

INSTITUTO POLITÉCNICO DE LISBOA
ESCOLA SUPERIOR DE TECNOLOGIA DA SAÚDE DE LISBOA

DETERMINANTES GENÉTICOS DA INFEÇÃO NA FEBRE Q CRÓNICA

Liliana Alexandra Marques De Castro

Orientadora: Doutora Susana David (Instituto Nacional de Saúde Doutor Ricardo Jorge- INSA)

Orientadora: Professora Doutora Anita Gomes (Escola Superior de Tecnologias de Saúde de Lisboa- ESTeSL)

Mestrado em Tecnologias Moleculares em Saúde

Lisboa, 2022

INSTITUTO POLITÉCNICO DE LISBOA
ESCOLA SUPERIOR DE TECNOLOGIA DA SAÚDE DE LISBOA

DETERMINANTES GENÉTICOS DA INFEÇÃO NA FEBRE Q CRÓNICA

Liliana Alexandra Marques De Castro

Orientadora: Doutora Susana David (Instituto Nacional de Saúde Doutor Ricardo Jorge- INSA)

Orientadora: Professora Doutora Anita Gomes (Escola Superior de Tecnologias de Saúde de Lisboa- ESTeSL)

Juri:

Presidente: Professora Doutora Maria Helena Soares (Escola Superior de Tecnologias de Saúde de Lisboa- ESTeSL)

Arguente: Professora Doutora Elsa Leclerc Duarte (Escola de Ciências e Tecnologia – Universidade de Évora)

Mestrado em Tecnologias Moleculares em Saúde

Lisboa, 2022

 
Mestrado em Tecnologias Moleculares em Saúde
DETERMINANTES GENÉTICOS DA INFEÇÃO NA FEBRE Q CRÓNICA
Liliana Alexandra Marques de Castro
2022

“All that is necessary for faith is the belief that by doing our best we shall come nearer to success.”

Rosalind Franklin

Agradecimentos

Ao terminar este trabalho, gostaria de manifestar o meu sincero reconhecimento pelo apoio, disponibilidade e colaborações concedidas pelas diversas pessoas que direta ou indiretamente tornaram possível a sua execução:

À Escola Superior de Tecnologia da Saúde de Lisboa (ESTeSL), na pessoa da sua Presidente, a Prof.^a Coordenadora Beatriz Fernandes pelas condições de aprendizagem que me tem proporcionado desde o início da minha admissão nesta instituição.

À direção do Mestrado em Tecnologias Moleculares em Saúde, Professora Doutora Helena Soares e a todos os docentes por todos os ensinamentos e pela forma como sempre nos cativaram para o mundo científico, o meu muito obrigada.

À Professora Anita Gomes, manifesto o meu agradecimento por ter aceite ser minha co-orientadora, pelo incentivo, colaboração, disponibilidade e preocupação demonstrada.

Ao Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), na pessoa Presidente do Conselho Diretivo Professor Doutor Fernando Almeida e à Doutora Glória Isidro, Coordenadora do Departamento de Genética do INSA pela oportunidade de realização deste trabalho.

À Doutora Susana David, expresse o meu agradecimento por ter aceite ser minha orientadora, pela preocupação e compreensão constante, pelas palavras de motivação, por todo o conhecimento transmitido e ajuda durante a execução deste trabalho. Estou-lhe ainda grata pela disponibilidade demonstrada, pela bibliografia disponibilizada, pela correção minuciosa desta dissertação e pelos reparos e sugestões que muitos contribuíram para o seu aperfeiçoamento.

Agradeço à Doutora Ana Santos, do Centro de Estudos de Vetores e Doenças Infeciosas (CEVDI), detentora do projeto “Febre Q: do diagnóstico à pesquisa eco epidemiológica de *Coxiella burnetii* no contexto da infeção humana”, financiado pela Fundação de Ciência e Tecnologia (financiamento: PTDC/SAU-SAP/115266/2009), no qual se insere este trabalho. Agradeço, igualmente, pelo fornecimento das amostras clínicas utilizadas neste trabalho.

À Unidade de Tecnologia e Inovação (UTI) do Instituto Nacional de Saúde Doutor Ricardo Jorge, na pessoa do Doutor Luís Vieira, bem como, aos elementos da sua equipa, Joana Mendonça, Miguel Machado e José Ferrão, pela realização da sequenciação necessária a este trabalho.

Ao José Carlos Robles-Fuentes e ao Ulisses Gaspar, pelo apoio técnico no âmbito dos seus trabalhos de *Erasmus* e mestrado, respetivamente, pelo seu contributo neste projeto, permitindo-me desenvolver o meu trabalho e dissertação à *posteriori*.

À Unidade de Investigação e Desenvolvimento, Laboratório de oncobiologia e vias de sinalização, nomeadamente à Doutora Patrícia Barros por nos ter facilitado a utilização do termociclador.

Aos meus amigos, por sempre acreditarem nas minhas capacidades e pelas palavras de motivação que não faltaram ao longo deste caminho.

À minha família, em especial à minha mãe e à minha irmã, pelo apoio incondicional, pela compreensão quando não pude estar tão presente, e por serem sempre o meu pilar.

Abstract

Q fever is a zoonosis of worldwide distribution caused by the bacterium *Coxiella burnetii*. The disease, resulting in high morbidity and mortality, presents itself in the self-limiting acute form or in the chronic form, with severe implications. A defective immune response may contribute to the development of chronic Q fever.

The present study aimed to identify, in the Portuguese population, genetic variants in the principal genes of the immune response to infection that could be associated with the development of chronic Q fever.

A retrospective case-control candidate gene association study (CGAS) was used for genotype-phenotype association analysis in an exploratory study and in a replication study. In the first, we used NGS to sequence coding and non-coding regions in the *IL12RB1*, *IFNGR1*, *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* and *TIMP1* genes. In the second, the genes *P2X7R* and *IFNG* were genotyped using Sanger sequencing. The study sample included 17 chronic disease cases, 43 acute cases, and 34 convenient samples (healthy).

SNPs with statistical significance resisting the Bonferroni correction were identified only in the *IFNGR2* gene. Linkage disequilibrium analysis revealed other variants in association.

In conclusion, our strategy was found to be suitable for the identification of genetic determinants of infection in chronic Q fever. However, functional studies are necessary in order to elucidate the role of the associated variants with chronic Q fever.

Key words: Chronic Q fever; CGAS; SNPs; NGS; Sanger sequencing

Resumo

A febre Q é uma zoonose de distribuição mundial causada pela bactéria *Coxiella burnetii*. A doença, que resulta numa elevada morbilidade e mortalidade, pode manifestar-se na forma aguda autolimitada ou na forma crónica de evolução grave. Considera-se que uma resposta imunitária deficitária poderá contribuir para o desenvolvimento da febre Q crónica.

O presente estudo procurou identificar, na população portuguesa, variantes genéticas dos principais genes envolvidos na resposta imune à infeção que pudessem estar associadas a um desenvolvimento de febre Q crónica.

Recorremos à análise retrospectiva caso-controlo usando métodos de associação genótipo-fenótipo com base em genes candidatos (*Candidate gene association studies* - CGAS) através de um estudo exploratório e de um estudo réplica. No primeiro, utilizamos o NGS para sequenciar regiões codificantes e não codificantes nos genes *IL12RB1*, *IFNGR1*, *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* e *TIMP1*. No segundo, foram genotipados, por sequenciação Sanger, variantes nos genes *P2X7R* e *IFNG*. As amostras incluíam 17 casos com doença crónica, 43 com doença aguda e 34 amostras de conveniência (saudáveis).

SNPs com significância estatística resistindo à correção de Bonferroni foram identificados apenas no gene *IFNGR2*. A análise dos padrões de desequilíbrio de ligação revelou ainda outras variantes em associação.

Concluindo, a estratégia adotada revelou-se adequada para a identificação de determinantes genéticos da infeção na febre Q crónica. Contudo, será imprescindível avançar com estudos funcionais de modo a esclarecer o papel das variantes associadas com a febre Q crónica.

Palavras Chave: Febre Q crónica; CGAS; SNPs; NGS; Sequenciação Sanger

Lista de abreviaturas e acrónimos

- A – Adenina
C – Citosina
CDC – *Centre for Disease Control and Prevention*
CEU – Habitantes do Estado do Utah com ascendência do norte e oeste da Europa
CGA – *Candidate Gene Association Studies* - Estudos de associação de gene candidato
CQ – Controlo de qualidade
DGS – Direção Geral de Saúde
DNA – *Deoxyribonucleic Acid* – Ácido Desoxirribonucleico
DNA *Ladder* – marcador de peso molecular padrão
dNTPs – Desoxinucleosídeos trifosfato
EDTA – *Ethylenediamine tetraacetic acid* – Ácido etileno diamino tetra acético
g – Gramas
G – Guanina
H₂O – Água (bidestilada)
HWE – *Hardy-Weinberg Equilibrium* – Equilíbrio de Hardy-Weinberg
IBS – População ibérica em Espanha
IC – Intervalo de confiança
IFNG – Gene do interferão gama
IFNG - Gene que codifica o IFN γ
IFNGR1 – Gene do receptor 1 do interferão gama
IFN γ – Interferão gama
IgG – Imunoglobulina G
IgM – Imunoglobulina M
IL12 – Interleucina 12
IL12RB1 - Gene do recetor β 1 da interleucina-12
INFR1 - Gene do recetor 1 do IFN γ
INFR2 – Gene do receptor 2 do interferão gama
INFR2 - Gene do recetor 2 do IFN γ
INSA - Instituto Nacional de Saúde Doutor Ricardo Jorge
Kb – Quilo base
LCV – *Large cell variants*
LD – Desiquilíbrio de ligação
LPS – Lipopolissacarideo
MAF – *Minor allele frequency* – Freqência do alelo minoritário
Mb – Mega par de bases
MgCl₂ – Cloreto de magnésio
mL – Mililitros
mM - Mlimolar
MMPs – *Matrix metalloproteinases* – Metaloproteinases de matriz

