos com DB e EA e que o indivíduo IV:9 (paciente com EA) possuía uma segregação deste região ligeiramente diferente dos outros indivíduos afectados por EA ou DB, o que sugere que talvez os indivíduos com EpAI também partilhem factores genéticos em comum com a BD e EA nesta família.

dbSNP	Posição física	DB			EA				EpAI			Restantes indivíduos																							
		III:2	III:5	III:7	III:3	III:8	III:10	IV:9	III:12	IV:2	IV:5	III:4	III:6	III:11	III:13	IV:1	IV:3	IV:4	IV:6	IV:7	IV:8	IV:10	IV:11												
rs2720602	15955791	AB	AB	AB	AB	AB	AA	BB	AB	BB	BB	AB	AA	AB	BB	BB	AA	AB	AA	AB	AA	AB	BB	BB	BB	BB	BB	AB	AB	BB	AA	AB	BB	BB	BB
rs10090333	16131941	AB	AB	AB	AB	AB	AB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs11998535	16387088	BB	BB	BB	BB	BB	AB	AB	BB	BB	BB	BB	AB	AB	BB	AB	AB	AB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs883355	16769576	AB	AB	AB	AB	AB	BB	AB	AB	AA	AB	AB	AB	AB	AB	AB	AB	AA	AB	AA	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs3850753	16920862	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs2014286	16951496	AA	AA	AA	AA	AA	AA	AB	AA	AA	AA	AA	AB	AB	AB	AA	AA	AA	AB	AA	AB	AA	AB	AA	AB	AA	AB	AA	AB	AA	AB	AA	AB	AA	AB
rs2705196	17053733	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	AB	BB	BB	AB	AB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs2904672	17102367	AB	AB	AB	AB	AB	AB	AA	AB	AA	AB	AA	AB	AA	AB	AA	AB	AB	AA	AB	AA	AB	AA	AB	AA	BB	AA	BB	AA	BB	AA	BB	AA	BB	AA
rs7460113	17342064	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	AA	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs2720574	17480455	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs7825094	17512758	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs17633197	17549692	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs2283119	17982022	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	BB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs10102201	18137713	AB	AB	AB	AB	AB	AB	AA	AB	BB	AA	AB	AB	AB	AB	AA	BB	AB	BB	AB	BB	AB	BB	AB	AB	BB	AA	BB	AA	BB	AA	BB	AA	BB	AA
rs7007390	18496552	BB	BB	BB	BB	BB	BB	BB	BB	AB	AA	AB	AB	AA	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs17469073	18862099	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AB	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs6586829	19326299	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs13261824	19358192	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB	AB	AB	AB	BB	BB	BB	BB	BB	BB	BB	AB	AB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB
rs10105363	19440598	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs7001825	20028139	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs952860	20076713	BB	BB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs1874542	20143087	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs2410635	20171132	AB	AB	AB	AB	AB	AB	AB	AB	AB	BB	BB	BB	BB	AB	BB	BB	BB	BB	BB	AA	AB	BB	AA	BB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB
rs2062348	20260042	AB	AB	AB	AB	AB	AB	AB	AB	BB	AB	AB	AB	AB	BB	BB	BB	BB	BB	BB	AB	BB	AB	AB	BB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB
rs7816876	20481982	AB	AB	AB	AB	AB	AB	AA	AB	AB	AA	AB	AA	AA	AB	AB	AB	AA	AA	AB	AA	AB	AA	AB	AA	BB	AA	BB	AA	BB	AA	BB	AA	BB	AA
rs10107413	20517834	AB	AB	AB	AB	AB	AB	AA	AB	AB	AA	AB	AB	AA	AA	AB	AB	AA	AA	AB	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs17093444	20871399	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AB	AA	AA	AA	AB	AA	AA	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs998527	20960271	BB	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs10503701	21164440	AB	AB	AA	AB	AB	AB	AB	AB	AB	AA	AA	AA	AA	AA	BB	AB	AA	AB	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs2959366	21294044	AB	AB	AB	AB	AB	AB	AB	AB	AB	AB	AA	AB	AB	BB	AB	BB	AB	AB	BB	AB	BB	AB	BB	AB	AB	AB	BB	AB	AB	AB	BB	AB	AB	AB
rs4506202	21642357	AB	AB	AB	AB	AB	AB	BB	AB	AB	BB	AB	BB	BB	AB	AA	AB	BB	AA	AB	BB	AA	AB	BB	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs7462141	21769865	BB	BB	BB	BB	BB	BB	BB	BB	AB	BB	AA	BB	AB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs2306646	21902532	AB	AB	AB	AB	AB	AB	AB	AB	BB	AB	BB	AA	AB	BB	BB	AA	BB	BB	AB	AA	BB	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs2309308	22007544	AA	AA	AA	AA	AA	AA	AB	AA	AA	AA	AB	BB	BB	AB	AB	AB	AB	AB	AA	AA	AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs4871834	22706104	AA	AA	AA	AA	AA	AA	AB	AA	AA	AA	AB	AB	AB	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs1543019	22822269	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB
rs4242391	23056128	AB	AB	AB	AB	AB	AB	AA	AB	AB	AA	AA	AA	AA	AB	BB	AB	AA	AB	AA	AB	AA	AB	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs11135720	23206616	AA	AA	AA	AA	AB	AA	AA	AB	AA	AA	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs12386863	23235171	AA	AA	AA	AA	AB	AA	AA	AA	AA	AB	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
rs954705	23268169	AB	AB	AB	AB	AB	BB	BB	BB	AB	AB	AB	AA	BB	AB	AB	AA	BB	AB	AB	AA	AB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB	BB

**Figura 11.** Representação da análise visual dos genótipos de alguns SNPs na região cromossômica 8p21.3 – 22 para os indivíduos da família. A caixa azul abrange a área onde foi observado *linkage* com o modelo fenotípico “BD e EA”, a caixa verde onde se observou *linkage* na análise “manual” dos haplótipos e a caixa vermelha a região de *linkage* observada na análise do modelo “lesões oculares”. A cor amarela representa o genótipo mais comum entre os indivíduos afectados por DB e EA e a cor branca os restantes genótipos, para cada SNP representado.

## 4.6 Pesquisa de mutações

### 4.6.1 Genes candidatos

Como o factor genético na região 8p21.3-22 que predispõe esta família a lesões oculares é, muito provavelmente, o mesmo que causa a DB e a EA, e como este modelo também engloba indivíduos com DB e EA, além de indivíduos com EpAI (**Anexo D6**), decidiu limitar-se a pesquisa de mutações ao pico de *linkage* obtido na análise com o modelo

fenotípico “lesões oculares” e, selecionar-se alguns genes dentro desta região, para serem sequenciados. Esta região compreende 16 genes, descritos na **Tabela 14**. Os genes escolhidos para serem sequenciados foram o *NAT1*, *NAT2* e *CHGN*. Os *NATs* foram selecionados por se saber que o metabolismo dos xenobióticos pode estar envolvido no despoletar de doenças auto-imunes e por o mecanismo de acetilação promovido pelo gene *NAT2* já ter sido associado à doença de Behçet. O *CHGN* foi selecionado porque esta é uma enzima envolvida na síntese da cadeia de sulfato de condroitina, um composto constituinte da cartilagem, que tem sido amplamente fornecido como suplemento dietético a indivíduos com problemas de osteoartrite e doença inflamatória dos intestinos<sup>110-113</sup>.

**Tabela 14.** Informação acerca dos genes selecionados para sequenciação

Genes	Símbolo	Localização	Função biológica	Nº exões	Comprimento do transcrito
Supressor tumoral mitocondrial 1	<i>MTUS1</i>	17,545,584 - 17,702,666	Inativação do receptor de tirosina cinase	20	3.759
Precursor da proteína 1 do tipo fibrinogénio	<i>FGL1</i>	17,766,169 - 17,797,693	Actividade mitogénica	9	936
Proteína do material pericentriolar 1	<i>PCMI</i>	17,824,788 - 17,929,534	Autoantigénio do centrosoma – Envolvimento em carcinomas	51	6.048
Precursor do ácido de ceramidase	<i>ASAH1</i>	17,958,214 - 17,986,757	Síntese do ácido de Ceramidase	13	969
N-acetiltransferase 1	<i>NAT1</i>	18,111,882 - 18,273,466	Metabolização de xenobióticos	2	1.210
N-acetiltransferase 2	<i>NAT2</i>	18,293,035 - 18,302,962	Metabolização de xenobióticos	1	870
Proteína 3 com domínios PH e SEC7	<i>PSD3</i>	18,432,351 - 18,915,476	Participação na substituição de nucleótidos de guanina	13	6.814
Proteína 11 abundante do sistema nervoso	<i>Q86YR2_HUMAN</i>	18,622,695 - 18,636,191	—	6	259
Proteína 4A com domínio SH2	<i>SH2D4A</i>	19,215,408 - 19,298,009	—	11	1.332
Condroitina β-1,4-N-acetilgalactosaminiltransferase 1	<i>CHGN</i>	19,305,953 - 19,584,552	Síntese da cadeia de sulfato de condroitina	9	3.674
CDNA FLJ25648 fis. clone SYN01015	<i>Q8NIG8_HUMAN</i>	19,500,730 - 19,584,552	—	3	474
Subunidade 10 do complexo integrador	<i>INTS10</i>	19,719,277 - 19,753,864	Medeia processamento dos snRNA	14	1.794
Precursor da lipase de lipoproteínas	<i>LPL</i>	19,841,232 - 19,867,912	Regulação dos níveis de triglicéridos no plasma	12	1.410
Transportador de aminas vesicular 1	<i>VAT1</i>	20,046,646 - 20,084,997	Regulador dos sistemas monoaminérgicos	15	1.578
Subunidade B da ATP sintetase vacuolar, isoforma do cérebro	<i>ATP6V1B2</i>	20,098,984 - 20,123,485	Mediador da acidez dos organelos intracelulares	14	1.536
Supressor tumoral com fecho de leucinas 1	<i>LZTS1</i>	20,147,956 - 20,205,754	Regulador mitótica	4	1.152

Nota: Para alguns genes não se obteve informação sobre a função pelo que se colocou um traço.

#### 4.6.2 Sequenciação dos genes candidatos

Para o gene *NAT2* obteve-se um padrão de segregação interessante para os seis indivíduos inicialmente sequenciados (**Anexo E**). Obtiveram-se sequências iguais entre os indivíduos com DB (III:2 e III:5) e EA (III:3 e III:10), diferentes das do indivíduo com EpAI

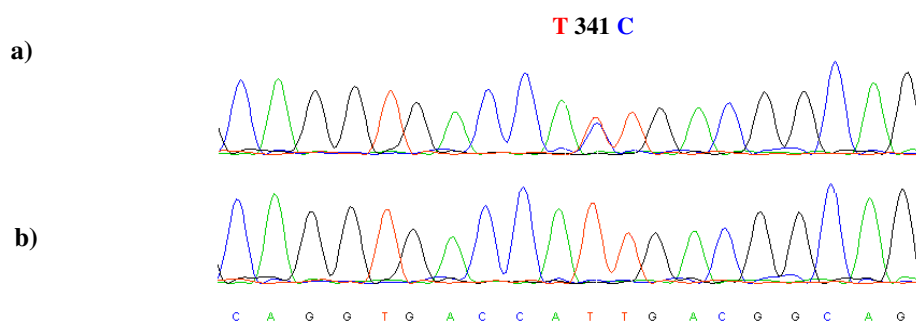
(IV:2) e do não afectado (IV:6), também eles com sequências diferentes entre si. A sequenciação do resto dos elementos da família, revelou a existência de 5 SNPs localizados ao longo do exão codificante deste gene e com elevada variabilidade entre os indivíduos. Os SNPs encontrados e a informação acerca dos mesmos encontram-se descritos na **Tabela 15**. Como se pode verificar, destes 5 SNPs, dois deles eram *missense* codificando o mesmo aminoácido e os restantes três *nonsense*. Na **Fig 12** encontra-se representada uma imagem de sequenciação para o SNP rs1801280, mostrando os genótipos encontrados para dois indivíduos diferentes.

Também se sequenciou o gene *NAT1* em todos os elementos da família, assim como o gene *CHGN* nos seis indivíduos que foram seleccionados para serem primeiro sequenciados, não se tendo encontrado variabilidade entre os indivíduos analisados.

**Tabela 15:** Polimorfismos observados no exão 2 do gene *NAT2* na família 1.

NCBI dbSNP	Localização no cromossoma 8 (pb)	Alteração nucleotídica	Alterações no aminoácido
rs1041983	18302075	C282T	—
rs1801280	18302134	T341C	I114T
rs1799929	18302274	C481T	—
rs1799930	18302383	G590A	R197Q
rs1799931	18302650	G857A	G286E

Nota: O traço na última coluna significa que não existe alteração do aminoácido codificado



**Figura 12.** SNP rs1801280 a) Sequência de um indivíduo heterozigótico TC (IV:2) para o polimorfismo. b) Sequência de um indivíduo homozigótico TT (III:2) sem alteração nucleotídica. A sequência de cDNA está traduzida debaixo dos traços.

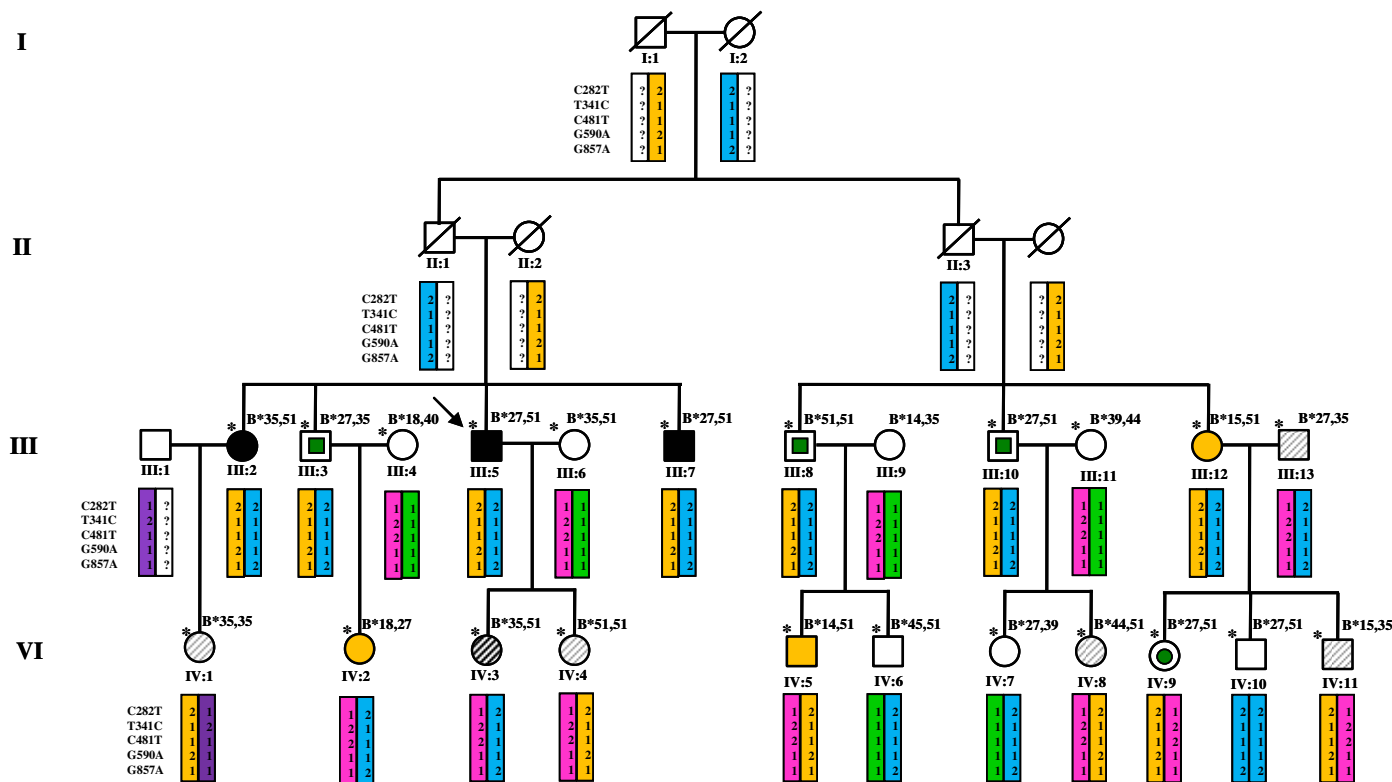
#### 4.7 Haplótipos do gene *NAT2*

A partir da informação proveniente da sequenciação obtida para os cinco SNPs do gene *NAT2*, foi possível a construção dos haplótipos deste gene para todos os indivíduos, de modo a obter-se informação acerca do seu estatuto de acetilação. Na **Tabela 16** encontra-se a informação acerca dos haplótipos encontrados e na **Fig 13** pode-se observar a transmissão dos mesmos. Como a distância entre SNPs era muito reduzida (da ordem dos 100pb) e a família apenas possuía 4 gerações, os programas usados para a análise calcularam os haplótipos considerando a sua transmissão em bloco para a progenia, sem eventos de recombinação entre os diferentes SNPs. Constatou-se a presença de seis haplótipos diferentes para o gene *NAT2*, com predominância de haplótipos que codificam para acetilação lenta, apenas o haplótipo *wild-type* (*NAT2\*4*) codificando para a acetilação rápida nesta família. Pode-se observar que na geração III todos os indivíduos afectados por DB e EA, assim como o indivíduo com EpAI (III:12) (que se tinha verificado na análise “manual” da região 8p21.3-22 ser semelhante a estes indivíduos nesta região) possuem ambos os haplótipos *NAT2\*6A* e *NAT2\*7B*. No entanto, o outro indivíduo com EA (IV:9) assim como os outros indivíduos com EpAI (IV:5 e IV:2) apenas possuem um destes haplótipos. A ocorrência isolada dos haplótipos *NAT2\*6A* e *NAT2\*7B* também se verificou em indivíduos não afectados. Analisando o fenótipo de acetilação, observa-se que todos os indivíduos afectados são acetiladores lentos e entre os indivíduos não afectados existem acetiladores lentos e intermédios.

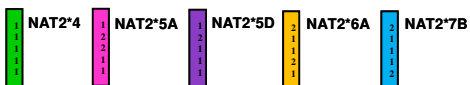
**Tabela 16.** Haplótipos do gene *NAT2* presentes na família, polimorfismos existentes em cada um deles e tipo de acetilação codificada.