MPM – Marcador de peso molecular
mRNA – Messenger Ribonucleic acid - Ácido ribonucleico mensageiro
NA – Não aplicável
ng – Nanogramas
NGS – *Next Generation Sequencing* - Sequenciação de nova geração
(NH₄)₂SO₄ – Sulfato de amónia
nm- Nanómetros
NRAMP1 – gene da resistência natural associada à proteína macrofágica
°C – Graus Célsius
OR *odds ratio*
P2RX7 - gene do recetor purigénico P2X7
pb – Par de Bases
PCR – *Polimerase chain reaction* – Reacção em Cadeia da Polimerase
RNA – *Ribonucleic acid* – Ácido ribonucleico
SCV– *small-cell variants*
SDC – *small dense cells*
SNP – *Single Nucleotide Polymorphism* – Polimorfismo de nucleótido único
T – Timina
tagSNP – SNP representativo numa região do genoma com alto desequilíbrio de ligação
Taq DNA Polimerase – Polimerase termorresistente do DNA derivada do *Thermus aquaticus*
TBE – Tris Borato EDTA
Th1 – *Type 1 T helper cells* – Células T auxiliares do tipo 1
Th2 – *Type 2 T helper cells* – Células T auxiliares do tipo 1
TIMP1 – Gene que codifica o inibidor tecidual 1 de metaloproteínas
TIMP1 – Gene que codifica o *Tissue Inhibitor of Metalloproteinase 1* – Inibidor tecidual da metaloproteinase 1
TLR1 – *Toll like receptor 1* – Recetores do tipo *toll 1*
TLR2 – *Toll like receptor 2* – Recetores do tipo *toll 2*
TLR4 – *Toll-like receptor 4* – Recetores do tipo *toll 4*
TLR1 e *TLR2* – Genes dos recetores *Toll-like 1* e *2*
TMEM50B – Gene da proteína transmembranar 50B
TNF – Fator de Necrose Tumoral
U – Unidade Internacional
V – Volts
WGAS – *Whole genome association studies* – Estudos de associação de genoma completo
X² – Qui-quadrado
% – Percentagem
µg – Microgramas
µL – Microlitros
µm – Micrómetro
µM – Micromolar

Lista de Publicações

Poster:

S David¹, L Castro, M Machado, U Gaspar, JC Robles-Fuentes, J Mendonça, J Ferrão, L Vieira AS Santos (2022, March 24th & 25th). *Genetic determinants of infection in chronic Q fever* [Poster presentation]. 19th Portugaliae Genetica Phenotypic Manifestation of Genetic Heterogeneity, I3S, Porto, Portugal

(Anexo 1)

Índice

Capítulo I - Introdução	1
1.1. Febre Q	1
1.1.1. Definição	1
1.1.2. Contexto Histórico	1
1.1.3. Caracterização do agente etiológico	1
1.1.4. Epidemiologia	4
1.1.4.1. Distribuição Mundial	4
1.1.4.2. Reservatórios e vias de transmissão	5
1.1.5. Patogénese	6
1.1.6. Febre Q aguda vs febre Q crónica	7
1.1.7. Diagnóstico laboratorial	8
1.1.8. Resposta imunológica do hospedeiro	9
1.1.9. Prova de princípio sobre a base genética das doenças infecciosas	12
1.1.10. Implicações para a escolha de genes candidatos na imunidade mediada pelo IFN γ	13
1.1.11. Variantes genéticas e a febre Q	14
1.2. Estudos de genes candidatos (<i>Candidate gene association studies</i> – CGAS)	17
1.2.1. Etapas de CGAS	17
Capítulo II – Objetivos	19
Capítulo III - Materiais e Métodos	20
3.1. Declaração de ética	20
3.2. Caracterização da população estudada	20
3.3. Amostras biológicas	21
3.4. Métodos	21
3.4.1. Estudo exploratório	21
3.4.1.1. Preparação das amostras	22
3.4.1.1.1. Desenho de primers	22
3.4.1.1.2. Amplificação por PCR	24
3.4.1.1.3. Eletroforese em gel de agarose	26
3.4.1.1.4. Quantificação visual dos produtos de PCR	27
3.4.1.1.5. <i>Pooling</i> das amostras	27
3.4.1.1.6. Purificação dos produtos de PCR	27
3.4.1.2. Sequenciação NGS	28
3.4.1.2.1. Preparação das bibliotecas NGS	28
3.4.1.2.1.1. Tagmentação	29
3.4.1.2.1.2. Amplificação de bibliotecas	29
3.4.1.2.1.3. Limpeza de bibliotecas	29
3.4.1.2.1.4. Normalização das bibliotecas	29
3.4.1.2.1.5. Diluição e desnaturação de bibliotecas	29
3.4.1.2.2. Fase I	30
3.4.1.2.3. Fase II	31
3.4.1.2.3.1. Mapeamento	31
3.4.1.2.3.2. Variant call	31
3.4.1.2.3.3. Anotação	31
3.4.1.2.3.4. Filtragem	32
3.4.1.2.4. Fase III	32
3.4.1.2.4.1. Controlo de Qualidade dos dados	33
3.4.1.2.4.2. Associação Genótipo-Fenótipo	34

3.4.1.2.4.3. Análise dos SNPs em associação	34
3.4.2. Estudo Réplica	35
3.4.2.1. Preparação das amostras de DNA	35
3.4.2.1.1. Seleção de variantes	35
3.4.2.1.2. Desenho de <i>primers</i>	36
3.4.2.1.3. Amplificação de DNA por PCR	36
3.4.2.1.4. Eletroforese em gel de agarose	37
3.4.3. Sequenciação automática pelo método de Sanger	38
3.4.3.1. Reação de Purificação	38
3.4.3.2. Reação de Sequenciação Sanger	38
3.4.4. Análise de dados	39
Capítulo IV – Resultados	41
4.1. Resultados do Estudo Exploratório	41
4.2. Resultados do Estudo Réplica	58
Capítulo V – Discussão	64
Capítulo VI – Conclusão	70
Capítulo VII – Perspetivas futuras	71
Referências bibliográficas:	72
Anexos	90

Índice de figuras

Figura 1.1. Ciclo de multiplicação da <i>Coxiella burnetii</i> numa célula eucariótica.....	3
Figura 1.2. Potenciais ciclos de transmissão de <i>Coxiella burnetii</i> entre diferentes hospedeiros.....	6
Figura 1.3. Características clínicas e estádios da febre Q.	8
Figura 1.4. Etapas dos estudos populacionais de associação genótipo-fenótipo com base em genes candidatos.....	18
Figura 3. 1. Organização dos <i>index</i> , visualização superior da placa de 96 poços.	30
Figura 3. 2. Fluxo da metodologia NGS em fases (I, II, III), indicação das ferramentas bioinformáticas usada em cada passo com os respetivos parâmetros de qualidade.	32
Figura 3. 3. Fluxo de trabalho da sequenciação Sanger por eletroforese capilar.	38
Figura 4. 1. Resultado eletroforese de cada fragmento dos genes <i>IL12RB1</i> , <i>IFNGR1</i> e <i>IFNGR2</i> , <i>NRAMP1</i> , <i>TLR1</i> , <i>TLR2</i> e <i>TIMP1</i>	41
Figura 4. 2. Resultado eletroforese dos 5 agrupamentos de 4 fragmentos cada.....	42
Figura 4. 3. Fase III do Controlo de Qualidade (CQ) dos dados de genotipagem	43
Figura 4. 4. Gráficos <i>pie chart</i> relativos às frequências alélicas da variante rs78407108 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	48
Figura 4. 5. Gráficos <i>pie chart</i> relativos às frequências genotípicas da variante rs78407108 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	48
Figura 4. 6. Gráficos <i>pie chart</i> relativos às frequências genotípicas da variante rs17879956 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	50
Figura 4. 7. Gráficos <i>pie chart</i> relativos às frequências genotípicas da variante rs17879956 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	50
Figura 4. 8. Gráficos <i>pie chart</i> relativos às frequências genotípicas da variante rs7277167 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	52
Figura 4. 9. Gráficos <i>pie chart</i> relativos às frequências alélicas da variante rs7277167 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	52
Figura 4. 10. Gráficos <i>pie chart</i> relativos às frequências genotípicas da variante rs9974603 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	54
Figura 4. 11. Gráficos <i>pie chart</i> relativos às frequências alélicas da variante rs9974603 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.....	54
Figura 4. 12. Padrões de desequilíbrio de ligação na região entre 1000pb a montante e 1000pb a jusante da variante rs9974603 de acordo com o HapMap, visualizado usando o software Haploview (Barret <i>et al.</i> , 2005) na subpopulação CEU.	56
Figura 4. 13. Resultado eletroforese dos fragmentos que contêm as variantes do gene <i>P2X7R</i> (rs1718119 e rs3751143) e do gene <i>INFG</i> (rs2430561 e rs1861493) para as amostras de doentes com febre Q crónica.....	59
Figura 4. 14. Eletroferograma parcial do gene <i>P2X7R</i> mostrando a variante rs1718119 assinalada a azul para os genótipos GG (I), GA (II) e AA (III). N= alelos G e A.....	60
Figura 4. 15. Eletroferograma parcial do gene <i>P2X7R</i> mostrando a variante rs3751143 assinalada a azul para os genótipos AA (I), AC (II) e CC (III). N= alelos A e C.....	60
Figura 4. 16. Eletroferograma parcial do gene <i>P2X7R</i> mostrando a variante rs2430561 assinalada a azul para o genótipo homocigótico AA (I).	61