Haplótipos do <i>NAT2</i>	C282T	T341C	C481T	G590A	G857A	Tipo de acetilação codificada
<i>NAT2*4</i>	C	T	C	G	G	Rápida
<i>NAT2*5A</i>	C	C	T	G	G	Lenta
<i>NAT2*5D</i>	C	C	C	G	G	Lenta
<i>NAT2*6A</i>	T	T	C	A	G	Lenta
<i>NAT2*7B</i>	T	T	C	G	A	Lenta

Nota: A negrito encontram-se as bases que estão alteradas em comparação com as bases presentes no alelo *wild type* *NAT2\*4*



**Legenda:**



**Figura 13.** Árvore genealógica da família com os alelos do sistema HLA-B e haplótipos do gene *NAT2*. Os membros da família foram sequenciados para cinco SNPs do *NAT2*, representando o número 1 a base existente no haplótipo *wild type* *NAT2\*4* e o número 2 a outra base possível. O haplótipo da esquerda corresponde ao que foi herdado via materna e o da direita via paterna. Os indivíduos incluídos na sequenciação do *NAT2* e genotipagem do HLA-B encontram-se marcados por (\*), tendo sido inferidos os genótipos e os haplótipos para os restantes indivíduos. Os diagnósticos clínicos encontram-se representados de modo semelhante à **Figura 4** (ver legenda da **Figura 4**)

**4.8 Interação entre o gene *NAT2* e o sistema HLA-B**

Quando se tem em consideração a interação dos alelos do sistema HLA-B do complexo MHC com haplótipos específicos do gene *NAT2* (**Fig 13**), não se observa qualquer distinção entre os indivíduos afectados e não afectados. No entanto, quando se tem em conta a interacção entre o sistema HLA-B e o tipo de acetilação promovida por estes haplótipos (ver **Tabela 17** e **Fig 13**), pode-se observar que todos indivíduos afectados por DB, EA e EpAI possuem haplótipos que codificam para uma acetilação lenta e um dos

alelos B\*27 ou B\*51. No que respeita aos indivíduos não afectados, com idade superior a 30 anos, observa-se que existe apenas um indivíduo (IV:10) que possui ambos - a acetilação lenta e um dos alelos HLA-B\*27 ou B\*51 - pois os restantes ou são acetiladores intermédios e têm o alelo B\*27 ou B\*51, ou são acetiladores lentos e não têm nenhum destes alelos. Nos indivíduos de idade inferior a 30 anos, observa-se que três deles (IV:3, IV:4 e IV:8) possuem manifestações incompletas de DB, um fenótipo de acetilação lenta e pelo menos um dos alelos B\*27 ou B\*51 do sistema MHC. O outro familiar jovem (IV:7) não demonstra manifestações patológicas, possui um genótipo de acetilação intermédia e tem o alelo B\*27.

**Tabela 17.** Alelos do sistema HLA-B e fenótipo de acetilação nos indivíduos da família.

Alelos HLA	Não afectados		Afectados	
	Acetilação rápida	Acetilação lenta	Acetilação rápida	Acetilação lenta
<b>B*27 ou 51</b>	3	1	0	10
<b>Outros alelos</b>	3	2	0	0

Nota: Só foram tidos em conta os indivíduos com idade superior a 30 anos na altura da observação.

## 5. DISCUSSÃO

### 5.1 Agregação familiar da DB, da EA e das EpAI

O estudo foi realizado numa família caucasiana com um número elevado de indivíduos afectados, tendo múltiplos casos de DB e doenças do complexo EpA. Embora actualmente só se encontrem duas gerações vivas, também se obteve alguma informação sobre as duas anteriores, o que, aliado ao facto da família ser numerosa e possuir vários indivíduos afectados, com diferentes ligações de parentesco, a tornam robusta para estudos de *linkage*. Devido à rara agregação entre a doença de Behçet e doenças do complexo EpA, esta família representa um objecto de estudo único, e, como tal, é de elevada importância científica o estudo genético da origem destes possíveis fenótipos clínicos.

Nesta família, observa-se antecipação genética para a doença de Behçet, pois sintomas de DB como as aftas orais e a pseudo-foliculite, surgem nas novas gerações numa idade ligeiramente inferior à da geração anterior.

### 5.2 Influência não linear dos alelos do sistema HLA-B

Relativamente à influência dos alelos HLA-B\*27 e B\*51 do complexo MHC na susceptibilidade à doença de Behçet e às Espondilartropatias, respectivamente, não se encontrou uma correlação perfeita nesta família. Esta relação é uma característica epidemiológica de ambas as doenças, como tal, esperava-se que a presença destes alelos em conjunto com outros factores genéticos influenciasse o desenvolvimento das doenças numa determinada direcção. Apesar de esta relação ter sido de algum modo verificada para o alelo B\*27, pois os pacientes portadores do genótipo HLA-B27(+)<sup>51(-)</sup> apresentavam EA ou EpAI, não existindo casos de Behçet, não se verificou esta relação para o alelo B\*51 pois entre os pacientes com genótipo HLA-B27(-)<sup>51(+)</sup> existiam casos de DB, EA e EpAI. Estes resultados apoiam uma associação mais forte entre o alelo B\*27 e as EpA do que entre o alelo B\*51 e a DB, o que vai ao encontro do observado em diversos estudos epidemiológicos<sup>6,33</sup>. O facto de existirem dois indivíduos afectados com DB e dois com EA com o mesmo genótipo HLA-B27(+)<sup>51(+)</sup>, também reforça a não linearidade desta relação. Estes resultados indicam que devem existir outros factores, para além do complexo MHC, a

influenciar a susceptibilidade para as doenças e que, através de uma interacção com este ou de um efeito independente, determinam o desenvolvimento das doenças numa determinada direcção. Este factor poderá ser um factor ambiental, mas a hipótese mais plausível é que seja um factor genético, já que se trata de uma única família, vivendo os seus elementos em regiões geográficas variadas.

Foi referido, num estudo recente, que a presença simultânea num indivíduo dos alelos HLA-B\*27 e B\*51 seria protectora a nível das uveítes na doença de Behçet, não desenvolvendo os portadores manifestações tão severas como os indivíduos HLA-B27(-)51(+)<sup>114</sup>. Nesta família, existem dois pacientes com DB portadores de genótipo HLA-B\*27(+)-51(+) e um doente com genótipo HLA-B27(-)51(+), tendo todos eles lesões oculares. Entre os indivíduos positivos para os dois alelos, um deles tem uma perda de visão de 90%, devido a uma uveíte posterior, estando quase cego, e o outro possui uma uveíte anterior e posterior biocular, o que demonstra um padrão grave de lesões oculares. O paciente HLA-B27(-)51(+) possui história de edema da papila, do qual não resultaram complicações sérias. Deste modo, embora esta não seja de todo uma amostra significativa e não se pretenda extrapolar conclusões destes dados, é de referenciar que não se observa o padrão verificado pelos autores nesse estudo.

Teria sido interessante observar os subtipos de alelos do HLA-B que existem na família. Poderia ajudar a elucidar a relação dos alelos com as doenças na família, já que se sabe que alguns alelos estão a elas maioritariamente associados<sup>4,52</sup>, e também porque o HLA-B\*2702 já foi relacionado com Behçet. Isto poderia explicar o porquê de alguns familiares com genótipo HLA-B\*27,51 desenvolverem DB e outros EA.

### **5.3 Estudo de *linkage***

#### **5.3.1 Limitações do programa usado**

Na análise de *linkage* não foi possível usar a informação proveniente da família toda em simultâneo, pois como foi referido, o Merlin, possui uma limitação relativamente ao número máximo de indivíduos por família que se consegue analisar de cada vez. Para a realização de análises NPL, existem outros programas que conseguem lidar com famílias

mais numerosas, tal como o Simwalk. No entanto, este programa não permite trabalhar com informação proveniente de tantos marcadores, pelo que se considerou o Merlin a melhor opção, apesar da já referida limitação. Houve, deste modo, alguma perda de informação, que provavelmente teria contribuído para aumentar os LOD scores nas regiões de *linkage*. Contudo, considera-se que esta perda foi minimizada, através de uma escolha criteriosa dos indivíduos incluídos em cada análise.

### 5.3.2 Ausência de *linkage* com o sistema MHC

Não constituiu surpresa o facto de não se ter observado um pico de *linkage* no cromossoma 6p21, região em que se localiza o sistema MHC-HLA-B e que tem sido consistentemente ligada à DB e às EpA, com LOD *scores* muito elevados, nos diferentes estudos de *linkage* já realizados<sup>65-67</sup>. Como se referiu nos resultados, não se observou uma perfeita associação entre os alelos deste sistema e as patologias existentes nesta família, pois não é o mesmo alelo do sistema HLA-B que está a ser transmitido aos pacientes com DB, EA e EpAI, tal como estabelecido pela genotipagem do HLA-B. Resultados semelhantes foram obtidos noutros estudos familiares na DB em que se observou a presença do alelo HLA-B\*51, tantos nos indivíduos afectados como nos não afectados<sup>115,116</sup>.

### 5.3.3 Evidência de *linkage* na região 8p21.3-22

Já na região 8p21.3-22, obteve-se um LOD *score* de 3.59 ( $p=2.0 \times 10^{-5}$ ), o que, de acordo com as linhas de orientação para avaliação dos resultados de *linkage*<sup>117</sup>, é um resultado significativo que nos permite concluir que reside um *locus* de susceptibilidade para a DB e EA nesta região. O facto de nas análises de transmissão individual dos fenótipos “DB” e “EA” não se ter obtido qualquer resultado significativo, reforça a ideia destas doenças partilharem um factor comum que é “silenciado” quando se faz a classificação “errada” dos indivíduos como afectados/não afectados. Por outro lado, ao fazer a análise em separado das doenças, considera-se um número reduzido de indivíduos afectados, o que se traduz em poder insuficiente para detectar *linkage*, mesmo que este exista.

Como esperado para uma análise de *linkage* numa única família<sup>118</sup>, a região de *linkage* mapeada para a BD e EA é bastante vasta, albergando um elevado número de genes.

Na análise de *linkage* para diferentes modelos fenotípicos, observaram-se LOD scores elevados em zonas que englobavam um número muito reduzido de marcadores ou mesmo em marcadores isolados. Como se está a analisar apenas uma família, sabe-se que o desequilíbrio de *linkage* atinge zonas extensas do genoma, pelo que, um número tão pequeno de marcadores ligados deverá ser indicador de um falso resultado positivo e, como tal, não merecedor de confiança. Deste modo, para ser considerado um pico significativo estipulou-se que para além do valor de LOD score > 3.0, o intervalo teria de ter no mínimo 4 cM, ou vários marcadores consecutivos em *linkage*. O pico de *linkage* obtido no cromossoma 8p21.3-22, para a análise “lesões oculares”, é pouco maior que este limiar (tem 4.5 cM) e não atingiu uma estatística significativa (LOD score = 2.47), apenas sugestiva. Como tal, caso este se tivesse tratado de um resultado isolado, como aconteceu com o pico de *linkage* obtido para o modelo aftas orais no cromossoma 8 entre os 117 – 119.8 cM, não se lhe teria atribuído importância biológica. No entanto, como este resultado reforçou o já obtido na análise do modelo “BD e EA”, assim como o obtido para o modelo “aftas orais” e foi confirmado na análise visual dos haplótipos, decidiu-se considerá-lo válido e limitar, inclusivamente, a procura de genes candidatos a esta região.

Teria sido muito interessante realizar-se a análise conjunta da transmissão da EA e das EpAI assim como da DB, EA e EpAI, pois poderia ajudar a suportar e a esclarecer os resultados obtidos. Se estas três doenças estão a segregar em conjunto nesta família, então, esperar-se-ia observar um LOD score semelhante ao do modelo “BD e EA”, com a primeira análise referida (pois existe o mesmo número de indivíduos afectados com DB e EpAI), e um resultado mais significativo quando se fizesse a segunda análise referida, pois estar-se-ia a atribuir o fenótipo afectado de maneira mais correcta.

#### **5.3.4 Diferentes modelos estatísticos aplicados**

Registaram-se mais discrepâncias entre as análises de *linkage* com as estatísticas  $S_{\text{pairs}}$  e  $S_{\text{all}}$ , quando se analisou individualmente a transmissão de sintomas característicos de Behçet (com as aftas orais, manifestações oculares e lesões cutâneas), do que quando se analisou a transmissão da DB e EA separadamente e em conjunto. Como  $S_{\text{pairs}}$  mede o número de alelos IBD partilhados por pares de parentes afectados, enquanto a estatística  $S_{\text{all}}$

captura informação tendo em conta a partilha entre grupos grandes de parentes afectados<sup>106</sup>, era esperado que para os modelos “BD” e “EA” não existisse muita diferença entre as duas estatísticas, já que em ambos os modelos se considera um pequeno número de indivíduos afectados. Relativamente ao facto de os resultados do modelo “DB e EA” serem mais consistentes do que os dos restantes modelos, isto deve-se a uma elevada consistência nos alelos IBD em todos os parentes afectados pela DB e pela EA, o que se traduz em resultados semelhantes quando se analisa par a par os indivíduos afectados, ou quando se tem em conta o conjunto deles. Já para os modelos em que se considera características isoladas da DB, observa-se que existe uma maior partilha de alelos IBD entre cada par de indivíduos do que entre o grupo todo, o que é natural, pois, embora os indivíduos apresentem em comum certas características clínicas como as aftas orais, ou as lesões cutâneas, possuem um quadro clínico bastante distinto.

Em relação aos modelos linear e exponencial de Kong and Cox, verificou-se que só com o modelo exponencial foram obtidos resultados significativos. O modelo linear foi desenhado para identificar um pequeno aumento na partilha de alelos entre indivíduos afectados, distribuídos por um grande número de famílias. Quando se está à procura de um aumento na partilha de alelos entre indivíduos afectados, num pequeno número de famílias, é preferível seleccionar o modelo exponencial de Kong e Cox uma vez que este providencia um melhor teste de *linkage*<sup>100</sup> sendo, por isso, o modelo que melhor se ajusta à família em estudo.

#### **5.4 Inexistência de LOH**

A análise de LOH indicou que o factor genético responsável pelos fenótipos na família não era uma deleção cromossómica heterozigótica. A perda de heterozigotia ocorre quando há deleções num dos cromossomas e reflecte, para cada amostra, zonas cromossómicas em que existe menos variabilidade do que a esperada, quando comparada com outras amostras. Este fenómeno pode indicar que apenas se está a observar a informação proveniente de um dos cromossomas, pois o outro encontra-se delectado. Esta análise foi muito importante pois caso existisse uma deleção num dos cromossomas (na zona de *linkage*) a sequenciação não deveria ser a técnica usada para detecção dessa

mutação, uma vez que apenas detecta deleções se estas ocorrerem em ambos os cromossomas (deleções homozigóticas). Nestes casos de agregação familiar de doenças, muitas vezes<sup>120-122</sup> o factor genético responsável é uma deleção heterozigótica. Deste modo, consideramos ser muito importante realizar esta análise antes de se despender recursos financeiros numa técnica como a sequenciação. Uma alternativa à sequenciação, nestes casos, seria a realização de um PCR em tempo real, metodologia que permitiria quantificar e qualificar as cadeias de DNA, mas que seria, no entanto, mais laboriosa.

### **5.5 Influência do NAT2 e do HLA-B na susceptibilidade à DB e às EpA**

O genótipo de acetilação lenta promovido pelo *NAT2*, revelou que por si só não era determinante para o fenótipo clínico dos elementos da família. No entanto, quando se verificou o seu efeito conjunto com a presença do alelo B\*27 e/ou B\*51 do sistema MHC classe I, constatou-se a existência de um padrão associado não só à DB e à EA, mas também às EpAI. Também se verificou que este padrão era uma característica das formas incompletas de DB em indivíduos de idade inferior a 30 anos. Contudo, nada se pode inferir dos resultados observados para estes indivíduos uma vez que estes ainda poderão vir, ou não, a desenvolver doença de Behçet, motivo pelo qual não entraram no modelo fenotípico “BD e EA” e “DB”.

Estes resultados suportam a ideia de que nesta família a presença do fenótipo de acetilação lenta, juntamente com um dos alelos B\*27 ou B\*51 do sistema HLA, seriam por si só suficientes para se desenvolver um quadro clínico de doença. Todavia, um dos indivíduos não afectados (IV:10), que possuía genótipo de acetilação lenta, não tinha apresentado sintomas de DB até à data do estudo. Atendendo à sua idade, 32 anos, ainda poderá vir a desenvolver a doença, até porque há casos na família em que o primeiro sintoma só se manifestou aos 35 anos (III:2). Também pode acontecer que, por interacção com outros factores, este indivíduo, apesar de genotipicamente codificar para acetilação lenta, tenha um fenótipo de acetilação intermédia ou rápida, pois, embora exista uma elevada correlação entre o fenótipo e o genótipo, pode haver discordâncias<sup>79,122</sup>. Foi inclusivamente referido num estudo que os fenótipos de indivíduos com haplótipos *NAT2*\*7 (presente neste indivíduo) necessitam de ser interpretados com precaução, uma vez que o

tipo de substracto e a sua concentração, que são expostos à enzima, os podem modificar<sup>123</sup>. Deste modo, seria muito interessante confirmar-se através de uma análise bioquímica, o estatuto fenotípico do indivíduo referido, assim como de alguns outros elementos da família. Esta análise também poderia servir para verificar se existe um fenótipo de acetilação único nos pacientes com DB derivado da interacção com o HLA-B\*51, tal como foi sugerido por Najim *et al.*<sup>93</sup>.

Os polimorfismos da acetilação xenobiótica desempenham um papel essencial na etiopatogénese de várias doenças, incluindo aquelas que são induzidas por factores ambientais<sup>79</sup> tendo sido já o fenótipo de acetilação lenta, associado a algumas doenças auto-imunes e inflamatórias na população<sup>85-89</sup>. Isto verifica-se, porque os xenobióticos não acetilados acumulam-se nos indivíduos que possuem um estatuto de acetilação lento e, consequentemente, são metabolizados por outras enzimas em reactivos intermédios. Estes compostos intermédios reactivos podem alterar proteínas do próprio (*self*) que se tornam estranhas ao organismo, sendo apresentadas ao sistema imunitário e estimulando as células T. Estas, por sua vez, iniciam os sinais clínicos e patológicos da auto-imunidade<sup>37</sup>. Para a DB já foram publicados alguns estudos, cujas conclusões não foram, no entanto, consistentes<sup>37,92,93</sup>. Relativamente às EpA ou EA em particular, não foi publicado qualquer estudo que sugerisse uma associação. As enzimas metabolizadoras de xenobióticos constituem uma das primeiras linhas de defesa contra químicos ambientais. Deste modo, os indivíduos que herdaram os alelos de risco para estas enzimas podem falhar na protecção contra carcinogénios do fumo de tabaco, dieta, processos industriais e poluição ambiental<sup>124</sup>.

É possível que estes factores ambientais conjuntamente com o fenótipo de acetilação lenta e com os alelos B\*27 e B\*51, nos indivíduos portadores, possam induzir mecanismos auto-imunes, devido à interacção dos xenobióticos não acetilados, ou das proteínas alteradas a estes associados, com os alelos do sistema HLA-B, contribuindo para o desenvolvimento da DB assim como de EpA<sup>37,92</sup> nesta família. Deste modo, o *NAT2* pode ser um dos factores genéticos que predis põem para o despoletar e/ou desenvolvimento destas doenças auto-imunes, estabelecendo, de algum modo, uma ligação entre os factores ambientais e outros factores genéticos, embora a sua contribuição para as doenças em si seja pequena.

Estes resultados foram surpreendentes pois partiu-se, desde o início, da hipótese da DB e EA partilharem um factor genético comum, tendo-se menosprezado os pacientes com

EpAI por se ter considerado terem problemas articulares de origem diferente da DB e EA. Doenças que envolvem problemas reumatológicos podem tornar-se problemáticas no que respeita à escolha do fenótipo, pois o diagnóstico frequentemente depende de vários sintomas distintos, não havendo, muitas vezes, acordo em relação ao que constitui um diagnóstico definitivo. A ausência de uma definição clara do fenótipo leva a incertezas acerca da classificação como afectados e não afectados e a potenciais inconsistências entre estudos<sup>101</sup>.

A família em estudo é muito heterogénea, o que levou a uma opção errada na escolha do fenótipo, pois ao considerar-se conjuntamente como afectados os indivíduos com doença de Behçet e espondilite anquilosante – na *análise* “DB e EA”, também se deveria ter considerado as Espondilartropatias indiferenciadas. Relativamente às manifestações de Behçet incompletas, como nesta família ocorrem maioritariamente em indivíduos de idade inferior a 30 anos, que não entraram na análise de *linkage* “BD e EA”, este factor acabou por não afectar significativamente os resultados. No entanto, apesar da incorrecta classificação como afectados/não afectados na análise principal, obteve-se uma elevada consistência na análise deste modelo com os modelos “aftas orais” e “lesões oculares”, embora estes não tenham tido uma significância estatística tão elevada. Estes resultados indicam que muito provavelmente este erro não provocou um enviesamento dos dados, tendo apenas diminuído o valor de LOD *score*.

Esperava encontrar-se um factor genético responsável pelos fenótipos clínicos, mais forte e determinante, tal como uma mutação. No entanto, encontrou-se um factor que é comum na população em geral, mas que nesta família, e em conjunto com alelos específicos do sistema HLA, é determinante no desenvolvimento das respectivas doenças. Sugere-se deste modo, a existência de outros factores genéticos, para além do *NAT2* e do complexo HLA-B, que determinam a susceptibilidade para cada uma das doenças em particular. Pelo facto de existirem poucos indivíduos afectados por DB e por EA, as análises individuais de *linkage* das doenças, não tiveram poder suficiente, no entanto, para reflectir esses factores.

Por outro lado, não se pode excluir a hipótese de o factor genético responsável pelos fenótipos clínicos existentes na família ser um gene em desequilíbrio de *linkage* com o *NAT2*. Pode-se observar que na geração III todos os indivíduos são afectados por DB, EA ou EpAI e todos possuem o genótipo de acetilação *NAT2\*6A/NAT2\*7B*. Uma explicação

alternativa poderia ser a existência de um factor genético com hereditariedade autossómica dominante, com penetrância incompleta, que estivesse a ser transmitido ligado aos alelos *NAT2\*6A* e *NAT2\*7B*. Deste modo, a descendência na geração IV, que herda um destes alelos (*NAT2\*6A* ou *NAT2\*7B*) através dos progenitores afectados, herda também este factor de penetrância incompleta que, por interacção com os alelos B\*27 ou B\*51, determina a susceptibilidade.

Os nossos resultados apontam, assim, para a existência de um novo determinante para DB e EA, numa região em desequilíbrio de *linkage* no cromossoma 8p21.3-22, ainda que não seja o *NAT2* por si só. Uma maneira de provar que o *NAT2* era o responsável sozinho pelo *linkage* observado, seria mostrar que os genes flanqueantes não estavam também a segregar com a doença. O gene a montante, o *NAT1* foi sequenciado para todos os indivíduos e não possuía variabilidade, como tal não é informativo, não sendo possível provar que a segregação está limitada ao *NAT2* via montante. A jusante do *NAT2*, só foi sequenciada a *CHGN*, que também não evidenciou variabilidade, pelo menos para os 6 indivíduos testados.

Para ajudar a esclarecer as conclusões tiradas, seria interessante realizar uma análise de *linkage* tendo em conta as variáveis presença dos alelos HLA-B\*27 ou B\*51 e o tipo de acetilação promovida pelo *NAT2*. Também se poderia experimentar fazer uma análise de *linkage* paramétrica, considerando um modelo autossómico dominante com penetrância incompleta com estes factores.

## 6. CONCLUSÃO

Neste trabalho descreve-se uma família portuguesa com múltiplos indivíduos afectados por doença de Behçet, espondilite anquilosante e Espondilartropatias indiferenciadas. A ocorrência de características clínicas partilhadas e a co-associação de múltiplas EpA na mesma família e, ocasionalmente, no mesmo sujeito, sugeria a existência de factores genéticos comuns que predispunham para as doenças dentro deste complexo<sup>7</sup>. Neste estudo, colocou-se a hipótese de também a doença de Behçet partilhar factores genéticos com as EpA e decidiu-se investigar o possível factor comum.

O estudo de *linkage* no genoma mapeou o factor genético responsável pelos fenótipos DB e EA à região cromossómica 8p21.3-22. A análise dos genes localizados nesta região possibilitou a detecção da existência de um estatuto de acetilação lenta promovida pelo gene *NAT2*, para além da presença dos alelos HLA-B\*27 e/ou B\*51, nos indivíduos afectados por DB, EA e EpAI, assim como em pacientes com manifestações características da DB mas com idade inferior a 30 anos. Estes resultados sugerem que a interacção entre o fenótipo de acetilação lenta promovido pela enzima *NAT2*, e os alelos B\*27 ou B\*51 do sistema HLA, pode ser determinante no desenvolvimento de doenças inflamatórias e auto-imunes em organismos geneticamente susceptíveis, salientando a importância de se estudar a interacção entre diversos factores genéticos (e se possíveis ambientais) na susceptibilidade a doenças complexas.

Este é um estudo pioneiro, já que nunca antes foi referida na literatura a possibilidade destas doenças possuírem uma base genética comum, nunca se tendo investido no sentido de identificar factores genéticos comuns. De igual modo, não existem quaisquer outros registos de agregação familiar de DB e EA, o que torna esta família única e, deste modo, de grande importância científica a realização do estudo. Deste estudo surgem eventuais implicações éticas, pois de acordo com os resultados obtidos pode-se concluir que é muito provável que os indivíduos da geração IV, com idade inferior a trinta anos e que possuem os factores mencionados, ainda venham a desenvolver a doença de Behçet. Estas inferências a serem confirmadas noutros estudos, poderiam servir como prognóstico para os indivíduos mais jovens dentro desta família e em famílias com um padrão clínico

semelhante. Também seria interessante, reavaliar estes indivíduos através de um estudo longitudinal, quando estes atingirem uma idade mais avançada.

Aproximadamente 4% da população, nos países industrializados, é afectada por doenças auto-imunes, e esta frequência parece estar a aumentar. Muitas destas doenças são severas, crónicas e debilitantes e para a maioria o tratamento disponível é inadequado e a cura permanente impossível de obter. Assim sendo, um conhecimento mais profundo dos mecanismos moleculares que levam à imunidade destrutiva é crucial, de modo a que se possam desenvolver melhores métodos de prevenção e tratamento deste grupo importante de doenças. A identificação dos genes que predispõem para as doenças, e das suas funções biológicas, são passos iniciais necessários a este conhecimento. Seria desejável, em última análise, não apenas identificar as pessoas que estão em maior risco de desenvolver doenças auto-imunes, mas também proporcionar meios mais específicos e eficientes de intervenção a nível molecular, antes que a destruição das células ou tecidos se torne demasiado extensa<sup>125</sup>.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

1. Oliveira SP, Horta AB, Serra MJ & Castro AS; *Doença de Behçet – experiência de um Serviço de Medicina Interna*; *Medicina Interna* **8**(3), 123-126 (2001)
2. Sieper J, Rudwaleit M, Khan MA & Braun J; *Concepts and epidemiology of spondyloarthritis*; *Best Pract Res Clin Rheumatol* **20**(3), 401–417 (2006)
3. Al-Otaibi LM, Porter SR & Poate TWJ; *Behçet's Disease: A Review*; *J Dent Res* **84**(3), 209-222 (2005)
4. Marshall SE; *Behçet's disease*; *Best Practice & Research Clinical Rheumatology* **18**(3), 291–311 (2004)
5. Gur A, Sarac AJ, Burkan YK, Nas K & Cevik R; *Arthropathy, quality of life, depression, and anxiety in Behçet's disease: relationship between arthritides and these factors*; *Clin Rheumatol* **25**(4), 524-31 (2006)
6. Olivieri, I., van Tubergen, A., Salvarani, C. & van der Linden, S.; *Seronegative spondyloarthritides*; *Best Practice & Research Clinical Rheumatology* **16**(5), 723-739 (2002)
7. Martínez-Borra J, González S & Larrea-López C; *Genetic Factors Predisposing to Spondylarthropathies*; *Arthritis & Rheumatism* **43**(3), 485-492 (2000)
8. Khan MA; *An overview of clinical spectrum and heterogeneity of spondyloarthropathies*; *Rheum Dis Clin North Am* **18**(1), 1-10 (1992)
9. Lau CS, Burgos-Vargas R, Louthrenoo W, Mok MY, Wordsworth P, Zheng QY; *Features of spondyloarthropathies around the world*; *Rheum Dis Clin North Am* **24**(4), 753-70 (1998)
10. Bluestein HG; *Anatomic and genetic considerations in the pathogenesis of ankylosing spondylitis*; *J Rheumatol Suppl.* **16**, 29-32 (1988)
11. Peixoto MJ, Gonzales T, Spínola H, Couto AR, Mora MG, Brehm A, & Santos MR; *Polimorfismos do Alelo HLA-B27 no Desenvolvimento das Espondilartropatias*; *Acta Méd Port* **18**, 283-294 (2005)
12. Smith JA, Märker-Hermann E & Colbert RA; *Pathogenesis of ankylosing spondylitis: Current concepts*; *Best Practice & Research Clinical Rheumatology* **20**(3) 571–591 (2006)

13. Rocha FG; *Manual da Espondilite (Anquilosante)*; ANEA, Associação Nacional da Espondilite Anquilosante; Secretariado Nacional para a Reabilitação e Integração das Pessoas com Deficiência (2002)
14. Braun J & Pincus T; *Mortality, course of disease and prognosis of patients with ankylosing spondylitis*; Clin Exp Rheumatol **20**, S16-S22 (2002)
15. Peters MJ, van der Horst-Bruinsma IE, Dijkmans BA & Nurmohamed MT; *Cardiovascular Risk Profile of Patients with Spondyloarthropathies, Particularly Ankylosing Spondylitis and Psoriatic Arthritis*; Semin Arthritis Rheum **34**, 585-592 (2004)
16. Sieper J, Braun J, Rudwaleit M, Boonen A & Zink A; *Ankylosing spondylitis: an overview*; Ann Rheum Dis **61**, iii8-iii18 (2002)
17. Liao HT, Lin KC, Chen CH, Liang TH, Lin MW, Tsai CY, Tak Yan Yu D & Chou CT; *Human leukocyte antigens in undifferentiated spondyloarthritis*; Semin Arthritis Rheum **37**(3), 198-201 (2007)
18. Zochling J, Brandt J & Braun J; *The current concept of spondyloarthritis with special emphasis on undifferentiated spondyloarthritis*; Rheumatology **44**, 1483-1491 (2005)
19. Elyan M & Khan MA; *Diagnosing ankylosing spondylitis*; J Rheumatol Suppl **78**, 12-23 (2006)
20. International Study Group for Behçet's Disease; *Criteria for diagnosis of Behçet's disease. International Study Group for Behçet's Disease*; Lancet **335**(8697), 1078-80 (1990)
21. Schirmer M, Baltaci M, Isayeva A, Dong Y, Zhang Y, Assaad-Khalil S, Wechsler B, Zouboulis CC, Altenburg A, Kaklamanis Ph, Kumar A, Davatchi F, Sadeghi-Abdollahi B, Shahram A, Nadji A, Chams-Davatchi C, Shams H, Zia'I N, Akhlagi M, Jamshidi AR, Haghdoost AK, Sharquie K, Hayani R, Ben-Chetrit E, Olivieri I, Salvarani C, Pipitone N, Ohno Sh, Namba K, Madanat W, Elmuntaser K, Benamour S, Ali A, Crespo J, Tribolet T, Vasconcelos C, Correia J, Carvalho L, Bastos M, Serra MJ, Resende C, Ramos F, Rosa M, Queirós V, Vedes J, Dias C, Patto JV, Duarte FP, Alekberova Z, Elonakov A, Al Dalaan A, Yew Kuang C, Grana Gil G, Chen WC, Emvalee A, Houman H, Ben Ghorbel I, Sliti-Khanfir M, Boyvat A, Calamia K & Davatchi F; *Revision of the International Criteria for Behçet's disease (ICBD)*; 12th International Conference on Behçet's disease, Lisboa, Portugal, 20-23 Setembro (2006)

22. van der Linden S, Valkenburg HA & Cats A; *Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria*; *Arthritis Rheum* **27**(4), 361-8 (1984)
23. Dougados M, van der Linden S, Juhlin R et al. *The European Spondylarthropathy Study Group. Criteria for the classification of spondylarthropaty*; *Arthritis and Rheumatism* **34**, 1218-1227 (1991)
24. Clegg DO; *Treatment of ankylosing spondylitis*; *J Rheumatol Suppl.* **78**, 24-31 (2006)
25. Khan MA; *Ankylosing spondylitis: a dual perspective of current issues and challenges*; *J Rheumatol Suppl* **78**, 1-3 (2006)
26. Chang HK, Lee DH, Jung SM, Choi SJ, Kim JU, Choi YJ, Baek SK, Cheon KS, Cho EH & Won KS; *The Comparison Between Behçet's Disease and Spondyloarthritides: Does Behçet's Disease Belong to the Spondyloarthropathy Complex?*; *J Korean Med Sci* **17**, 524-529 (2002)
27. Kotevoglu N; *Coexistence of ankylosing spondylitis and behçet's disease. Two cases with atypical presentation and course*; *Scand J Rheumatol* **32**, 184-185 (2003)
28. Etaouil N, Benyahya E, Bennis R & Mkinsi O; *Ankylosing spondylitis and Behçet's disease in combination. Two case reports*; *Joint Bone Spine* **69**(1), 96-98 (2002)
29. Nadji A, Davatchi F, Shahram F, Jamshidi A, Chams C, Akbarian M, Gharibdoost F & Mowla K; *The comparison of ankylosing spondylitis in Behçet's Disease and normal population. A control study*; *Adv Exp Med Biol.* **528**, 449-53 (2003)
30. Yazici H, Tuzlaci M & Yurdakul S; *A controlled survey of sacroiliitis in Behçet's disease*; *Ann Rheum Dis* **40**, 558-9 (1981)
31. Chamberlain MA & Robertson RJ; *A controlled study of Sacroillitis in Behçet's disease*; *Br J Rheumatol* **32**, 693-8 (1993)
32. Shimizu T, Ehrlich GE, Inaba G & Hayashi K; *Behçet's disease (Behçet's syndrome)*; *Semin Arthritis Rheum* **8**, 223-60 (1979)
33. Al-Mutawa SA & Hegab SM; *Behçet's disease*; *Clin Exp Med.* **4**, 103-131 (2004)
34. Fietta P; *Behçet's disease: Familial clustering and immunogenetics – Review*; *Clin Exp Rheumatol.* **23**, S96-S105 (2005)
35. Breban M; *Genetics of spondyloarthritis*; *Best Pract Res Clin Rheumatol* **20**(3), 593-9 (2006)

36. Verity DH, Wallace GR, Vaughan RW & Stanford MR; *Behçet's disease: from Hippocrates to the third millennium*; Br J Ophthalmol **87**, 1175-1183 (2003)
37. Tamer L, Tursen U, Eskandari G, Ates NA, Ercan B, Yildirim H, Atik U; *N-acetyltransferase 2 polymorphisms in patients with Behçet's disease*; Clin Exp Dermatol. **30**(1), 56-60 (2005)
38. Sims AM, Wordsworth BP & Brown MA; *Genetic Susceptibility to Ankylosing Spondylitis*; Current Molecular Medicine **4**, 13-20 (2004)
39. Fresko I, Soy M, Hamuryudan V, Yurdakul S, Yavuz S, Tümer Z & Yazici H; *Genetic anticipation in Behçet's syndrome*; Ann Rheum Dis **57**(1), 45-8 (1998)
40. Gül A, Inanç M, Ocal L, Aral O & Koniçe M; *Familial aggregation of Behçet's disease in Turkey*; Ann Rheum Dis **59**(8), 622-5 (2000)
41. O'Donovan M, Jones I & Craddock N; *Anticipation and repeat expansion in bipolar disorder*; Am J Med Genet C Semin Med Genet **123**(1),10-17 (2003)
42. Lindblad K & Schalling M; *Expanded repeat sequences and disease*; Semin Neurol. **19**(3), 289-99 (1999)
43. de Blecourt J, Polman A & de Blecourt-Meindersma; *Hereditary factors in rheumatoid arthritis and ankylosing spondylitis*; Ann Rheum Dis **20**, 215-20 (1961)
44. Brown MA; *Non-major-histocompatibility-complex genetics of ankylosing spondylitis*; Best Practice & Research Clinical Rheumatology **20**(3), 611–621 (2006)
45. Breban M; *Genetics of spondyloarthritis*; Best Practice & Research Clinical Rheumatology **20**(3), 593-599 (2006)
46. Blecourt J, Polman A & Blecourt-Meindersma T; *Hereditary factors in rheumatoid arthritis and ankylosing spondylitis*; Ann Rheum Dis **20**, 215-220 (1961)
47. Gül A, Uyar FA, Inanç M, Ocal L, Barrett JH, Aral O, Koniçe M & Saruhan-Direskeneli G; *A weak association of HLA-B\*2702 with Behçet's disease*; Genes Immun **3**(6), 368-372 (2002)
48. Takeno M, Kariyone A, Yamashita N, Takiguchi M, Mizushima Y, Kaneoka H & Sakane T; *Excessive function of peripheral blood neutrophils from patients with Behçet's disease and from HLA-B51 transgenic mice*; Arthritis Rheum **38**(3), 426-433 (1995)