Figura 4. 17. Eletroferograma parcial do gene *P2X7R* mostrando a variante rs1861493 assinalada a azul para os genótipos AA (I), AG (II) e GG (III). N= alelos A e G. 61

Índice de Tabelas

Tabela 1-1. SNPs identificados pelo Método CGAs como estando associados a febre Q na população holandesa.....	16
Tabela 3. 1. Fragmentos a amplificar dos genes <i>IL12RB1</i> , <i>IFNGR1</i> , <i>IFNGR2</i> ; <i>NRAMPI</i> ; <i>TLR1</i> , <i>TLR2</i> e <i>TIMP1</i>	23
Tabela 3. 2. Primers escolhidos para amplificar os fragmentos dos genes <i>IL12RB1</i> , <i>IFNGR1</i> , <i>IFNGR2</i> , <i>NRAMPI</i> , <i>TLR1</i> , <i>TLR2</i> e <i>TIMP1</i>	24
Tabela 3. 3. Misturas de reação para os fragmentos a amplificar dos genes <i>IL12RB1</i> , <i>IFNGR1</i> , <i>IFNGR2</i> ; <i>NRAMPI</i> ; <i>TLR1</i> , <i>TLR2</i> e <i>TIMP1</i>	25
Tabela 3. 4. Condições de PCR adequadas para amplificação dos fragmentos pela <i>TaKaRa LA Taq Hot Start</i>	26
Tabela 3. 5. Variantes escolhidas para o estudo réplica	35
Tabela 3. 6. Sequência dos primers e tamanho dos fragmentos amplificados dos genes <i>P2RX7</i> e <i>IFNG</i> 36	
Tabela 3. 7. Mistura reacional para amplificação dos fragmentos dos genes <i>P2RX7</i> e <i>IFNG</i>	37
Tabela 3. 8. Condições da amplificação por PCR dos fragmentos dos genes <i>P2RX7</i> e <i>IFNG</i>	37
Tabela 3. 9. Mistura da reação de sequenciação Sanger	38
Tabela 3. 10. Condições da sequenciação cíclica pelo método de Sanger.....	39
Tabela 4. 1. Resultados estatisticamente significativos para o teste alélico de associação com a febre Q crónica.....	45
Tabela 4. 2. Frequências alélicas e genotípicas das variantes do gene <i>IFNGR2</i> com associação significativa que resistiram ao critério de Bonferroni (0,00016339869) em diferentes populações do Projeto 1000 Genomas e na população portuguesa estudada	47
Tabela 4. 3. Variantes em desequilíbrio de ligação ($R^2 \geq 0,8$) com os SNPs que mostraram uma relação estatisticamente significativa, resistindo ao critério de Bonferroni, em associação com a febre Q crónica.....	57
Tabela 4. 4. Número de amostras sequenciadas pela Sequenciação de Sanger	58
Tabela 4. 5. Associação da variante rs1718119 com o aumento do risco de febre Q crónica sob diferentes modelos de hereditariedade.....	62
Tabela 4. 6. Associação da variante rs3751143 com o aumento do risco de febre Q crónica sob diferentes modelos de hereditariedade.....	62
Tabela 4. 7. Associação da variante rs2430561 com o aumento do risco de febre Q crónica sob diferentes modelos de hereditariedade	63
Tabela 4. 8. Associação das variantes rs1861493 com o aumento do risco de febre Q crónica sob diferentes modelos de hereditariedade	63

Capítulo I - Introdução

1.1. Febre Q

1.1.1. Definição

A febre Q é uma zoonose de distribuição mundial causada pela bactéria intracelular *Coxiella burnetii*. Esta é uma doença complexa e polimórfica que pode manifestar-se na forma aguda autolimitada, com um curso ligeiro a moderado e prognóstico benigno, ou na forma persistente ou crónica, geralmente localizada e de evolução grave ou potencialmente fatal. A febre Q tanto ocorre em casos isolados, como em contexto de surtos, alguns com importantes implicações em saúde pública e grande impacto económico (Santos, 2015). O *Centre for Disease Control and Prevention* (CDC) classificou a *C. burnetii* como agente de bioterrorismo do Grupo B (CDC, 2011), e a febre Q ocupa um lugar entre as 13 zoonoses de prioridade global (ILRI, 2012).

1.1.2. Contexto Histórico

As primeiras referências à febre Q datam de 1935, quando Edward Holbrook Derrick, descreveu uma série de casos de doença febril, provavelmente infecciosa, que afetava os trabalhadores de um matadouro em Brisbane, Austrália. Dada a etiologia incerta da doença, esta foi denominada por *Query fever* (Derrick, 1937). O agente implicado na febre Q, viria, pouco depois, a ser isolado (Burnet e Freeman, 1937; Cox e Bell, 1939) e designado de *Coxiella burnetii*, em homenagem a Cox e Burnet, investigadores responsáveis pelo seu isolamento (Philip, 1948).

1.1.3. Caracterização do agente etiológico

A *Coxiella burnetii* é um pequeno cocobacilo de aproximadamente 0,4-1 µm de tamanho que tem as características da parede celular de uma bactéria gram-negativa. É um organismo intracelular obrigatório *in vivo* e tem um tropismo por células mieloides, preferencialmente, monócitos e macrófagos (Maurin e Raoult, 1999).

C. burnetii é a única espécie que oficialmente integra o género *Coxiella*. Este agente foi originalmente classificado na ordem *Rickettsiales*, família *Rickettsiaceae*, tribo *Rickettsiae*, por partilhar com as rickettsias algumas características morfológicas e ecológicas. Todavia, a aplicação de técnicas moleculares à filogenia questionou essa taxonomia tradicional, demonstrando que *C. burnetii* se distancia dos demais microrganismos da ordem *Rickettsiales*, que são α -proteobactérias, aproximando-se de algumas γ -proteobactérias, como a *Legionella spp.* e *Francisella tularensis*, o que levou à sua recente reclassificação na ordem *Legionellales*. O género *Coxiella* poderá vir a

sofrer mais uma revisão para integrar microrganismos que têm vindo a ser descritos como semelhantes a *C. burnetii*, como é o caso de vários *Coxiella-like organisms* (CLO) (Mori *et al.*, 2017).

A *Coxiella burnetii* possui várias características distintivas que permitem a sua sobrevivência, nomeadamente uma alta transmissibilidade (Sandoz *et al.*, 2016) e a variação antigénica do seu lipopolissacarídeo (LPS) (Williams *et al.*, 1981).

A alta transmissibilidade da *C. Burnetii* é devida, parcialmente, ao facto de poder adotar três formas distintas: *large cell variants* (LCV), *small-cell variants* (SCV) e *small dense cells* (SDC). As três formas são morfológica e metabolicamente diferentes, e também apresentam diferenças na sua resistência física e química (Arricau-Bouvery e Rodolakis, 2005).

As SDC e as SCV representam as pequenas variantes morfológicas das bactérias, são metabolicamente inativas e muito resistentes a agentes químicos e físicos, como o calor, radiação UV, dessecação e pressão osmótica, o que favorece a sua persistência e transmissão no meio ambiente (Arricau-Bouvery e Rodolakis, 2005; OIE, 2018). Para além disso, as SDC podem ser consideradas pseudo-esporos (Guatteo, 2013). Pelo contrário, as LCV são metabolicamente ativas e correspondem à forma intracelular do organismo (Oyston e Davies, 2011), não intervêm na transmissão entre indivíduos, mas sim na disseminação da bactéria no organismo (Guatteo, 2013).

A formação das diferentes variantes morfológicas está ligada ao ciclo de vida de *C. burnetii*, uma estratégia desenvolvida para sobreviver dentro e fora do vacúolo parasitóforo. A passagem das SCV para LCV na célula hospedeira está bem documentada, no entanto a formação das SCV e das SDC ainda não foi devidamente estudada (Arricau-Bouvery e Rodolakis, 2005).

O ciclo de multiplicação da *C. burnetii* no interior de uma célula eucariótica encontra-se descrito na Figura 1.1. As SCV ou os pseudo-esporos ligam-se à membrana celular dos monócitos que circulam na corrente sanguínea e macrófagos localizados nos tecidos, penetrando nas células por fagocitose. No interior da célula residem num endossoma, onde ocorre a acidificação do meio (pH 5,5). O pH baixo parece desencadear a conversão das SCV ou dos pseudo-esporos a LCV. Após a formação da LCV ocorre a fusão do endossoma com um lisossoma, formando-se um fagolisossoma. Posteriormente, ocorre nova acidificação do meio (pH 4,5) e as LCV diferenciam-se novamente em SCV ou endosporos. Por fim sucede-se a libertação dos esporos e das SCV para o meio extracelular (Arricau-Bouvery e Rodolakis, 2005; Guatteo, 2013).

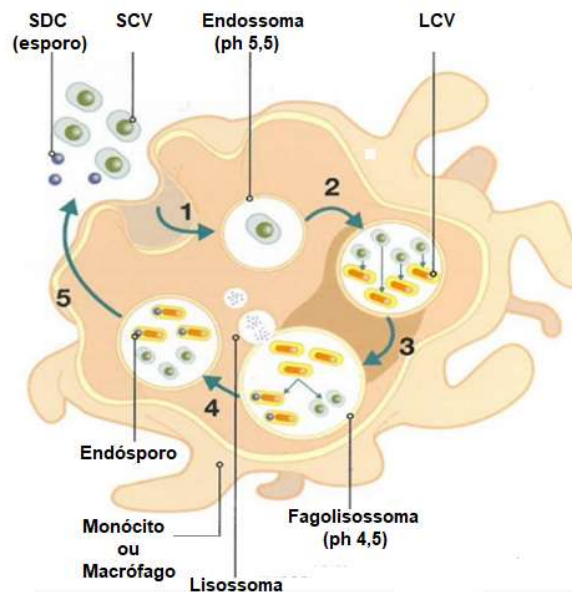


Figura 1.1. Ciclo de multiplicação da *Coxiella burnetii* numa célula eucariótica.