49. Mau W, Zeidler H, Mau R, Majewski A, Freyschmidt J, Stangel W & Deicher H; *Clinical features and prognosis of patients with possible ankylosing spondylitis. Results of a 10-year followup*; J Rheumatol **15**(7), 1109-1114 (1988)
50. Khan MA; *Update on spondyloarthropathies*; Ann Intern Med. **136**(12), 896-907 (2002)
51. Prakash S, Mehra NK, Bhargava S & Malaviya AN; *HLA B27 related 'unclassifiable' seronegative spondyloarthropathies*; Ann Rheum Dis **42**(6), 640-643 (1983)
52. Brown MA, Pile KD, Kennedy LG, Calin A, Darke C, Bell J, Wordsworth BP & Cornélis F; *HLA class I associations of ankylosing spondylitis in the white population in the United Kingdom*; Ann Rheum Dis **55**(4), 268-270 (1996)
53. Hammer RE, Maika SD, Richardson JA, Tang JP & Taurog JD; *Spontaneous inflammatory disease in transgenic rats expressing HLA-B27 and human beta 2m: an animal model of HLA-B27-associated human disorders*; Cell **63**(5), 1099-1112 (1990)
54. Ahmad T, Wallace GR, James T, Neville M, Bunce M, Mulcahy-Hawes K, Armuzzi A, Crawshaw J, Fortune F, Walton R, Stanford MR, Welsh KI, Marshall SE & Jewell DP; *Mapping the HLA association in Behçet's disease: a role for tumor necrosis factor polymorphisms?*; Arthritis Rheum **48**(3), 807-13 (2003)
55. Choukri F, Chakib A, Himmich H, Hüe S & Caillat-Zucman S; *HLA-B\*51 and B\*15 alleles confer predisposition to Behçet's disease in Moroccan patients*; Hum Immunol **62**(2), 180-5 (2001)
56. Fallingborg J, Ambrosius Christensen L & Grunnet N; *HLA antigens in a family with Behçet's syndrome*; Acta Med Scand **220**(4), 375-8 (1986)
57. Müller CA, Engler-Blum G, Gekeler V, Steiert I, Weiss E & Schmidt H; *Genetic and serological heterogeneity of the supertypic HLA-B locus specificities Bw4 and Bw6*; Immunogenetics **30**(3), 200-7 (1989)
58. Robinson WP, van der Linden SM, Khan MA, Rentsch HU, Cats A, Russell A & Thomson G; *HLA-Bw60 increases susceptibility to ankylosing spondylitis in HLA-B27+ patients*. Arthritis Rheum **32**(9), 1135-41 (1989)
59. Vargas-Alarcón G, Londoño JD, Hernández-Pacheco G, Pacheco-Tena C, Castillo E, Cardiel MH, Granados J & Burgos-Vargas R; *Effect of HLA-B and HLA-DR genes on susceptibility to and severity of spondyloarthropathies in Mexican patients*; Ann Rheum Dis **61**(8), 714-7 (2002)

60. Ott J & Bhat A; *Linkage analysis in heterogeneous and complex traits*; Eur Child Adolesc Psychiatry **8** (Suppl 3), 43-46 (1999)
61. Taere MD & Barret JH; *Genetic Epidemiology 2. Genetic linkage studies*; Lancet **366**, 1036-44 (2005)
62. Carlson CS, Eberle MA, Kruglyak L & Nickerson DA; *Mapping complex disease loci in whole-genome association studies*; Nature **429**, 446-452 (2004)
63. Teltsh O, Kanyas K, Karni O, Levi A, Korner M, Ben-Asher E, Lancet D, Hamdan A, Lerer B & Kohn Y; *Genome-wide linkage scan, fine mapping, and haplotype analysis in a large, inbred, Arab Israeli pedigree suggest a schizophrenia susceptibility locus on chromosome 20p13*; Am J Med Genet B Neuropsychiatr Genet **147**(2), 209-15 (2008)
64. Brown MA, Pile KD, Kennedy LG, Campbell D, Andrew L, March R, Shatford JL, Weeks DE, Calin A & Wordsworth BP; *A Genome-Wide Screen for Susceptibility Loci in Ankylosing Spondylitis*; Arthritis & Rheumatism **41**(4), 588-595 (1998)
65. Laval SH, Timms A, Edwards S, Bradbury S, Brophy S, Milicic A, Rubin L, Siminovitch KA, Weeks DE, Calin A, Wordsworth BP & Brown MA; *Whole-Genome Screening in Ankylosing Spondylitis: Evidence of Non-MHC Genetic-Susceptibility Loci*; Am J. Hum. Genet **68**, 918-926 (2001)
66. Zhang G, Luo J, Bruckel J, Weisman MA, Schumacher HR, Khan MA, Inman RD, Mahowald M, Maksymowych WP, Martin TM, Yu DTY, Stone M, Rosenbaum JT, Newman P, Lee J, McClain JA, West OC, Jin L & Reveille J; *Genetic Studies in Familial Ankylosing Spondylitis Susceptibility*; Arthritis & Rheumatism **50**(7), 2246-2254 (2004)
67. Lee YH, Rho YH, Choi SJ, Ji JD & Song GG; *Ankylosing spondylitis susceptibility loci defined by genome-search meta-analysis*; J Hum Genet. **50**(9), 453-9 (2005)
68. Carter KW, Pluzhnikov A, Timms AE, Miceli-Richard C, Bourgain C, Wordsworth BP, Jean-Pierre H, Cox NJ, Palmer LJ, Breban M, Reveille JD & Brown MA; *Combined analysis of three whole genome linkage scans for Ankylosing Spondylitis*; Rheumatology (Oxford) **46**(5), 763-71 (2007)
69. Karasneh J, Gül A, Ollier WE, Silman AJ & Worthington J; *Whole-genome Screening for Susceptibility Genes in Multicases Families With Behçet's Disease*; Arthritis & Rheumatism **52**(6), 1836-1842 (2005)

70. Hamuryudan V, Yurdakul S, Ozbakir F, Yazici H & Hekim H; *Monozygotic twins concordant for Behçet's syndrome*; *Arthritis Rheum* **34**(8), 1071-2 (1991)
71. Brown MA, Kennedy LG, MacGregor AJ, Darke C, Duncan E, Shatford JL, Taylor A, Calin A & Wordsworth P; *Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment*; *Arthritis Rheum* **40**(10), 1823-8 (1997)
72. Palmer LJ & Cardon LR; *Genetic Epidemiology 4. Shaking the tree: mapping complex disease genes with linkage disequilibrium*; *Lancet* **366**, 1223-34 (2005)
73. Carlson CS, Eberle MA, Kruglyak L & Nickerson DA, Mapping complex disease loci in whole-genome association studies; *Nature* **429**, 446-452 (2004)
74. Payton MA & Sim E; *Genotyping human arylamine N-acetyltransferase type 1 (NAT1): the identification of two novel allelic variants*. *Biochem Pharmacol* **55**(3), 361-6 (1998)
75. Westwood IM, Kawamura A, Fullam E, Russell AJ, Davies SG & Sim E; *Structure and mechanism of arylamine N-acetyltransferases*; *Curr Top Med Chem* **6**(15), 1641-54 (2006)
76. Wakefield L, Long H, Lack N & Sim E; *Ocular defects associated with a null mutation in the mouse arylamine N-acetyltransferase 2 gene*; *Mamm Genome* **18**(4), 270-6 (2007)
77. Husain A, Zhang X, Doll MA, States JC, Barker DF & Hein DW; *Identification of N-acetyltransferase 2 (NAT2) transcription start sites and quantitation of NAT2-specific mRNA in human tissues*; *Drug Metab Dispos* **35**(5), 721-7 (2007)
78. Blum M, Grant DM, McBride W, Heim M & Meyer UA; *Human arylamine N-acetyltransferase genes: isolation, chromosomal localization, and functional expression*; *DNA Cell Biol* **9**(3), 193-203 (1990)
79. Rychlik-Sych M, Skretkiewicz J, Gawrońska-Szklarz B, Górnik W, Sysa-Jedrzejowska A & Skretkiewicz-Szarmach K; *Acetylation genotype and phenotype in patients with systemic lupus erythematosus*; *Pharmacol Rep* **58**(1), 22-9 (2006)
80. Sabbagh A & Darlu P; *Inferring haplotypes at the NAT2 locus: the computational approach*; *BMC Genet* **6**(30), 1-3 (2005)
81. <http://www.louisville.edu/medschool/pharmacology/NAT.html>
82. Nacak M, Erbagci Z & Aynacioglu AS; *Human arylamine N-acetyltransferase 2 polymorphism and susceptibility to allergic contact dermatitis*; *Int J Dermatol* **45**(3), 323-6 (2006)

- 83.** Hein DW, Doll MA, Fretland AJ, Leff MA, Webb SJ, Xiao GH, Devanaboyina US, Nangju NA & Feng Y; *Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms*; *Cancer Epidemiol Biomarkers Prev.* **9**(1), 29-42 (2000)
- 84.** Yuliwulandari R, Sachrowardi Q, Nishida N, Takasu M, Batubara L, Susmiarsih TP, Rochani JT, Wikaningrum R, Miyashita R, Miyagawa T, Sofro AS & Tokunaga K; *Polymorphisms of promoter and coding regions of the arylamine N-acetyltransferase 2 (NAT2) gene in the Indonesian population: proposal for a new nomenclature*; *J Hum Genet.* **53**(3), 201-9 (2008)
- 85.** Najim RA, Farid YY, Samad TA & Shihab SA; *Acetylator phenotype in Iraqi patients with systemic lupus erythematosus*; *East Mediterr Health J* **11**(5-6), 1003-8 (2005)
- 86.** Zscheschang P, Hiepe F, Gromnica-Ihle E, Roots I & Cascorbi I; *Lack of association between arylamine N-acetyltransferase 2 (NAT2) polymorphism and systemic lupus erythematosus*; *Pharmacogenetics* **12**(7), 559-63 (2002)
- 87.** von Schmiedeberg S, Fritsche E, Rönnau AC, Specker C, Golka K, Richter-Hintz D, Schuppe HC, Lehmann P, Ruzicka T, Esser C, Abel J & Gleichmann E; *Polymorphisms of the xenobiotic-metabolizing enzymes CYP1A1 and NAT-2 in systemic sclerosis and lupus erythematosus*; *Adv Exp Med Biol* **455**, 147-52 (1999)
- 88.** Pawlik A, Ostanek L, Brzosko I, Gawroska-Szklarz B, Brzosko M & Dabrowska-Zamojcin E; *Increased genotype frequency of N-acetyltransferase 2 slow acetylation in patients with rheumatoid arthritis*; *Clin Pharmacol Ther* **72**(3), 319-25 (2002)
- 89.** May DG, Black CM, Olsen NJ, Csuka ME, Tanner SB, Bellino L, Porter JA, Wilkinson GR & Branch RA; *Scleroderma is associated with differences in individual routes of drug metabolism: a study with dapson, debrisoquin, and mephenytoin*; *Clin Pharmacol Ther* **48**(3), 286-95 (1990)
- 90.** Taniguchi A, Urano W, Tanaka E, Furihata S, Kamitsuji S, Inoue E, Yamanaka M, Yamanaka H & Kamatani N; *Validation of the associations between single nucleotide polymorphisms or haplotypes and responses to disease-modifying antirheumatic drugs in patients with rheumatoid arthritis: a proposal for prospective pharmacogenomic study in clinical practice*; *Pharmacogenet Genomics* **17**(6), 383-90 (2007)

91. Chen M, Xia B, Chen B, Guo Q, Li J, Ye M & Hu Z; *N-acetyltransferase 2 slow acetylator genotype associated with adverse effects of sulphasalazine in the treatment of inflammatory bowel disease*; *Can J Gastroenterol* **21**(3), 155-8 (2007)
92. Aynacioglu AS, Bozkurt A, Nacak M, Kortunay S, Tunc R, Dincel A & Kayaalp SO; *N-acetyltransferase polymorphism in patients with Behçet's disease*; *Eur J Clin Pharmacol* **57**(9), 659-62 (2001)
93. Najim RA, Sharquie KE & Al-Janabi MH; *Acetylator phenotype in Behçet's disease*; *J Dermatol* **30**(6), 464-71 (2003)
94. Tamer L, Yilmaz A, Yildirim H, Ayaz L, Ates NA, Karakas S, Oz O, Yildirim O & Atik U; *N-acetyltransferase 2 phenotype may be associated with susceptibility to age-related cataract*; *Curr Eye Res* **30**(10), 835-9 (2005)
95. Meyer D, Parkin DP, Seifart HI, Maritz JS, Engelbrecht AH, Werely CJ & van Helden PD; *NAT2 slow acetylator function as a risk indicator for age-related cataract formation*; *Pharmacogenetics* **13**(5), 285-9 (2003)
96. Wakefield L, Long H, Lack N & Sim E; *Ocular defects associated with a null mutation in the mouse arylamine N-acetyltransferase 2 gene*; *Mamm Genome* **18**(4), 270-6 (2007)
97. GeneChip® Human Mapping 500K Array Set\_Data sheet, Affymetrix
98. Affymetrix GeneChip® Genotyping Analysis Software Version 4.1 —User's Guide
99. Abecasis GR, Cherny SS, Cookson WO & Cardon LR; *Merlin - rapid analysis of dense genetic maps using sparse gene flow trees*; *Nat Genet* **30**(1), 97-101 (2002)
100. Kong A & Cox NJ; *Allele-Sharing Models: LOD Scores and Accurate Linkage Tests*, *Am. J. Hum. Genet* **61**(5), 1179-88 (1997)
101. Dawn Teare M & Barrett JH; *Genetic Epidemiology 2 - Genetic linkage studies*; *Lancet* **366**(9490), 1036-44 (2005)
102. Halpern J & Whittemore AS; *Multipoint linkage analysis*; *Hum hered* **49**, 194-196 (1999)
103. Whittemore AS & Halpern J; *A class of tests for linkage using affected pedigree members*; *Biometrics* **50**, 118-127 (1994)
104. Kruglyak L, Daly MJ, Reeve-Daly MP & Lander ES; *Parametric and nonparametric linkage analysis: a unified multipoint approach*; *Am J Hum Genet* **58**(6), 1347-63 (1996)

- 105.** Nyholt DR; *INVITED EDITORIAL - All LODs Are Not Created Equal*; Am J Hum Genet **67**(2), 282-8 (2000)
- 106.** Affymetrix GeneChip Chromosome Copy Number Analysis Tool 4.0 – User Guide
- 107.** Mitchell AA, Chakravarti A & Cutler DJ; *On the probability that a novel variant is a disease-causing mutation*; Genome Res **15**(7), 960-6 (2005)
- 108.** Sobel E & Lange K; *Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics*; American Journal of Human Genetics **58**, 1323-1337 (1996)
- 109.** Almasy L & Blangero J; *Multipoint quantitative trait linkage analysis in general pedigrees*; Am J Hum Genet **62**, 1198-1211 (1998)
- 110.** Gregory PJ, Sperry M & Wilson AF; *Dietary supplements for osteoarthritis*; Am Fam Physician **77**(2), 177-84; (2008)
- 111.** Hungerford MW & Valaik D; *Chondroprotective agents: glucosamine and chondroitin*; Foot Ankle Clin **8**(2), 201-19 (2003)
- 112.** Hardingham T; *Chondroitin sulfate and joint disease*; Osteoarthritis Cartilage **6** Suppl A, 3-5; (1998)
- 113.** Kelly GS; *The role of glucosamine sulfate and chondroitin sulfates in the treatment of degenerative joint disease*; Altern Med Rev **3**(1), 27-39 (1998)
- 114.** Ahn JK & Park YG; *Human leukocyte antigen B27 and B51 double-positive Behçet's uveitis*; Arch Ophthalmol **125**(10), 1375-80 (2007)
- 115.** Hayasaka S, Kurome H & Noda S; *HLA antigens in a Japanese family with Behçet's disease*; Graefes Arch Clin Exp Ophthalmol. **232**(10), 589-90 (1994)
- 116.** Villanueva JL, Gonzalez-Dominguez J, Gonzalez-Fernandez R, Prada JL, Peña J & Solana R; *HLA antigen familial study in complete Behçet's syndrome affecting three sisters*; Ann Rheum Dis **52**(2), 155-7 (1993)
- 117.** Lander E & Kruglyak L; *Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results*; Nat Genet **11**(3), 241-7 (1995)
- 118.** Miceli-Richard C, Zouali H, Said-Nahal R, Lesage S, Merlin F, De Toma C, Blanche H, Sahbatou M, Dougados M, Thomas G, Breban M, Hugot JP & Groupe Français d'Etude Génétique des Spondylarthropathies; *Significant linkage to spondyloarthritis on 9q31-34*; Hum Mol Genet **13**(15), 1641-8 (2004)

- 119.** Sfar S, Abid A, Mahfouh W, Ouragini H, Ouechtati F, Abdelhak S & Chouchane L; *Genetic analysis of hereditary multiple exostoses in Tunisian families: a novel frame-shift mutation in the EXT1 gene*; Mol Biol Rep (2008) (printing);
- 120.** Lü Y, Zhu HG, Ye WM, Zhang MB, He D & Chen WT; *A new mutation of PTCH gene in a Chinese family with nevoid basal cell carcinoma syndrome*; Chin Med J (Engl) **121**(2), 118-21 (2008)
- 121.** Kirito K, Sakoe K, Shinoda D, Takiyama Y, Kaushansky K & Komatsu N; *A novel RUNX1 mutation in familial platelet disorder with propensity to develop myeloid malignancies*; Haematologica **93**(1), 155-6 (2008)
- 122.** Chiou HL, Wu MF, Chien WP, Cheng YW, Wong RH, Chen CY, Lin TS & Lee H; *NAT2 fast acetylator genotype is associated with an increased risk of lung cancer among never-smoking women in Taiwan*; Cancer Lett **223**(1), 93-101 (2005)
- 123.** Fretland AJ, Leff MA, Doll MA, Hein DW; *Functional characterization of human N-acetyltransferase 2 (NAT2) single nucleotide polymorphisms*; Pharmacogenetics **11**(3), 207-215 (2001)
- 124.** Torkaman-Boutorabi A, Hoormand M, Naghdi N, Bakhshayesh M & Milanian I; *Genotype and allele frequencies of N-acetyltransferase 2 and glutathione S-transferase in the Iranian population*; Clin Exp Pharmacol Physiol **34**(11), 1207-11 (2007)
- 125.** Thorsby E & Lie BA; *HLA associated genetic predisposition to autoimmune diseases: Genes involved and possible mechanisms*; Transpl Immunol. Review **14**(3-4), 175-82 (2005)

## **Anexo A**

---

### **Declaração da comissão de ética do IPR**



Exmo. Senhor  
Dr. José Vaz Patto  
Investigador principal  
I.P.R

Lisboa, 4 de Dezembro de 2007

Exmo. Senhor,

A Comissão de Ética para a Saúde do Instituto Português de Reumatologia, vem informar V. Exa. que aprovou na sua reunião do dia 4 de Dezembro de 2007, a realização do estudo Foi apreciado e aprovada a realização do estudo "Estudo de factores genéticos, moleculares e celulares numa família portuguesa com Doença de Behçet e Espondilite Anquilosante", no Instituto de Reumatologia.

Com os meus cumprimentos.

A Comissão de Ética

*Paula Maria Rodrigues Pais*

PP/MP

## **Anexo B**

---

### **Informação adicional sobre a análise no programa Merlin**

O programa Merlin necessita de três ficheiros de *input*: o ficheiro do pedigree (ficheiro.ped) onde se identifica a família, o indivíduo, a relação de parentesco entre todos os indivíduos, o sexo de cada indivíduo e ainda o seu fenótipo; o ficheiro de mapeamento que contém o nome e localização cromossómica de cada marcador (ficheiro.map) e o ficheiro de dados onde se encontra o genótipo de cada marcador usado (ficheiro.dat). O ficheiro de dados pode ser apenas um com informação de todos os cromossomas ou alternativamente, vários ficheiros, cada um com informação de apenas um cromossoma.

Os comandos usados para correr a análise de *linkage* não paramétrica foram:

```
>merlin -d ficheiro.dat -p ficheiro.ped -m ficheiro.map --megabytes 1200 --grid 0.1 --pairs  
--npl -exp -pdf -zscore --prefix nome_do_ficheiro_de_ouput
```

Onde a opção **--grid** determina o espaçamento usado entre os marcadores (em cM), a opção **--pairs** e **--npl** o cálculo da estatística  $S_{\text{pairs}}$  e  $S_{\text{all}}$  de Whittemore e Halpern, respectivamente, a opção **--exp** o modelo exponencial de Kong e Cox (a análise com o modelo linear é sempre realizada por defeito) e a opção **--zscore** o cálculo da estatística NPL de Kruglayak.

Adicionalmente usou-se as opções **--start** e **--stop** (com indicação à frente da posição em cM) para limitar a análise dum determinado cromossoma a uma região específica (usando o ficheiro de dados do respectivo cromossoma).

## **Anexo C**

---

### **Árvores genealógicas com modelos fenotípicos**

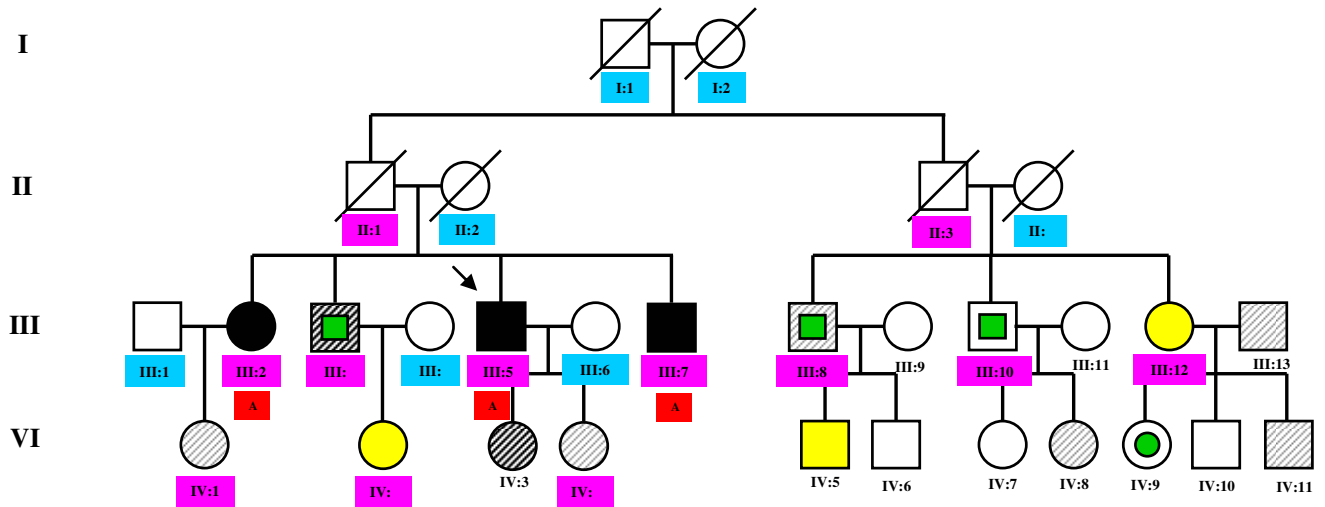
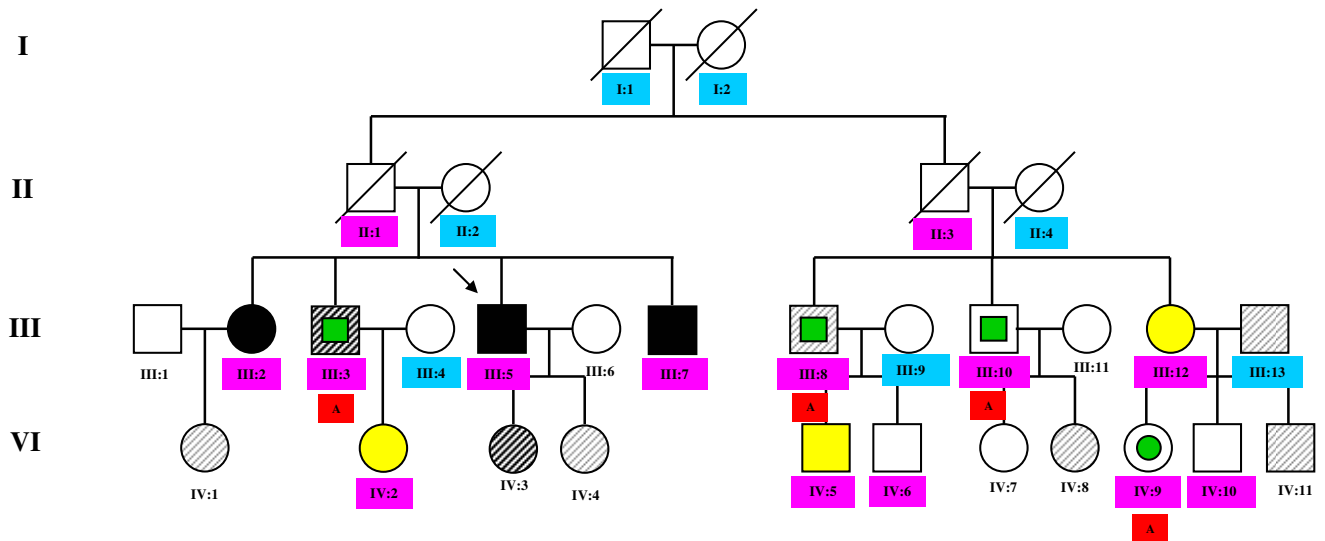


Figura C1. Árvore genealógica com representação do modelo fenotípico “DB”



**Legenda:**

- |                           |                                      |  |
|---------------------------|--------------------------------------|--|
| □ Não afectados           | ■ Espondilite Anquilosante           | ■ Indivíduos fundadores incluídos na análise     |
| ● Doença de Behçet        | ■ Espondiloartropatia indiferenciada | ■ Indivíduos não fundadores incluídos na análise |
| ▨ Um sintoma de Behçet    | ➤ Doente índice                      | ■ Indivíduos afectados                           |
| ▩ Dois sintomas de Behçet |                                      |  |

Figura C2. Árvore genealógica com representação do modelo fenotípico “EA”

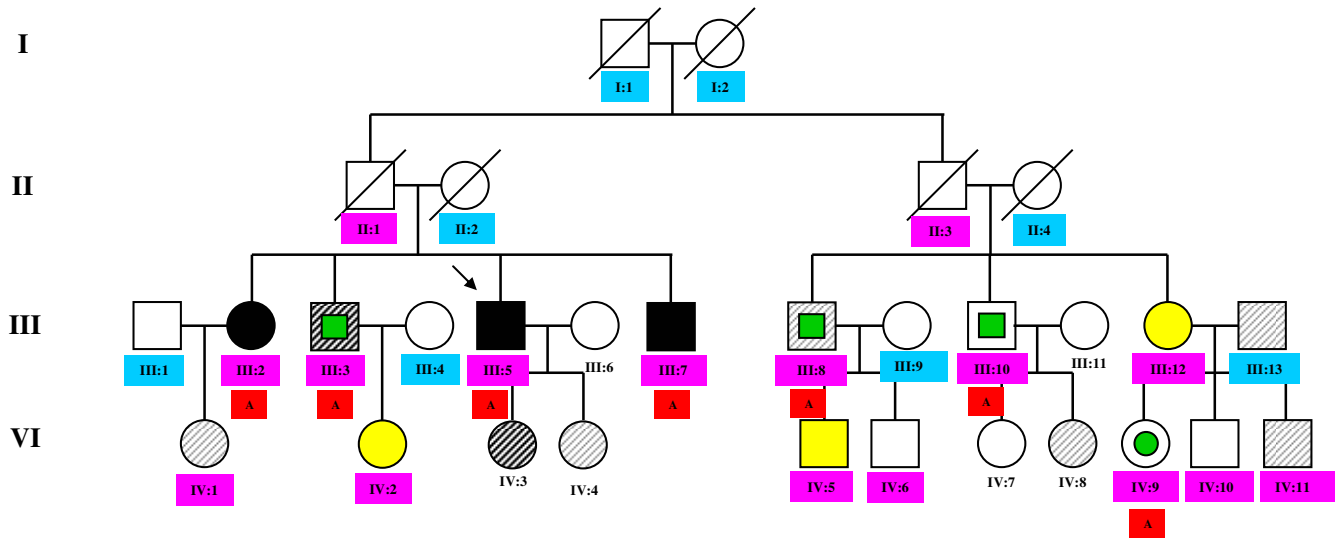
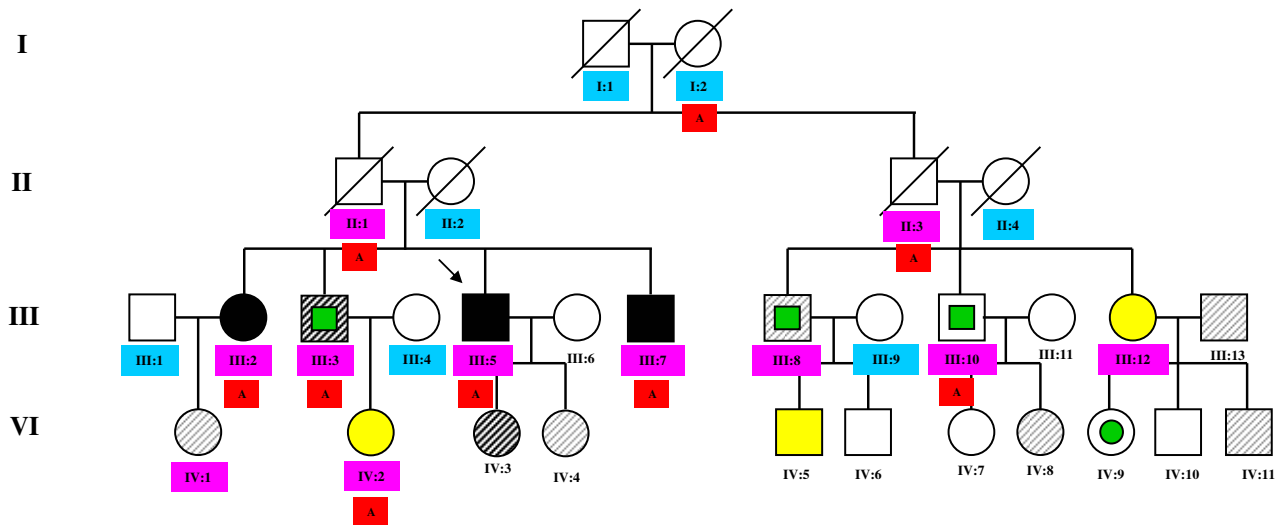


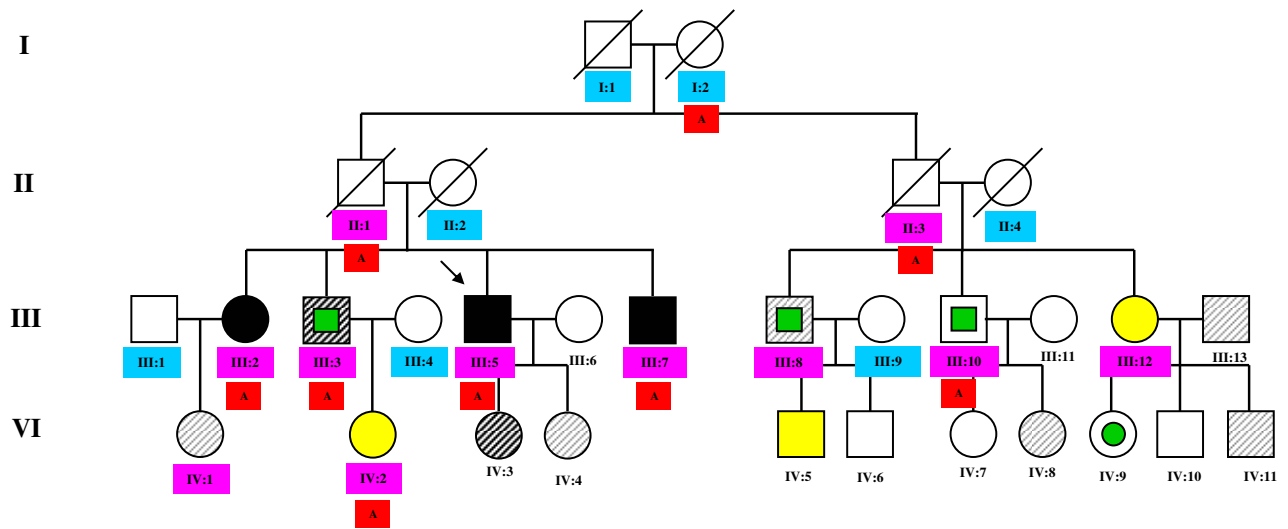
Figura C3. Árvore genealógica com representação do modelo fenotípico “DB e EA”



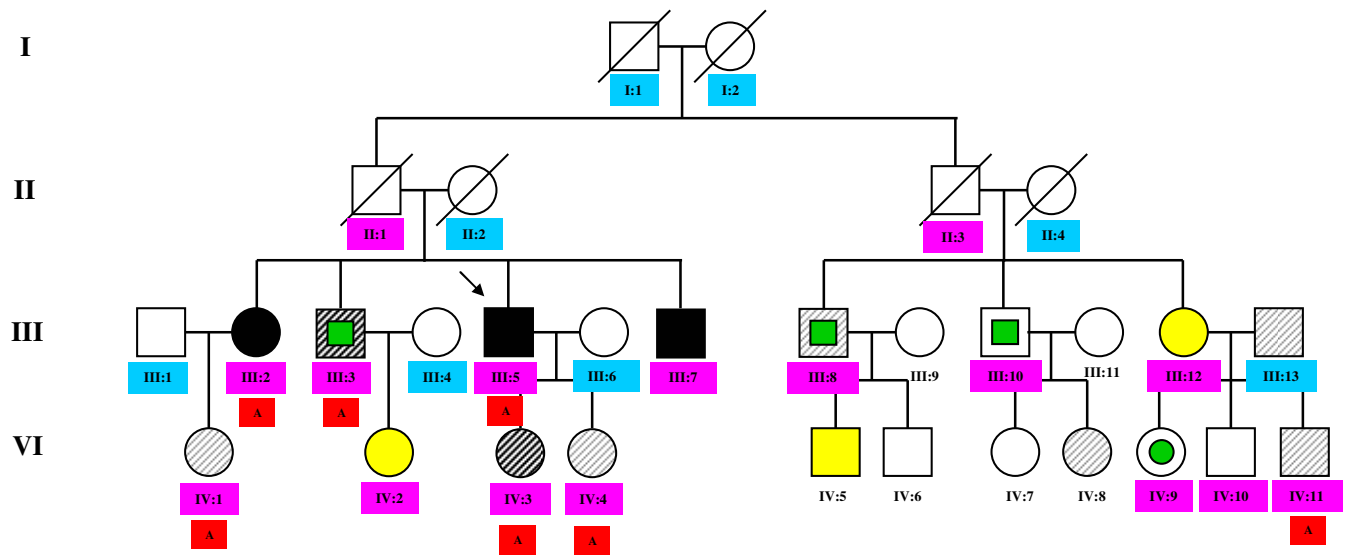
Legenda:

- |   |                         |   |                                    |   |  |
|---|-------------------------|---|------------------------------------|---|--|
| □ | Não afectados           | ⊞ | Espondilite Anquilosante           | ■ | Indivíduos fundadores incluídos na análise     |
| ● | Doença de Behçet        | ■ | Espondiloartropatia indiferenciada | ■ | Indivíduos não fundadores incluídos na análise |
| ⊞ | Um sintoma de Behçet    | ⊞ | Doente índice                      | ■ | Indivíduos afectados                           |
| ⊞ | Dois sintomas de Behçet |   |                                    |   |  |

Figura C4. Árvore genealógica com representação do modelo fenotípico “Lesões Oculares”