Legenda: -1. Entrada do esporo ou SCV na célula eucariótica e acidificação do endossoma. 2- Multiplicação da SVC por divisão binária e diferenciação em LCV. 3- Fusão do endossoma com o lisossoma e acidificação do fagolisossoma. 4- Multiplicação do LCV por divisão binária, diferenciação LCV em SCV e desenvolvimento do endósporo. 5- Liberação do endósporo e SCV para o exterior da célula.

Fonte: Adaptado de Guatteo, 2013.

O lipopolissacarídeo (LPS) é considerado um dos principais determinantes da expressão de virulência e infecção de *C. burnetii*, desempenhando um papel importante não apenas no crescimento e desenvolvimento do microrganismo, mas também na patogênese e imunidade da febre Q (Sireci *et al.*, 2021). O LPS de *C. burnetii* está envolvido em estratégias de evasão imune, uma vez que, a sua estrutura peculiar leva a efeitos repressivos sobre os mecanismos de defesa das células (Abnave *et al.*, 2017).

A bactéria pode apresentar duas variações antigênicas do seu LPS de membrana: LPS da fase I virulenta (LPS I) e de fase II avirulenta (LPS II). As estirpes de *C. burnetii* de tipo selvagem podem ser encontradas em animais, artrópodes ou humanos infetados (Maurin e Raoult, 1999; OIE, 2018). Estas possuem um LPS de comprimento total complexo com unidades sacarídicas de cadeia "O" específicas que tornam a bactéria virulenta (Hackstadt, 1986). Após a passagem repetida em culturas de ovos embrionados ou em linhas celulares, ou seja, na ausência de intervenção do sistema imunológico, as estirpes de *C. burnetii* do tipo selvagem, convertem-se na fase II, produzindo um LPS de fase II determinando a perda de virulência (Lukáčová *et al.*, 2008). O LPS de fase II é facilmente eliminado pela resposta imune em hospedeiros imunocompetentes (Hackstadt, 1986). Esta passagem é acompanhada por

modificações tanto na composição quanto na estrutura da macromolécula LPS (Toman *et al.*, 2009). O LPS II apresenta uma estrutura rugosa (R), ao contrário do LPS I, caracterizado por um aspeto liso (S). Além disso, o LPS II possui os açúcares, virenose e di-hidroxistreptose, não encontrados noutros LPS e que, portanto, são biomarcadores únicos do LPS de *C. burnetii* (Narasaki e Toman, 2012). A variação de fase antigénica está relacionada com deleções génicas que levam a perda de açúcares que compõe o LPS, reduzindo o seu tamanho. O LPS II está inserido no LPS I e é imunodominante, dessa forma, os anticorpos anti-fase II estão sempre presentes, independentemente da fase da doença (Rousset *et al.*, 2012).

A variação de fase também é usada no diagnóstico de febre Q, uma vez que os estádios agudo e crónico correspondem a duas fases antigénicas distintas na resposta humoral. Na infeção aguda, a *C. burnetii* induz anticorpos contra a fase II (antígenos proteicos), enquanto que a infeção crónica, frequentemente manifestada como endocardite, está associada à produção de altos títulos de anticorpos dirigido contra a fase I (antígeno LPS) (Peacock *et al.*, 1983).

1.1.4. Epidemiologia

1.1.4.1. Distribuição Mundial

Após a sua descoberta na Austrália, vários surtos epidémicos e casos isolados de febre Q começaram a ser progressivamente identificados, sendo esta, atualmente uma doença de distribuição mundial, com exceção da Nova Zelândia (Hilbink *et al.*, 1993; Eldin *et al.*, 2017).

O último grande surto de febre Q ocorreu na Holanda entre 2007 e 2009 e provou que esta zoonose pode tornar-se uma séria ameaça à saúde pública com um impacto económico elevado (Roest *et al.*, 2011; Van der Hoek *et al.*, 2012). Foram reportados cerca de 4000 casos agudos e pelo menos duas centenas de casos crónicos. Este surto determinou ainda avultados gastos em medidas profiláticas e de contenção, na ordem dos milhões de euros, que culminaram com o abate de mais de 50000 animais gestantes, o que representou em muitas explorações cerca de 60% do efetivo pecuário (Van der Hoek *et al.*, 2012).

Ocorreram, ainda, outros surtos significativos, nomeadamente, em Espanha (Errasti *et al.*, 1984), Suíça (Dupuis *et al.*, 1987), Inglaterra (Guigno *et al.*, 1992), Alemanha (Schneider *et al.*, 1993), Eslováquia (Kováčová *et al.* 1998), Polónia (Chmielewski *et al.*, 2013), Austrália (Graves e Stenos, 2009; Graves e Islam, 2016), e

ainda em alguns países de África (Angelakis *et al.*, 2014) e do Médio Oriente (White *et al.*, 2013).

Segundo dados da Direção Geral de Saúde, em Portugal, os primeiros registos de casos de infeção por *C. burnetii* surgiram em 1948, tendo a febre Q passado a integrar a lista das Doenças de Declaração Obrigatória (DDO) desde 1999, ainda que nunca se tenha verificado nenhum surto epidemiológico. A febre Q é considerada endémica no continente, com o maior número de casos notificados na região Centro e Sul e com uma incidência média de 0,10 casos por 100000 habitantes ano (1999-2007) (DGS, 2010).

Apesar da febre Q ser uma doença humana de declaração obrigatória, defende-se que em muitos casos a doença fica por diagnosticar pela inespecificidade do quadro clínico e fraco contexto epidemiológico com que muitos doentes se apresentam, considerando-se que de uma maneira geral a realidade da febre Q está subestimada em muitos dos países onde ocorre (Santos, 2007; Santos, 2009; Santos, 2019).

1.1.4.2. Reservatórios e vias de transmissão

A *C. burnetii* encontra-se largamente disseminada na natureza, tendo como principais reservatórios as carraças, as aves e os mamíferos. Às carraças é atribuído um papel essencial na transmissão do agente entre os animais silvestres e na amplificação dos ciclos selvagem aos animais domésticos (Mori *et al.*, 2017). Relativamente à transmissão de *C. burnetii* ao Homem, apesar de qualquer um destes grupos de animais constituir uma fonte potencial de infeção, é no contacto com mamíferos domésticos, sobretudo gado bovino, ovino e caprino, que reside a base fundamental da epidemiologia da febre Q (Maurin e Raoult, 1999; Raoult, 2000). Os animais infetados são geralmente assintomáticos, salvo alguns distúrbios reprodutivos, que passam, muitas vezes, despercebidos. O agente é eliminado nas secreções e excreções como leite, fezes, urina, sémen, secreções vaginais e produtos do parto/aborto. Neste contexto é particularmente importante a parição, por estimular a proliferação de *C. burnetii* na placenta e as épocas de parição, devido às quantidades massivas de microrganismos que são eliminados para o ambiente durante o parto ou aborto (Maurin e Raoult, 1999).

A infeção humana é adquirida essencialmente pela inalação de aerossóis ou poeiras contendo esporos do agente resultantes do contacto direto com animais infetados e seus produtos ou com o ambiente contaminado por estes. A alta resistência do microrganismo às condições ambientais e a possibilidade de difusão por correntes

de ar, pode também originar infecções indiretas, de diagnóstico difícil pela ausência de epidemiologia compatível (Dupuis *et al.*, 1987; Tissot-Dupont *et al.*, 2004). O contacto com indivíduos infetados é ainda outra forma de adquirir febre Q, embora, na prática, esta seja uma situação rara. Outras vias alternativas de infecção são também referidas na literatura, mas tidas como excepcionais, como a via sexual, transfusional e transplacentária (Santos, 2007). Na Figura 1.2. estão representadas as potenciais vias de transmissão da *C. burnetii*.

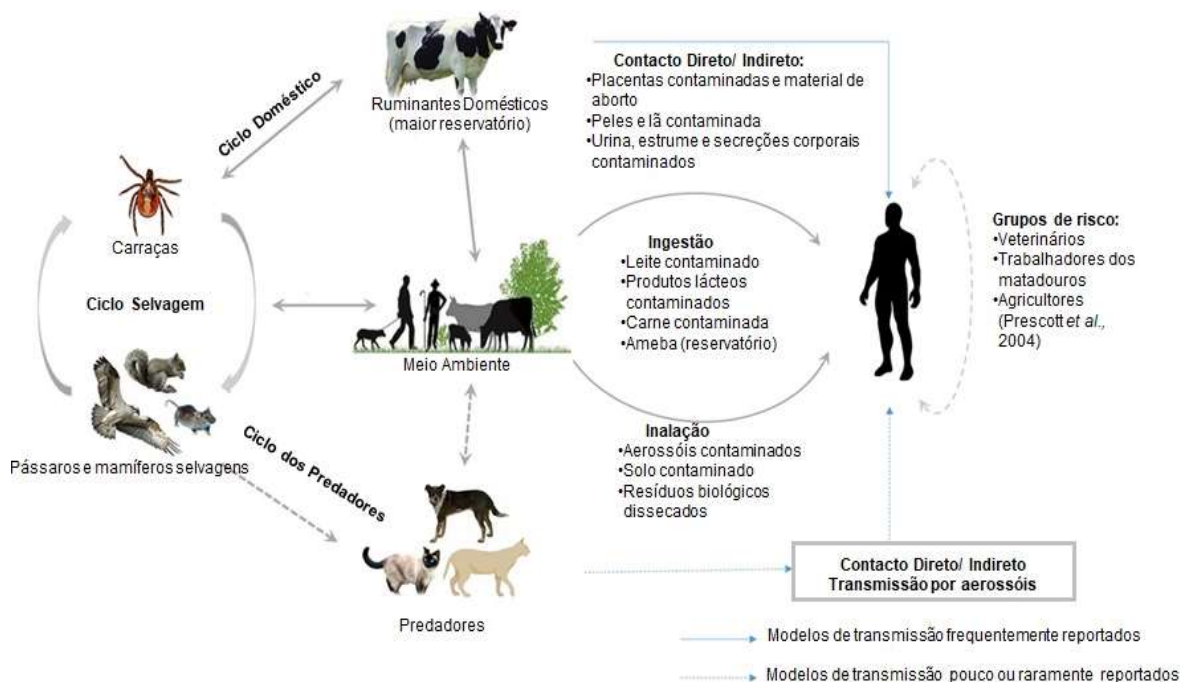


Figura 1.2. Potenciais ciclos de transmissão de *Coxiella burnetii* entre diferentes hospedeiros.

Fonte: Adaptado de Sahu *et al.*, 2021.

A doença ocorre entre os 30 e os 60 anos, apesar de algumas referências a idades mais extremas (Marrie, 1990; Maltezos e Raoult, 2002) afetando principalmente o sexo masculino. No entanto, apesar da predominância no sexo masculino, a seroprevalência apresenta um rácio homem: mulher de 1:1, motivo pelo qual se pensa que após a exposição a *C. burnetii* os dois géneros são afetados de forma igual, mas o sexo masculino é mais sintomático (Raoult *et al.*, 2005).

1.1.5. Patogénese

A infecção primária ocorre após o contacto do indivíduo não imune com *C. burnetii* (Guatteo, 2013). Dado que a via aerogénica é a principal via de infecção, os pulmões são o primeiro local onde ocorre a proliferação do agente, que posteriormente se dissemina a outros órgãos através da circulação sanguínea (Maurin e Raoult, 1999).

Após o período de incubação, geralmente de duas a três semanas, a infecção desenvolve-se em cerca de 60% dos indivíduos de forma assintomática e apenas os restantes 40% apresentam sinais clínicos (Dupuis *et al.*, 1985; Guatteo, 2013). Em aproximadamente 2% das infecções, *C. burnetti* pode multiplicar-se mesmo após a resposta imune à infecção primária (Guatteo, 2013). Quando o sistema imunitário é insuficiente para controlar a infecção, desenvolve-se a forma crónica da doença (Maurin & Raoult, 1999). A patologia e características clínicas da febre Q em humanos é apresentada na Figura 1.3.

1.1.6. Febre Q aguda vs febre Q crónica

O sinal clínico característico da febre Q aguda em humanos é a síndrome febril que, em regra, regride espontaneamente sem tratamento (Guatteo, 2013), no entanto, 2 a 5% dos indivíduos desenvolvem complicações que requerem tratamentos médicos (Dupuis *et al.*, 1985; Guatteo, 2013; Sahu *et al.*, 2021). O quadro clínico caracteriza-se normalmente por febre, cansaço, dores de cabeça e mialgias, todavia os casos mais severos são frequentemente acompanhados de hepatite e/ou pneumonia atípica. (Angelakis e Raoult, 2010; Guatteo, 2013; Sahu *et al.*, 2021). Embora a maior parte das situações apresente um prognóstico favorável, as hospitalizações associadas à febre Q aguda evidenciam a morbilidade que lhe está associada (Santos, 2019).

A maior relevância e gravidade é devida à persistência da infecção e ao desenvolvimento de situações crónicas o que acontece em cerca de 1% a 5% dos casos agudos (Schoffelen *et al.*, 2015; Sahu *et al.*, 2021). Normalmente, a doença crónica instala-se no período de seis meses a um ano (Peacock *et al.*, 1983), no entanto, esta pode, igualmente, manifestar-se anos após a infecção inicial (Gami *et al.*, 2004).

A gravidez, imunodepressão (patológica ou induzida), doença valvular, vascular ou osteoarticular são tidos como fatores de risco para o desenvolvimento da febre Q crónica (Angelakis e Raoult, 2010). Cerca de 60% a 70% dos casos crónicos, apresentam endocardite (Maurin e Raoult, 1999), quase sempre de evolução lenta, mas muito insidiosa (Santos, 2019). Pessoas com deficiências nas válvulas cardíacas apresentam maior predisposição para o desenvolvimento de endocardites (Guatteo, 2013). A infecção vascular é a segunda apresentação mais frequente neste estágio da doença, no entanto, podem ser observadas outras manifestações clínicas tais como infecções osteoarticulares, incluindo osteomielite, osteoartrite, hepatite crónica ou um mal-estar geral caracterizado por uma fadiga crónica (Angelakis e Raoult, 2010). Além disso, a infecção em mulheres grávidas pode causar nados-mortos ou abortos

espontâneos sem sintomas clínicos anteriores (Million e Raoult, 2015; Ghaoui *et al.*, 2018).

A febre Q crónica pode ser fatal se não tratada atempadamente, mas mesmo com uma intervenção antibiótica adequada, o regime necessário é agressivo e exige um tratamento prolongado com a possibilidade de efeitos secundários, mas com o perigo de recidivas se suspenso antecipadamente (Santos, 2007; Santos, 2019).

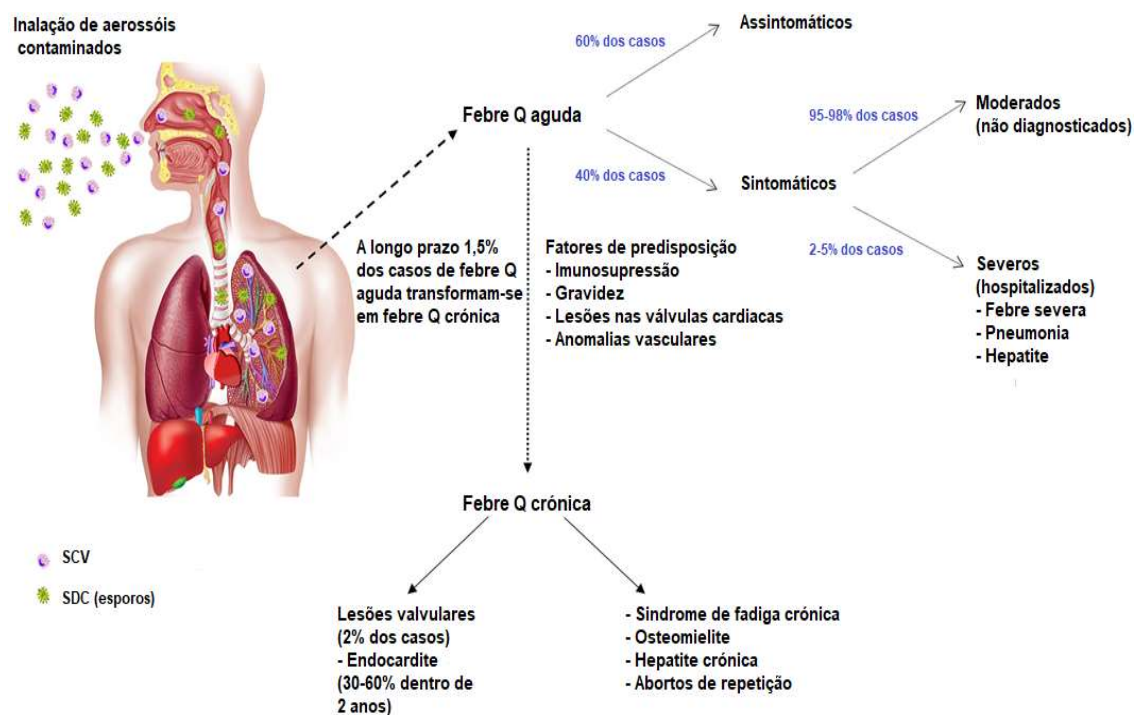


Figura 1.3. Características clínicas e estádios da febre Q.

Fonte: Adaptado de Sahu *et al.*, 2021.

1.1.7. Diagnóstico laboratorial

O diagnóstico laboratorial de febre Q apoia-se essencialmente na serologia, com a avaliação dos títulos de IgG e IgM anti-*C. burnetii* em fase I e fase II, sendo a imunofluorescência indireta (IFA) uma das técnicas de referência mais usadas para o efeito. Um caso de febre Q aguda é confirmado sempre que se verifique a presença de títulos de IgM e IgG de fase II numa amostra isolada, bem como, seroconversão ou aumento de 4x no título de IgG em duas amostras consecutivas, colhidas com um intervalo de duas a três semanas. As situações de febre Q crónica são sugeridas por um alto título de IgG de fase I, mas este deve ser sempre interpretado no contexto clínico do doente. Este último aspeto é particularmente importante, realçando-se que a serologia não deve, por si só, definir um caso crónico. Na verdade, títulos altos de anticorpos anti-fase I podem estar presentes sem relação direta com situações crónicas.

Por exemplo podem ocorrer em doentes agudos, de forma transitória já numa fase de remissão da doença, e em casos muito excepcionais podem mesmo permanecer durante vários anos após a exposição, sem que haja desenvolvimento de doença (Santos, 2019).