**Figura C5.** Árvore genealógica com representação do modelo fenotípico “Aftas Oraes”



**Legenda:**

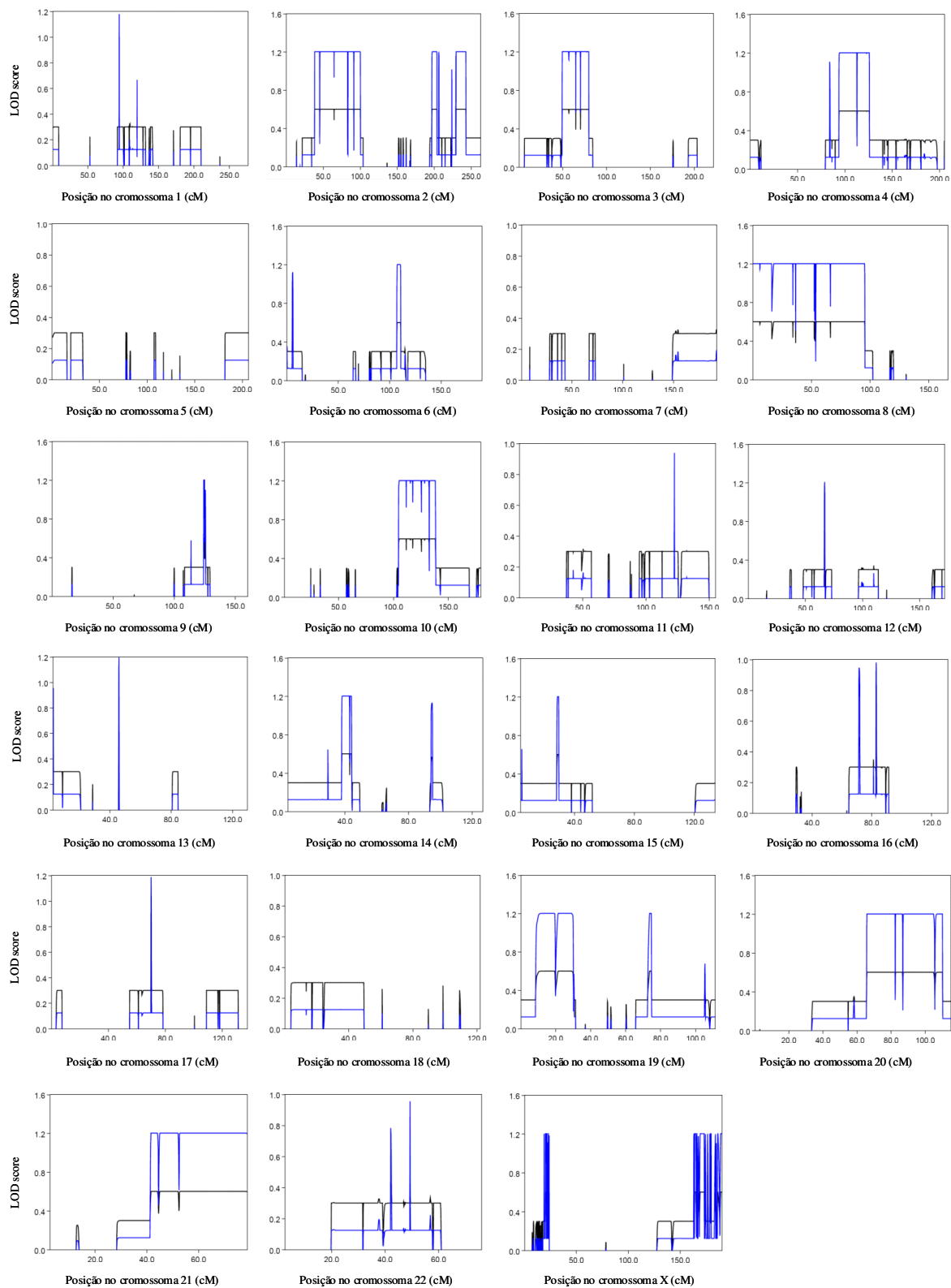
- |                           |                                      |  |
|---------------------------|--------------------------------------|--|
| □ Não afectados           | ◻ Espondilite Anquilosante           | ■ Indivíduos fundadores incluídos na análise     |
| ● Doença de Behçet        | ◻ Espondiloartropatia indiferenciada | ■ Indivíduos não fundadores incluídos na análise |
| ◻ Um sintoma de Behçet    | ⚡ Doente índice                      | ■ Indivíduos afectados                           |
| ◻ Dois sintomas de Behçet |                                      |  |

**Figura C6.** Árvore genealógica com representação do modelo fenotípico “Lesões Cutâneas”

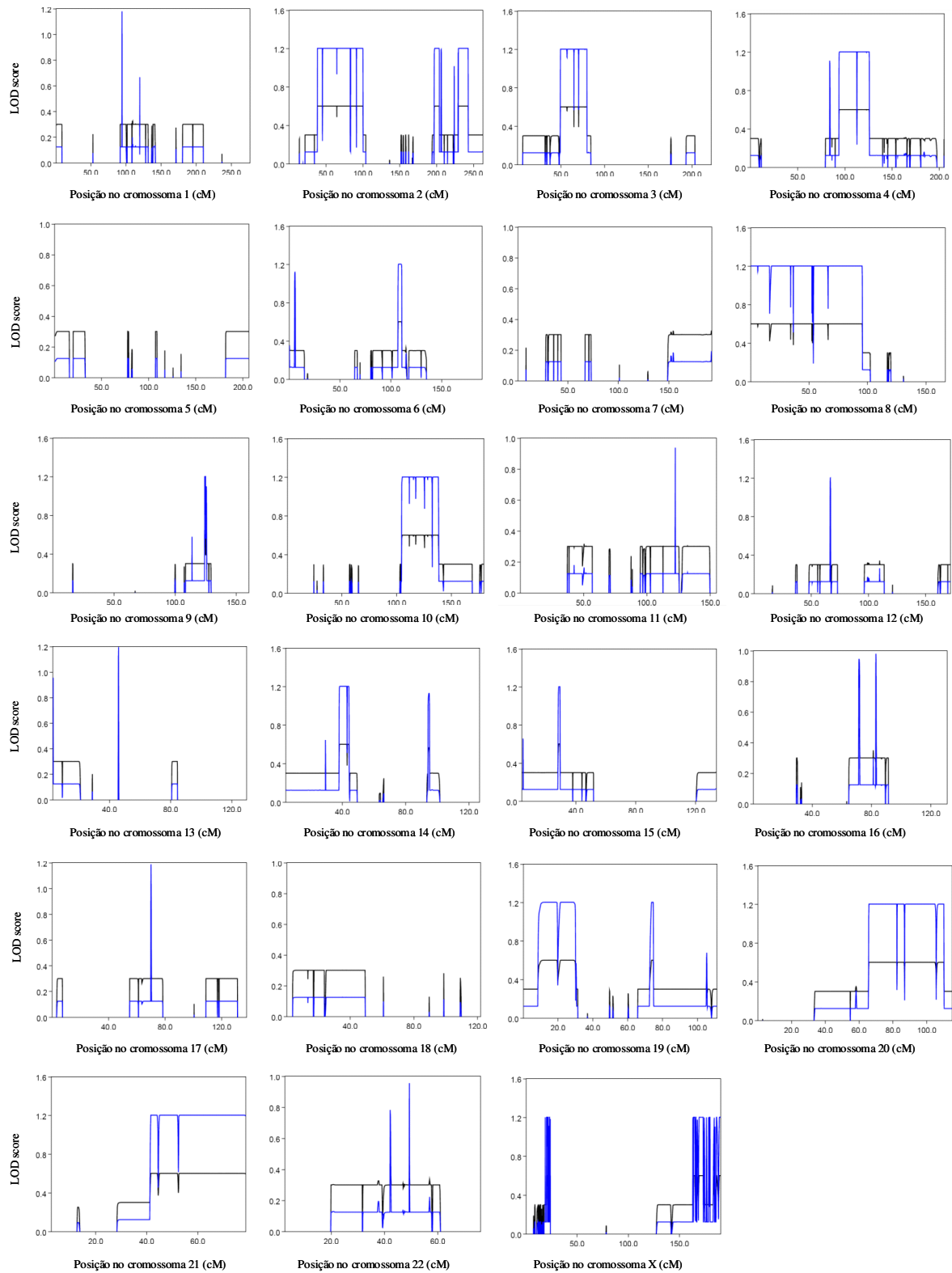
## **Anexo D**

---

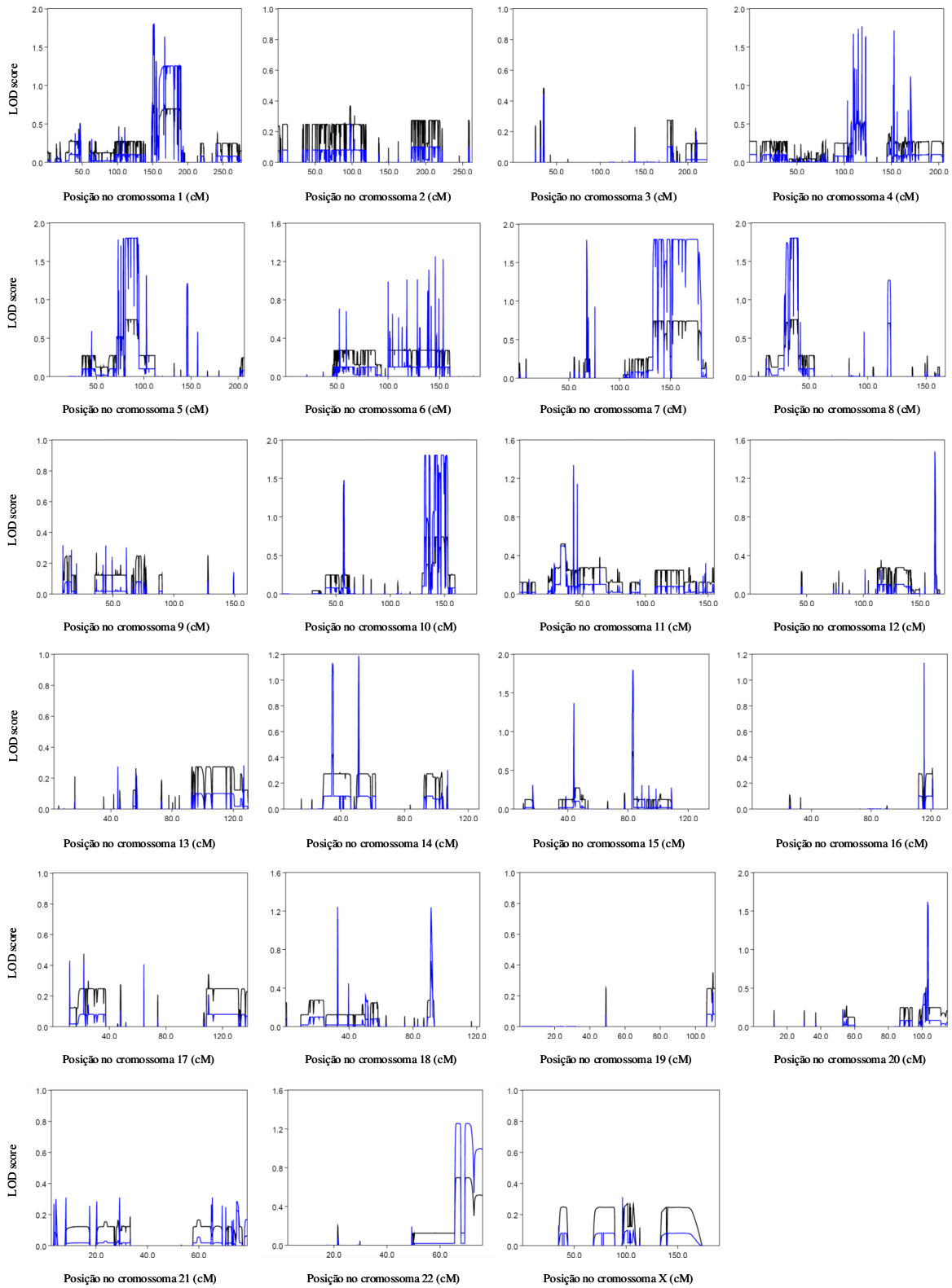
### **Gráficos de *linkage***



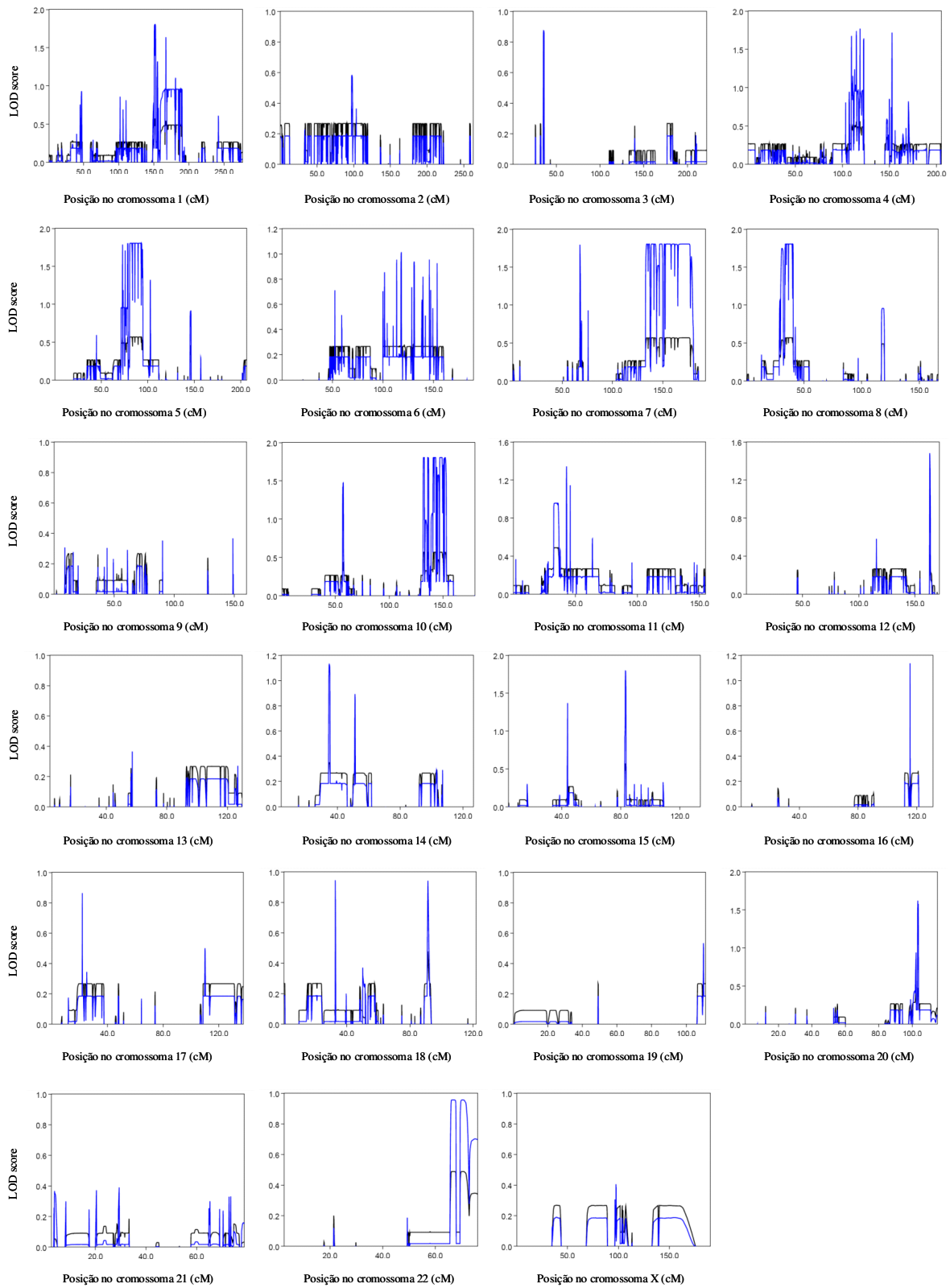
**Figura D1.** Análise NPL *multi-point* no genoma total com a estatística *Sall* de Whittmore e Halpern, para o modelo fenotípico “DB”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



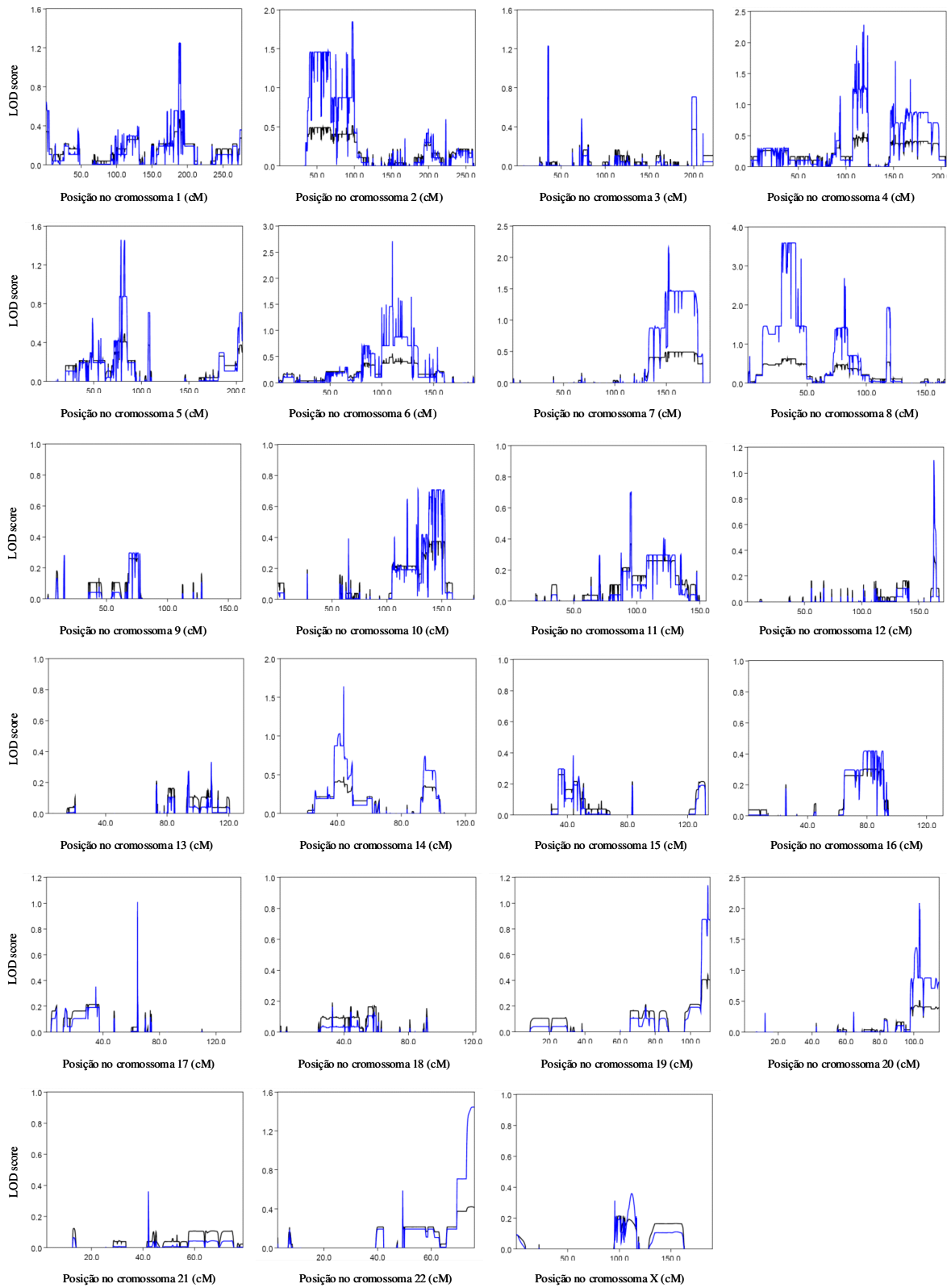
**Figura D2.** Análise NPL *multi-point* no genoma total com a estatística *Spairs* de Whittmore e Halpern, para o modelo fenotípico “DB”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