Quanto ao diagnóstico laboratorial direto, o isolamento do agente (por inoculação em células de crescimento *in-vitro*) ou a deteção molecular (por PCR), são testes particularmente úteis para a confirmação de caso, tanto numa fase inicial da infecção aguda e ainda quando os resultados serológicos são negativos, como numa situação crónica. No primeiro caso, a amostra mais adequada é o sangue periférico recolhido em tubo com anticoagulante EDTA. Outras amostras que sejam consideradas pertinentes atendendo à focalização da doença aguda podem também ser utilizadas, tais como expetoração, lavado broncoalveolar, biópsias, etc. Nas situações crónicas, para além de sangue em EDTA, são particularmente úteis biópsias ou peças cirúrgicas recolhidas face ao quadro da infecção, tais como válvulas cardíacas no caso de endocardites, material vascular de aneurismas, fragmentos ósseos em situações de osteomielites, material placentário ou abortivo, etc. Em qualquer dos casos, é importante complementar os testes diretos com a pesquisa de anticorpos circulantes por serologia (Santos, 2007; Santos 2019).

1.1.8. Resposta imunológica do hospedeiro

No hospedeiro, a *C. burnetii* ultrapassa as primeiras barreiras defensivas do corpo; consegue sobreviver no muco nasal, flora do trato urogenital e às enzimas dos fluídos corporais. Uma vez dentro do organismo, a *C. burnetii* é intercetada, principalmente por monócitos, macrófagos, células *Natural Killer* (NK) e células dendríticas (Ghigo *et al.*, 2009).

O curso da infecção por *C. burnetii* depende da eficácia da resposta imune. A imunidade mediada por células desempenha um papel central e as células T são críticas para a eliminação de *C. burnetii* (Honstetter *et al.*, 2004; Andoh *et al.*, 2007; Shannon e Heinzen, 2009; Capo e Mege, 2012; Cunha *et al.*, 2015; Eldin *et al.*, 2017).

A ativação das células mieloides induzida pela ligação da *C. burnetii* de fase I aos recetores celulares, nomeadamente ao TLR4, origina uma reorganização do citoesqueleto e as células infetadas respondem à infecção com uma regulação geral positiva de citocinas relacionadas a uma resposta imunitária mediada por células T auxiliares do tipo 1 (*type 1 T helper cells* -Th1) (Honstetter *et al.*, 2004; Zamboni *et al.*, 2004; Shannon e Heinzen, 2009; Che *et al.*, 2011). Essa resposta inflamatória resulta

no controlo precoce da infeção pela produção de intermediários reativos de nitrogénio e oxigénio que levam à morte intracelular do agente patogénico (Brennan *et al.*, 2004; Andoh *et al.*, 2007). No entanto, *C. burnetii* desenvolveu a capacidade de limitar a resposta pró-inflamatória para resistir à resposta imune do hospedeiro (Capo *et al.*, 1999; Shannon *et al.*, 2005; Barry *et al.*, 2011; Capo e Mege, 2012; Graham *et al.*, 2013; Sobotta *et al.*, 2016). A subversão da resposta protetora está relacionada com a estrutura do LPS. Os LPS de fase I mascaram o reconhecimento bacteriano pelo recetor TLR4 o que resulta numa resposta inflamatória menos intensa que permite a persistência de *C. burnetii* sem maturação significativa de células dendríticas que passam a produzir níveis baixos de IL-12 e TNF α (Zamboni *et al.*, 2004; Shannon *et al.*, 2005; Waag, 2007; van Schaik *et al.*, 2013). Em oposição, os LPS de fase II são prontamente detetados pelo TLR2, induzindo a ativação de células dendríticas e secreção de altos níveis de IL-6, de IL-12 e TNF α . Os macrófagos são, então, ativados para mediar a depuração bacteriana (Shannon *et al.*, 2005; van Schaik *et al.*, 2013).

Os monócitos e macrófagos são os principais alvos de *C. burnetii* (Eldin *et al.*, 2017). As células pertencentes à linhagem monócito-macrófago expressam propriedades funcionais polarizadas (Gordon e Taylor, 2005). Essa polarização parece estar intimamente relacionada com a capacidade de controlo da infeção por *C. burnetii*, explicando a persistência bacteriana em infeções crónicas (Benoit *et al.*, 2008). Os macrófagos polarizados M1 segregam citocinas inflamatórias (IL-1b, TNF, IL-12 e IL-6), quimiocinas (CCL2, CCL5 e CXCL8) e expressam os recetores de superfície CCR7, CD80, TLR-2 e TLR-4. Eles também segregam altos níveis de iNOS e intermediários reativos de nitrogénio, como nitritos (Honstetter *et al.*, 2004; Benoit *et al.*, 2008). Os macrófagos M2 segregam altos níveis de citocinas inflamatórias (IL-10, IL-1, antagonista do recetor IL-1Ra e TGF-b1), quimiocinas (CCL16, CCL17, CCL18 e CCL24) e expressam os recetores de superfície CXCR1, recetor de manose (MR), CD14 e arginase-1 que induz o bloqueio da via da iNOS. Classicamente, macrófagos polarizados em M1 são induzidos por LPS, IFN γ e TNF α , participam da resistência contra patogenos intracelulares e estão envolvidos em respostas dos linfócitos T auxiliares (*T helper 1* - Th1). Em contraste, macrófagos polarizados em M2 são induzidos por IL-4, IL-13 ou IL-10 e promovem respostas dos linfócitos T auxiliares 2 (*T helper 2* - Th2). Considerando isso, acredita-se que o curso da infeção difere de acordo com a polarização do macrófago em resposta à infeção por *C. burnetii* (Benoit *et al.*, 2008). Se moléculas associadas a M1 forem expressas por macrófagos, a replicação bacteriana será controlada (Ghigo *et al.*, 2002; Benoit *et al.*, 2008). Enquanto que a estimulação de uma resposta M2 será responsável pela persistência de *C. burnetii* em

macrófagos que se tornam altamente permissivos à replicação de *C. burnetii* (Ghigo *et al.*, 2001; Ghigo *et al.*, 2003; Benoit *et al.*, 2008).

Na febre Q, a resposta efetiva do sistema imunológico é caracterizada pela formação de lesões granulomatosas protetoras, ricas em macrófagos em diferentes fases de maturação, incluindo células epitelióides e células gigantes multinucleadas (Honstetter *et al.*, 2004; Shannon e Heinzen, 2009). A formação dos granulomas está sob o controlo do interferão gama IFN γ e é característica da febre Q aguda que se resolve espontaneamente (Raoult *et al.*, 2005; Boucherit *et al.*, 2012). Apenas algumas bactérias isoladas podem ser encontradas em granulomas durante a fase aguda (Honstetter *et al.*, 2004). No entanto, o controlo imunológico de *C. burnetii* pode não levar à sua erradicação do hospedeiro infetado (Honstetter *et al.*, 2004). DNA de *C. burnetii* pode ser encontrado em monócitos circulantes ou na medula óssea de pessoas infetadas meses ou anos antes (Capo *et al.*, 2003). Além disso, *C. burnetii* pode persistir no fígado, osso, válvulas cardíacas e endocárdio (Harris *et al.*, 2000). Há também a hipótese de que o útero pode ser um local de infecção latente, portanto, a reativação durante a gravidez pode ocorrer (Langley *et al.*, 2003; Carcopino *et al.*, 2007). Por todas essas razões, a persistência, recorrência ou reemergência de *C. burnetii* são uma preocupação constante após uma infecção aguda (Sabatier *et al.*, 1997).

As infecções crónicas são devidas a uma imunidade celular defeituosa que é inadequada para erradicar o patógeno. Pacientes com febre Q crónica apresentam aumento da secreção de IL-1 e IL-10 (Capo *et al.*, 1996; Ghigo *et al.*, 2001; Honstetter *et al.*, 2003; Raoult *et al.*, 2005) e supressão da resposta celular mediada por Th1 (Waag, 2007). A *C. burnetii* continua a multiplicar-se apesar das altas concentrações de todas as três classes de anticorpos (IgG, M e A) para bactérias de fase I e II (Sabatier *et al.*, 1997). Biópsias de órgãos não mostram granulomas (Boucherit *et al.*, 2012), mas grandes vacúolos contendo *C. burnetii* detetados em tecidos infetados, como válvulas cardíacas e fígado, e em aneurismas (Maurin e Raoult, 1999). Os monócitos desses pacientes não são capazes de eliminar a *C. burnetii* (Dellacasagrande *et al.*, 2000) e não migram pelo endotélio (Raoult *et al.*, 2005).

Além da resposta mediada por células, a imunidade mediada por anticorpos também parece ser importante na infecção por *C. burnetii* (Shannon e Heinzen, 2009). Imunoglobulinas específicas são segregadas após a infecção (Maurin e Raoult, 1999). No entanto, uma grande heterogeneidade é descrita entre os hospedeiros em relação ao tempo de início da resposta sorológica, bem como à magnitude do pico do título de anticorpo e do seu decaimento. Nos estágios iniciais da infecção, especialmente durante

o aumento da resposta de anticorpos, a concentração altera-se rapidamente num período de tempo relativamente curto (Wielders *et al.*, 2015). Anticorpos específicos para *C. burnetii* de fase II, tanto IgM quanto IgG, podem ser detetados em duas ou três semanas. Os títulos de anticorpos permanecem aumentados até 13 semanas após a infecção (Maurin e Raoult, 1999; Roest *et al.*, 2013a) e o seu declínio leva muito tempo com concentrações decrescentes lentamente (Wielders *et al.*, 2015). O anticorpo *C. burnetii* antifase I aumenta cerca de quatro semanas mais tarde quando comparado com a fase II (Roest *et al.*, 2013a). Assim, dentro de três semanas após a infecção, apenas níveis baixos de IgM antifase I são detetados (Dupuis *et al.*, 1985). Anticorpos específicos de fase I, nomeadamente IgG e IgA, são predominantes nas formas crônicas da doença (Fournier e Raoult, 1999). Globalmente, pode-se considerar que a fase II de IgG atinge o maior título de pico e que também é o mais persistente (Wielders *et al.*, 2015). A duração da resposta do anticorpo não está exatamente determinada. No entanto, pode-se concluir a partir de estudos de campo que os anticorpos *C. burnetii* são altamente persistentes com duração de vários meses ou até anos (Roest *et al.*, 2013a; Teunis *et al.*, 2013). Assim, tanto a imunidade humoral quanto a celular desempenha um papel na infecção por *C. burnetii* (Maurin e Raoult, 1999).