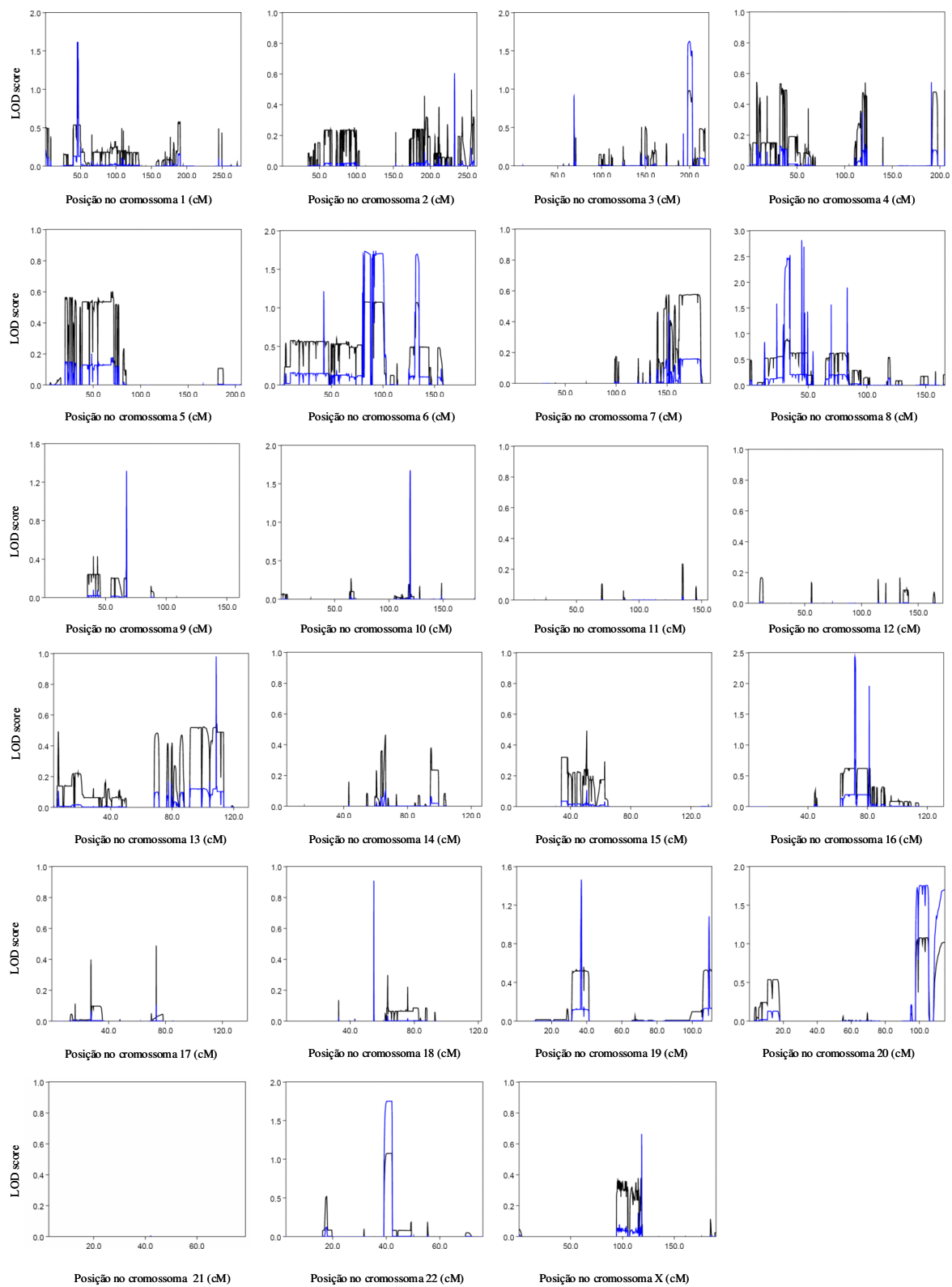
**Figura D3.** Análise NPL *multi-point* no genoma total com a estatística *Sall* de Whittmore e Halpern, para o modelo fenotípico “EA”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



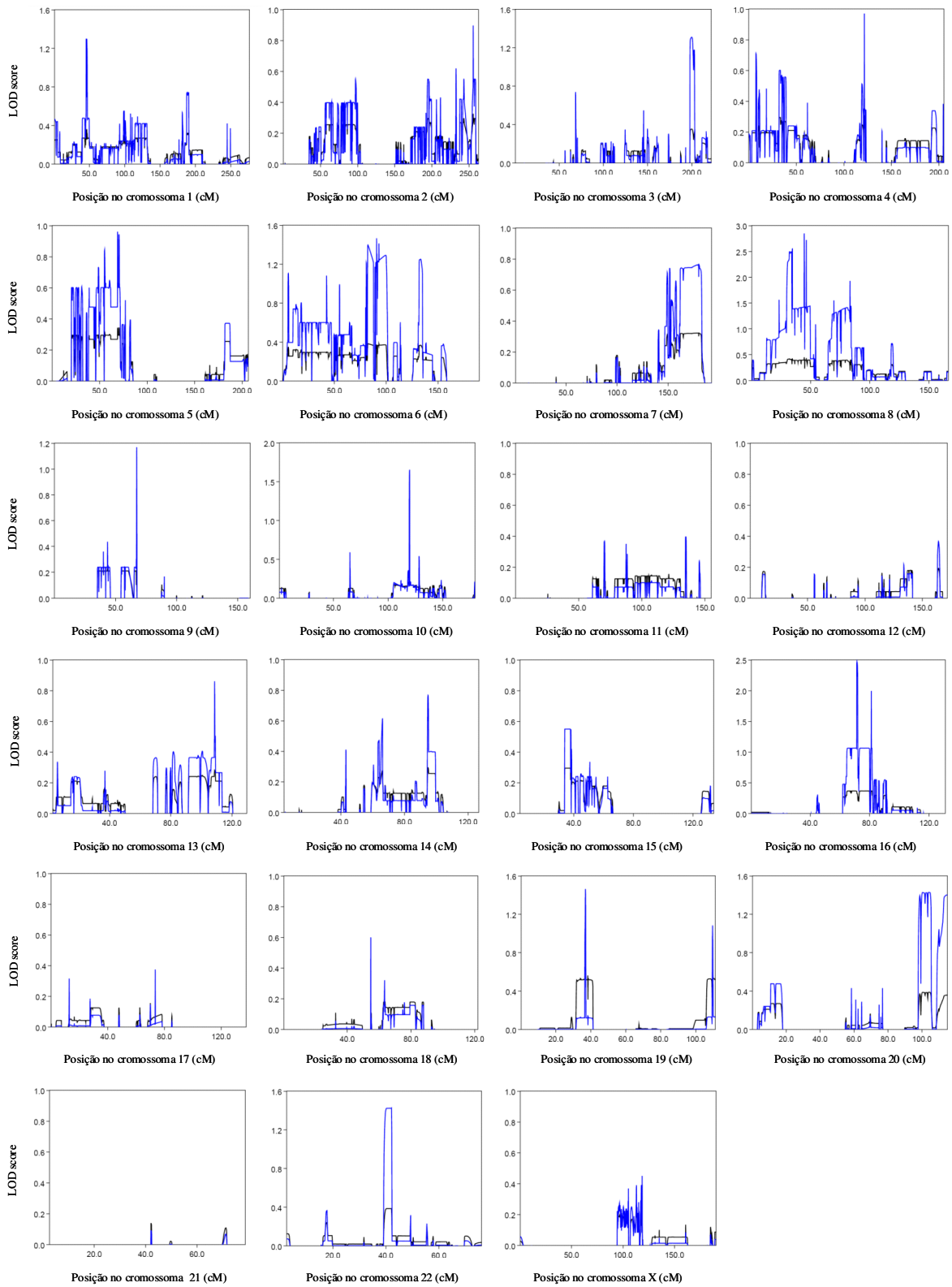
**Figura D4.** Análise NPL *multi-point* no genoma total com a estatística *Spairs* de Whittmore e Halpern, para o modelo fenotípico “EA”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



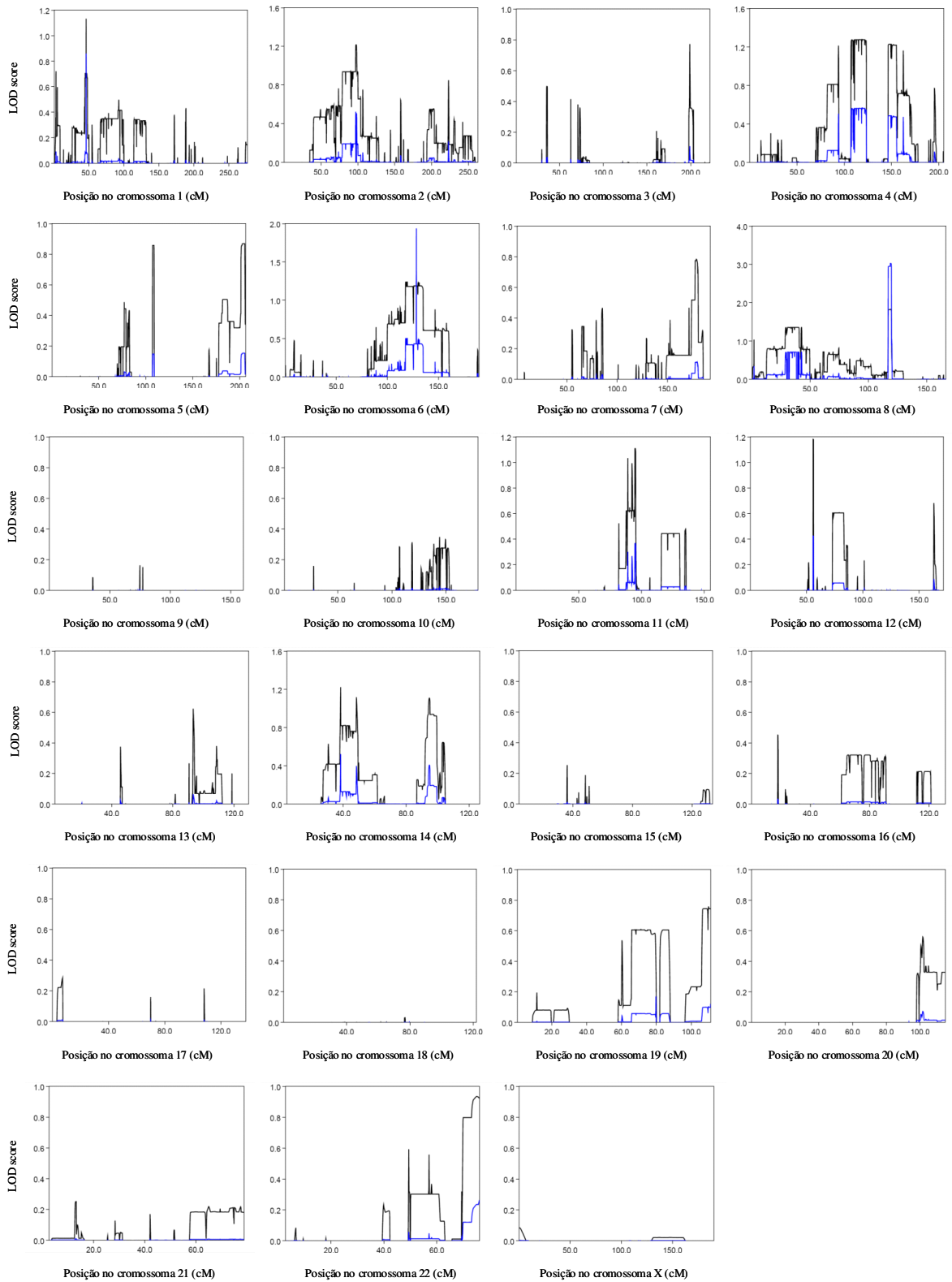
**Figura D5.** Análise NPL *multi-point* no genoma total com a estatística *Spairs* de Whittmore e Halpern, para o modelo fenotípico “DB e EA”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



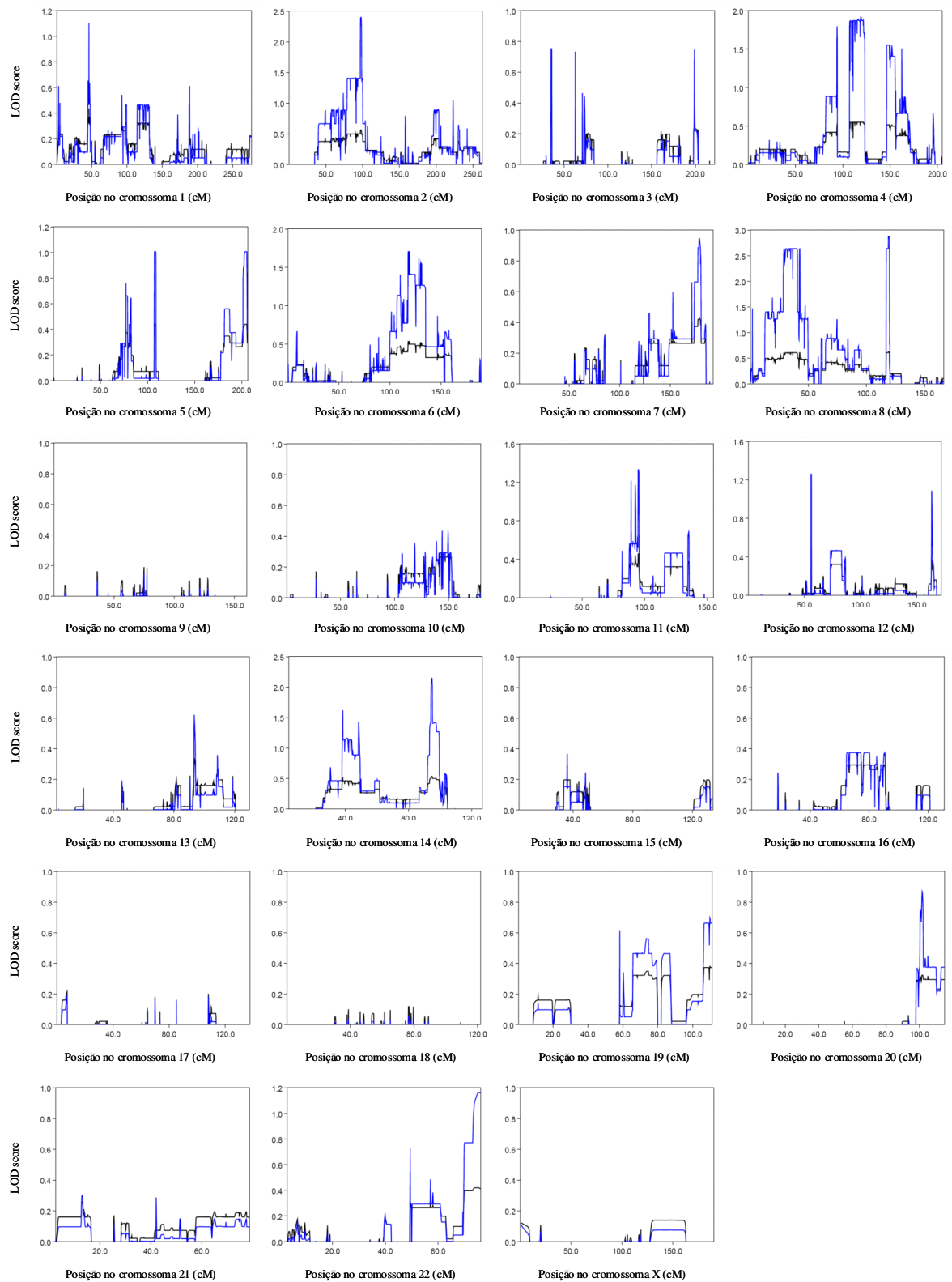
**Figura D6.** Análise NPL *multi-point* no genoma total com a estatística *Sall* de Whittmore e Halpern, para o modelo fenotípico “Lesões oculares”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



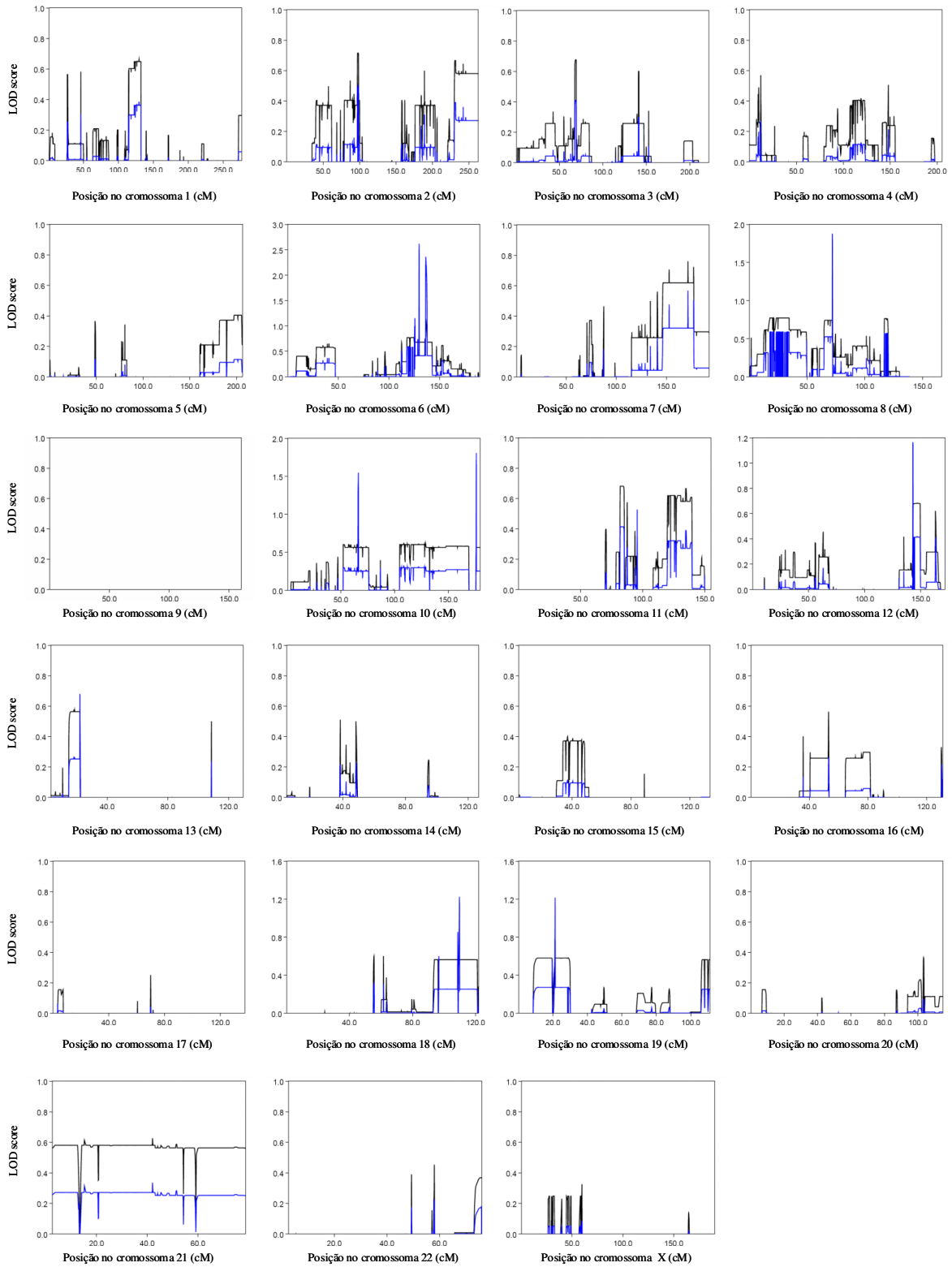
**Figura D7.** Análise NPL *multi-point* no genoma total com a estatística *Spairs* de Whittmore e Halpern, para o modelo fenotípico “Lesões oculares”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



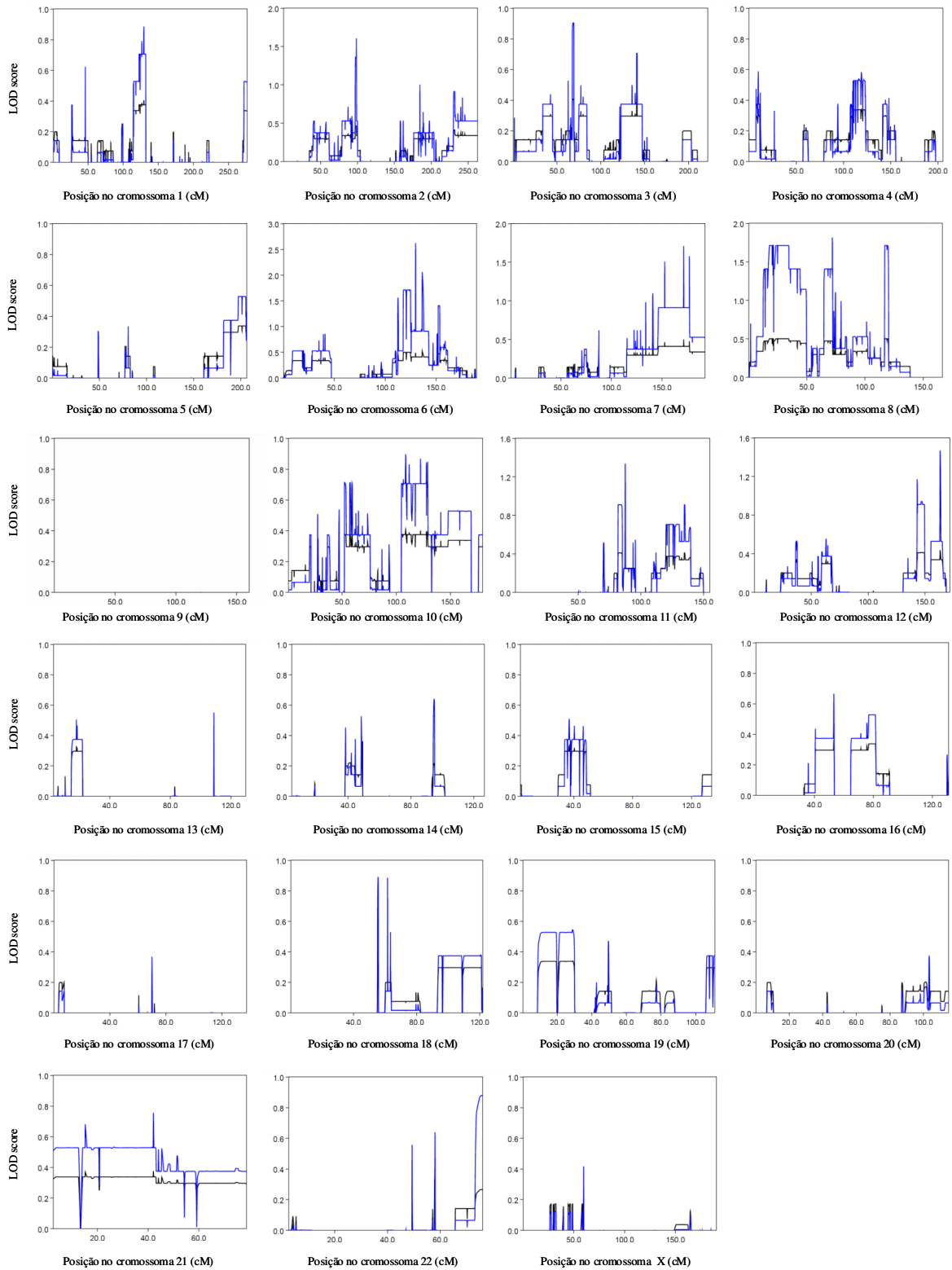
**Figura D8.** Análise NPL *multi-point* no genoma total com a estatística *Sall* de Whittmore e Halpern, para o modelo fenotípico “Aftas orais”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



**Figura D9.** Análise NPL *multi-point* no genoma total com a estatística *Spairs* de Whittmore e Halpern, para o modelo fenotípico “Aftas orais”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.



**Figura D10.** Análise NPL *multi-point* no genoma total com a estatística *Sall* de Whittmore e Halpern, para o modelo fenotípico “Manifestações cutâneas”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.

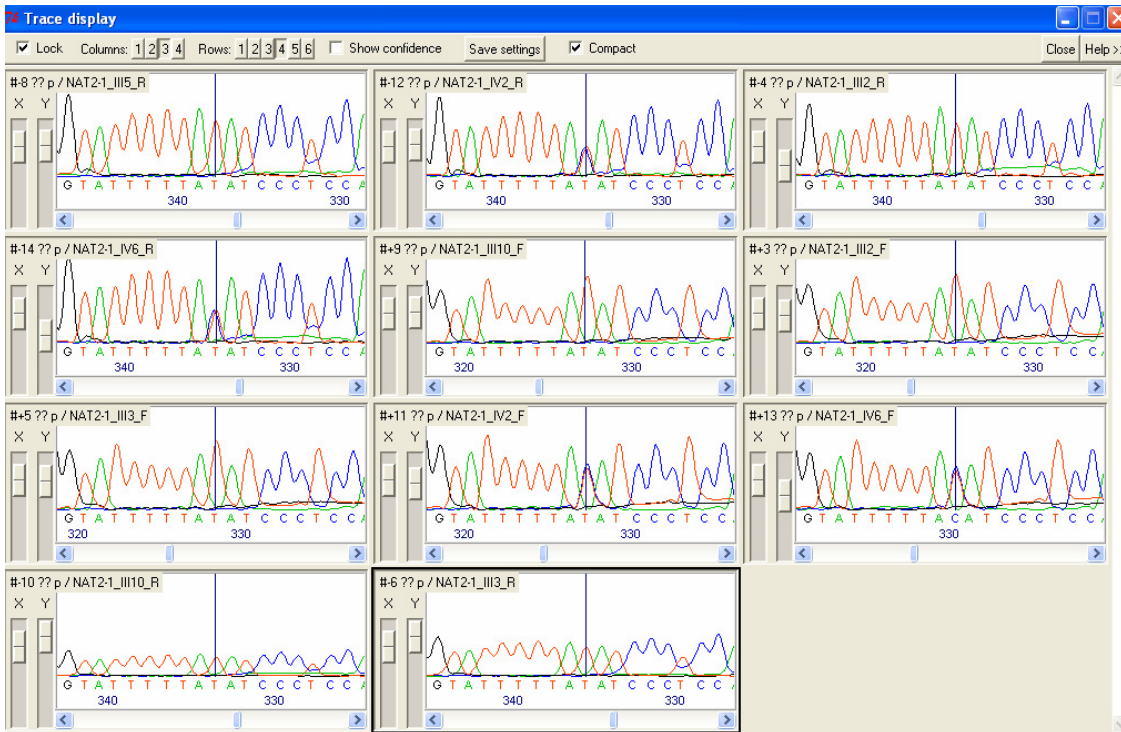


**Figura D11.** Análise NPL *multi-point* no genoma total com a estatística *Spairs* de Whittmore e Halpern, para o modelo fenotípico “Manifestações cutâneas”. A linha preta representa o modelo linear e a linha azul o modelo exponencial de Kong e Cox.

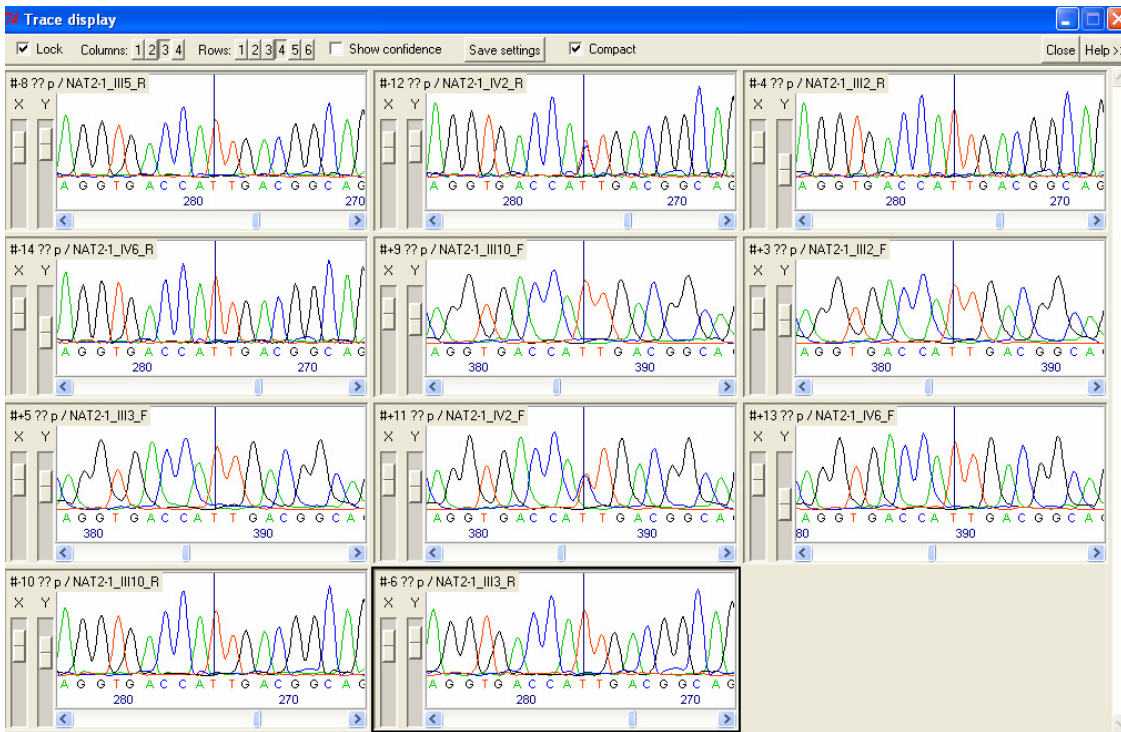
## **Anexo E**

---

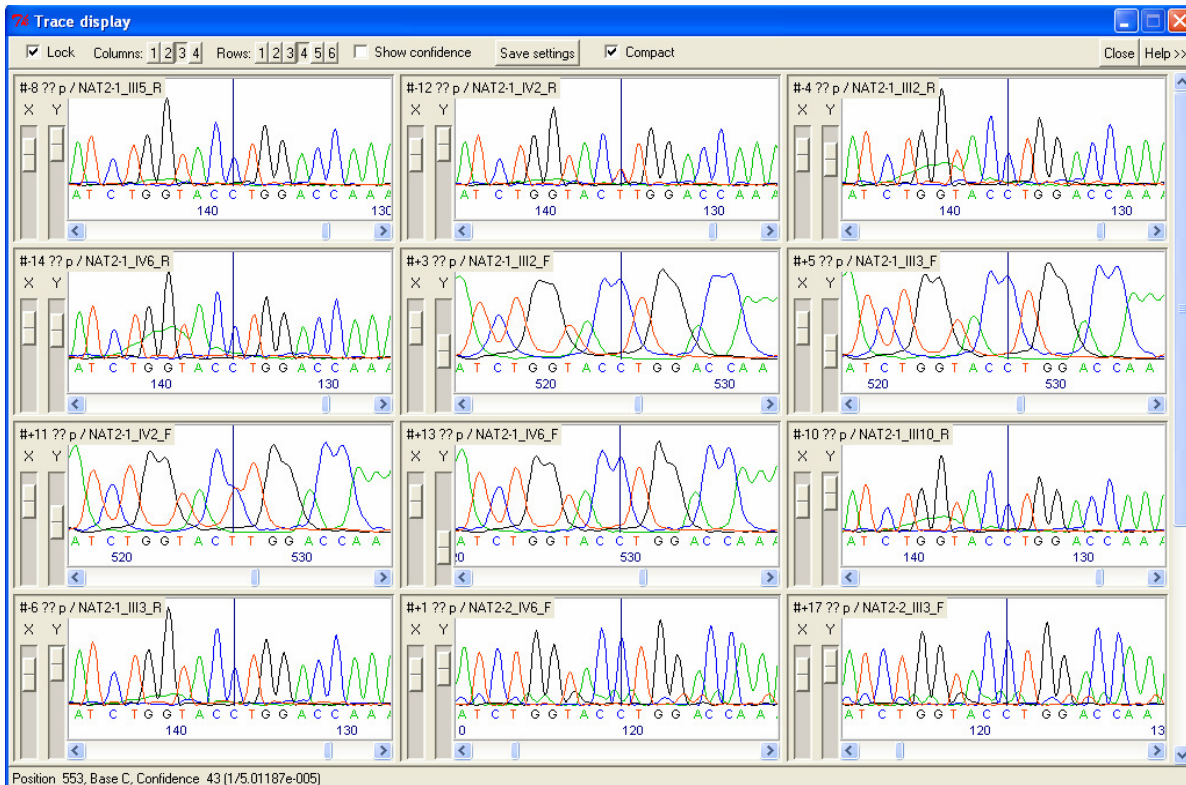
### **Imagens de sequenciação**



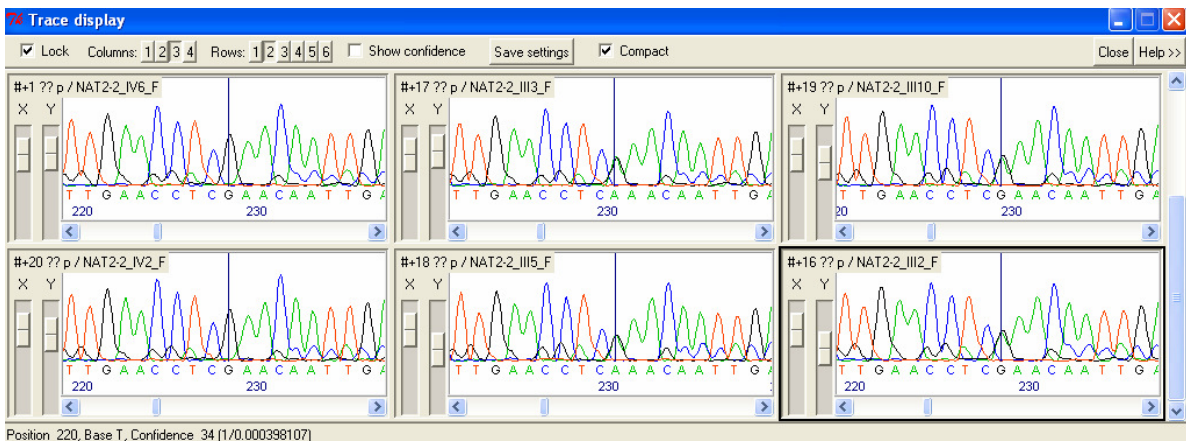
**Figura E1.** Gráficos com sequências dos seis indivíduos primeiramente analisados (III:2, III:3, III:5, III:10, IV:2 e IV:6) ilustrando o SNP rs1041983 (C282T). Estão representadas as cadeias *forward* e *reverse* para todos os indivíduos com a exceção da cadeia *forward* para o indivíduo III:5.



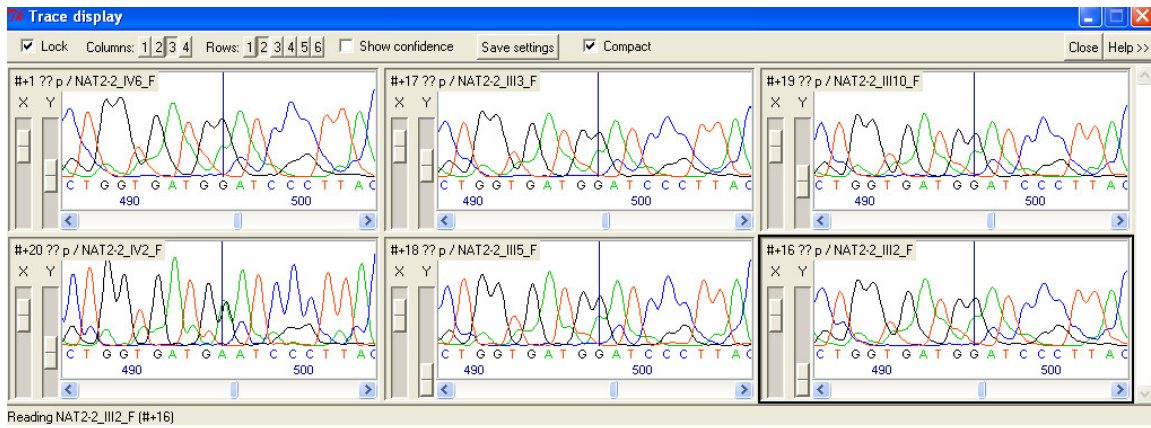
**Figura E2.** Gráficos com sequências dos seis indivíduos primeiramente analisados (III:2, III:3, III:5, III:10, IV:2 e IV:6) ilustrando o SNP rs1801280 (T341C). Estão representadas as cadeias *forward* e *reverse* para todos os indivíduos com a exceção da cadeia *forward* para o indivíduo III:5.



**Figura E3.** Gráficos com sequências dos seis indivíduos primeiramente analisados (III:2, III:3, III:5, III:10, IV:2 e IV:6) ilustrando o SNP rs179929 (C481T). Estão representadas as cadeias *forward* e *reverse* para todos os indivíduos.



**Figura E4.** Gráficos com sequências dos seis indivíduos primeiramente analisados (III:2, III:3, III:5, III:10, IV:2 e IV:6) ilustrando o SNP rs179930 (G590A). Está representada apenas a cadeia *forward*.



**Figura E5.** Gráficos com seqüências dos seis indivíduos primeiramente analisados (III:2, III:3, III:5, III:10, IV:2 e IV:6) ilustrando o SNP rs179931 (G857A). Está representada apenas a cadeia *forward*.