1.1.9. Prova de princípio sobre a base genética das doenças infecciosas

A genética humana introduziu um novo paradigma na compreensão das doenças infecciosas ao permitir o entendimento das consequências de erros inatos nos genes do sistema imunitário sobre a suscetibilidade e gravidade da doença (Alcaïs *et al.*, 2009). A prova de princípio surgiu com o desvendar do papel da genética do hospedeiro na suscetibilidade e na gravidade das doenças causadas por micobactérias, incluindo na tuberculose e na lepra (Casanova e Abel 2002; Alter *et al.*, 2011; Boisson-Dupuis *et al.*, 2011; Grant *et al.*, 2013; Bustamante *et al.*, 2014; Boisson-Dupuis *et al.*, 2015). Mutações em nove genes da via IL12 - IFN γ , que resultam num compromisso da imunidade mediada pelo IFN γ , foram identificadas em indivíduos com suscetibilidade mendeliana a infecções micobacterianas (Mendelian susceptibility to mycobacterial disease, MSMD [MIM 209950]). Inclusivamente, a utilização do IFN γ em tratamentos que visam o restabelecimento da deficiência inata desta citocina entrou na prática terapêutica (Holland *et al.*, 1994; Holland, 2001). No caso da tuberculose pulmonar em adultos a relação não parece ser tão linear (Grant *et al.*, 2013). De entre os fatores genéticos mais convincentes da suscetibilidade à tuberculose pulmonar está o gene *NRAMP1* (alias SLC11A1) (Vidal *et al.*, 1993; Bellamy *et al.*, 1998; Greenwood *et al.*, 2000; Malik *et al.*, 2005; Li *et al.*, 2006; Abel *et al.*, 2014).

No entanto o papel da genética do hospedeiro continua por desvendar no caso de outras doenças infecciosas inclusivamente nas doenças por microrganismos intracelulares tais como na febre Q (Asner *et al.*, 2014).

1.1.10. Implicações para a escolha de genes candidatos na imunidade mediada pelo IFN γ

O sucesso do modelo genético na tuberculose levou-nos a refletir sobre algumas semelhanças na patofisiologia da infecção por *M. tuberculosis* e *C. burnetii*, na escolha de genes candidatos para estudos de associação genótipo – fenótipo da febre Q. Embora *M. tuberculosis* e *C. burnetii* pertençam a grupos bacterianos distantes, das bactérias ácido-álcool resistentes e gram negativas, respetivamente, a infecção por esses patógenos compartilha várias características fisiopatológicas comuns. Ambos são bactérias intracelulares que preferencialmente infetam macrófagos. A característica maior da infecção é a formação do granuloma protetor, estruturas ricas em macrófagos que atuam no controlo da disseminação bacteriana. O granuloma caseoso clássico é considerado a “hallmark” (marca registrada) da tuberculose (Ramakrishnan, 2012). A formação de granulomas também é característica da febre Q (Eldin *et al.*, 2017).

Existe um viés quanto ao papel protetor do granuloma, que também pode constituir um nicho para o crescimento bacteriano e a persistência dentro do hospedeiro (Flynn *et al.*, 2011; Gideon e Flynn, 2011; Silva Miranda *et al.*, 2012). A polarização do macrófago é o mecanismo pelo qual *M. tuberculosis* e *C. burnetii* são capazes de subverter a seu favor esse viés na resposta imune. Estados polarizados, como os estados de ativação do macrófago clássicos (M1) ou alternativos (M2), estão diretamente ligados a aumentos na aptidão bacteriana para a sobrevivência intracelular, replicação bacteriana ou morte bacteriana (Marino *et al.*, 2015). As espécies de *Mycobacterium* e *Coxiella* regulam diferencialmente a polarização do macrófago para M1 ou M2, moldando assim um fenótipo de doença aguda/proliferativa (M1) ou um fenótipo de doença crónica/persistente (M2) (Thiriot *et al.*, 2020). A resposta inflamatória por fim resulta na formação de um granuloma e na regulação positiva da apresentação de antígenos e subsequentes respostas das células NK e T (Doherty e Anderson, 2003). Em última análise, os macrófagos do tipo M1/M2 direcionam as células T para atividades semelhantes a Th1 ou Th2, respectivamente (Mills e Ley, 2014; Mills, 2015). O papel do IFN γ em todo este processo é determinante. Curiosamente, enquanto que na tuberculose a associação das formas graves à deficiência de IFN γ parece ser um dado adquirido, no caso da febre Q há evidência do contrário, ou seja, um aumento da

produção do IFN γ parece acompanhar as formas graves da febre Q crónica (Schoffelen *et al.*, 2017).

1.1.11. Variantes genéticas e a febre Q

As doenças causadas por bactérias intracelulares e microrganismos fastidiosos estão bem caracterizadas relativamente à contribuição de fatores sociodemográficos e dos seus vetores artrópodes ou animais envolvidos no estabelecimento de infecção. No entanto, com exceção da tuberculose, pouco se sabe sobre o papel dos fatores ligados ao hospedeiro, como as variantes genéticas associadas aos fenótipos clínicos da doença (Asner *et al.*, 2014).

Relativamente à infecção por *C. burnetti*, existem alguns relatos na literatura da existência de variantes nos genes da resposta imunitária associadas com o desenvolvimento de febre Q. Na tabela 1.1. estão indicadas as variantes genéticas com associação significativa à febre Q já identificadas através de estudos baseados em genes candidatos (*candidate genes association studies* - CGAS). Não obstante, a comunidade científica ainda conhece muito pouco sobre a real contribuição das variantes genéticas em genes chaves no estabelecimento de infecção e na febre Q crónica.

Helbig e colaboradores relataram a existência de diferenças imunogenéticas, que suportam o conceito de diferentes estados imunes na febre Q crónica, determinadas por variações genéticas na resposta imune dos hospedeiros, em vez de apenas pelas propriedades da *Coxiella burnetii* (Helbig *et al.*, 2005).

A maior contribuição nesta área tem sido dos grupos holandeses em resposta aos importantes surtos sofridos no país, concluindo que uma resposta imunitária deficitária com base genética poderá contribuir para o desenvolvimento da febre Q crónica (Tabela 1.1). O desenvolvimento da febre Q crónica foi associado a polimorfismos de nucleotídeo único (*Single Nucleotide Polymorphism* - SNPs) em genes que codificam recetores de reconhecimento de padrões (nomeadamente recetores *Toll-like*- TLR), componentes da via fagolisossomal e metaloproteinases de matriz (MMPs) (Buijs *et al.*, 2021). A via IL-12 - IFN γ parece ter, igualmente, um papel fulcral na resolução da febre Q (Schoffelen *et al.*, 2017).

O IFN γ tem um papel importante na defesa contra *C. burnetii* uma vez que está envolvido na resposta pró-inflamatória com formação de granuloma e morte ou controlo da bactéria dentro de monócitos/macrófagos ativados (Schoffelen *et al.*, 2017). Embora

alguns estudos defendam que pacientes com febre Q crónica têm uma resposta inadequada do IFN γ à *C. burnetii*, levando à infecção persistente (Dellacasagrande *et al.*, 1999., Andoh *et al.*, 2007; Capo e Mege, 2012), faltam evidências científicas que apoiem essa hipótese. Pelo contrário, estudos mais recentes mostraram a elevada produção de IFN γ antigénio-específico na febre Q crónica (Limonard *et al.*, 2012; Schoffelen *et al.* 2013; Schoffelen *et al.*, 2017). A produção e resposta ao IFN γ estão intactas em pacientes com febre Q crónica, sendo mais elevadas do que em indivíduos saudáveis, o que não explica a falha na eliminação da infecção. Os polimorfismos no gene IL-12p40 estão associados a febre Q crónica o que indica um papel da via do IFN γ , embora as suas consequências continuem obscuras (Schoffelen *et al.*, 2017).

Na via IL-12 - IFN γ , SNPs associados à resposta imunitária à infecção por microrganismos intracelulares, foram detetados, essencialmente, nos genes dos recetores de IFN γ (*INFGR1* e *INFGR2*) (Hijikata *et al.*, 2012; Bustamente *et al.*, 2014; He *et al.*, 2017; Schoffelen *et al.*, 2017).

Partindo da análise da literatura da Tabela 1.1.e do conhecimento dos genes que terão um papel importante da fisiopatologia das infeções intracelulares, decidimos efetuar um estudo exploratório para identificar variantes associadas ao desenvolvimento de febre Q crónica nos genes: *TLR1* e *TLR2* (genes dos recetores *Toll-like* 1 e 2) implicados no reconhecimento de padrões, fagocitose e resposta de sinalização (Everett *et al.*, 2007; Ammerdorffer *et al.*, 2016); *IL12RB1* (gene do recetor β 1 da interleucina-12), *INFGR1* (gene do recetor 1 do IFN γ) e *INFGR2* (gene do recetor 2 do IFN γ), componentes da via IL-12 – IFN γ , implicados na formação do granuloma (Hijikata *et al.*, 2012; He *et al.*, 2017; Schoffelen *et al.*, 2017); *NRAMP1* (gene da resistência natural associada à proteína macrófágica), implicado na via fagolisossomal (Helbig *et al.*, 2003; Archer *et al.*, 2015); e *TIMP1* (gene que codifica o inibidor tecidual 1 de metaloproteinases) implicado na remodelação e reparação do tecido após a destruição por metaloproteinases de matriz (MMPs) (Chen *et al.*, 2017). Numa segunda fase, efetuamos, ainda, um estudo réplica para confirmar a associação de variantes previamente identificadas no gene *IFNG* (gene que codifica o IFN γ) (Jansen *et al.*, 2019) e no gene da via fagolisossomal, *P2RX7* (gene do recetor purinérgico P2X7) (Wielders *et al.*, 2015).

Tabela 1-1. SNPs identificados pelo método CGAs como estando associados a febre Q na população holandesa

Referência bibliográfica	Gene	SNP ID	Nomes HGVS	Consequência mais severa	Nº de casos	Nº de controlos	Modelo Estatístico	Tipo de Hereditariedade	Valor de p	OR	Crítérios de seleção
Jansen et al., 2017	MMP7	rs11568818	NC_000011.10:g.102530930T>C; NC_000011.10:g.102530930T>A	Variante de região reguladora	139	220	Teste de Wilcoxon	—	< 0,05	1,63	Controlos: infeção com febre Q no passado sem progressão para infeção crónica; Risco cardiovascular. Casos: febre Q crónica comprovada ou provável.
	MMP9	rs17576	NC_000020.11:g.46011586A>G	Variante missence	139	220	Teste de Wilcoxon	—	< 0,05	1,67	
Jansen et al., 2019	ATG5	rs2245214	NC_000006.12:g.106214866C>G	Variante intrónica	173	184	Regressão logística univariada	Modelo Recessivo	0,04	0,5	Controlos com fatores de risco cardiovascular e exposição à C. burnetii. Casos com febre Q crónica confirmada ou provável.
	MAP1LC3A	rs1040747	NC_000020.11:g.34558028C>G; NC_000020.11:g.34558028C>A	Variante intrónica	173	184	Regressão logística univariada	Modelo Recessivo	0,01	0,6	
	P2RX7	rs1718119	NC_000012.12:g.121177300G>A; NC_000012.12:g.121177300G>T	Variante missence	173	184	Regressão logística univariada	Modelo Recessivo	0,02	0,6	
		rs3751143	NC_000012.12:g.121184501A>C; NC_000012.12:g.121184501A>G	Variante missence	173	184	Regressão logística univariada	Modelo recessivo	0,03	4,10	
	RAB5A	rs8682	NC_000003.12:g.19984917C>G	Variante UTR 3'	173	184	Regressão logística univariada	Modelo recessivo	< 0,001	0,00	
	RAB7A	rs13081864	NC_000003.12:g.128771430C>T	Variante intrónica	173	184	Regressão logística univariada	Modelo dominante	0,04	1,60	
	TNF	rs1800610	NC_000006.12:g.31576050G>A	Variante intrónica	173	184	Análise de qui-quadrado	Sem modelo	0,03	—	
Schoffelen et al., 2015	MYD88	rs4988453	NC_000003.12:g.38137763C>A; NC_000003.12:g.38137763C>T	Local de ligação de fator de transcrição	92	220	Regressão logística	Modelo Dominante	0,09	2,00	Controlos com fatores de risco cardiovascular e exposição à C. burnetii. Casos com febre Q confirmada (92). Controlos com fatores de risco cardiovascular e exposição à C. burnetii. Casos com febre Q confirmada (92) ou suspeita (47)
					139	220	Regressão logística	Modelo Dominante	0,04	2,15	
	TLR 1	rs5743611	NC_000004.12:g.38798593C>G	Variante missence	92	220	Regressão logística	Modelo Dominante	0,03	0,48	
Wielders et al., 2015	IFNG	rs1861493	NC_000012.12:g.68157416G>A	Variante intrónica	85	1008	Teste de qui-quadrado	—	0,006	3,53	Controlos: comunidade saudável de zona endémica. Casos com febre Q aguda.
		rs2430561	NC_000012.12:g.68158742T>A	Variante intrónica	85	1008	Teste de qui-quadrado	—	0,027	2,13	

1.2. Estudos de genes candidatos (*Candidate gene association studies* – CGAS)

A acessibilidade a dados de sequenciação do genoma humano promoveu a epidemiologia genética para um domínio de interface chave nas ciências da saúde. Integrado dentro do escopo da epidemiologia genética, a associação genótipo- fenótipo, por meio de abordagens de genoma completo (*whole genome association studies* – WGAS) ou gene candidato (*candidate gene association studies* – CGAS) contribui para o amplo processo de compreensão da arquitetura genética das doenças (David, 2021).

Os estudos populacionais de associação genótipo-fenótipo com base em genes candidatos (CGAS) conheceram um recente incremento resultante das abordagens baseadas na sequenciação de nova geração (*Next generation sequencing* – NGS), que se têm revelado eficazes na determinação das sequências de DNA dos genes nas regiões codificantes e regiões não codificantes incluindo regiões flangeadoras, locais de *splicing* e locais reguladores. Este avanço tecnológico foi determinante para a genética humana no cruzamento da investigação em genética clínica com as abordagens baseadas em estudos populacionais (David, 2020).

Considerando que WGAS é uma estratégia de associação livre de hipóteses *a priori* e não direcionada, que estuda a totalidade do genoma, a abordagem CGAS do gene candidato é lógica na medida em que os genes são selecionados com base numa hipótese *a priori* sobre a sua implicação causal na doença (MacArthur, 2014). Estes estudos envolvem a sequenciação dos genes candidatos em grupos de indivíduos com fenótipos clínicos cuidadosamente definidos, e a análise funcional subsequente de variantes identificadas como estando em associação.

1.2.1. Etapas de CGAS

Existem várias etapas críticas para o CGAS (Figura 1.4.), nomeadamente a elaboração do protocolo científico, o projeto de estudo, a definição de fenótipo, a seleção de genes candidatos, a seleção de variantes candidatas, a análise de haplótipos e de desequilíbrio de ligação, a estimativa do tamanho da amostra, a significância estatística e o valor-p, o controlo de qualidade dos dados de genotipagem, a análise exploratória dos dados, os testes de associação estatística e a subsequente análise funcional (David, 2021).



Figura 1.4. Etapas dos estudos populacionais de associação genótipo-fenótipo com base em genes candidatos.

Fonte: David, 2020

Capítulo II – Objetivos

O objetivo principal do presente estudo é a identificação na população portuguesa de determinantes genéticos da infeção na febre Q crónica por métodos de associação genótipo-fenótipo com base em genes candidatos (*Candidate gene Association Studies* – CGA). A seleção de genes candidatos para estes estudos foi realizada com base no conhecimento da resposta imunológica à infeção e em estudos anteriores acima descritos.

Com vista a alcançar o objetivo principal, foi utilizada uma abordagem retrospectiva caso-controlo para verificar a frequência com que ocorrem certas variantes genética nos diferentes grupos. Foram efetuadas duas abordagens distintas: um estudo exploratório para identificar novas variantes associadas a fenótipos clínicos da doença e um estudo réplica para confirmar a associação de variantes previamente identificados. Deste modo, foram postulados objetivos específicos para cada estudo.

Objetivos do estudo exploratório:

- Genotipar os genes *IL12RB1*, *IFNGR1*, *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* e *TIMP1* por *Next-Generation Sequencing*- NGS.
- Identificar, nos genes analisados, variantes genéticas associadas ao desenvolvimento de febre Q crónica em amostras provenientes da população portuguesa infetada por *Coxiella burnetii*.

Objetivos do estudo réplica:

- Genotipar das variantes rs1718119 e rs3751143 do gene *P2X7R* e as variantes rs2430561 e rs1861493 do gene *INFG* através da sequenciação pelo método de Sanger.
- Efetuar a análise de associação destas variantes com a febre Q crónica na população portuguesa.

Capítulo III - Materiais e Métodos

3.1. Declaração de ética

Os procedimentos do presente estudo foram autorizados pela Comissão de Ética para a Saúde do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA) e realizado de acordo com a Declaração de Helsínquia de 1975. Os participantes forneceram o seu consentimento informado por escrito para que as respetivas amostras pudessem ser utilizadas na investigação epidemiológica da febre Q. Nos casos em que não foi possível recolher esse documento as amostras foram utilizadas à guarda do nº6 do Artigo 19º da Lei nº 12/2005, de 26 de janeiro referente à informação genética pessoal e informação de saúde: *“No caso de uso retrospectivo de amostras ou em situações especiais em que o consentimento das pessoas envolvidas não possa ser obtido devido à quantidade de dados ou de sujeitos, à sua idade ou outra razão comparável, o material e os dados podem ser processados, mas apenas para fins de investigação científica ou obtenção de dados epidemiológicos ou estatísticos”*. A codificação dos participantes foi respeitada de acordo com as normas homologadas para a investigação no INSA.

3.2. Caracterização da população estudada

Para os estudos de associação genética com a febre Q crónica, foram analisados 60 casos de febre Q em pacientes portugueses diagnosticados, entre 2005 e 2018 (exceto um caso de 1995).

A febre Q crónica foi diagnosticada serologicamente pela presença de anticorpos anti-fase I com título de $IgG \geq 800$ (Bossi *et al.*, 2004) num total de 17 casos com uma média de idades de $52,0 \pm 19,7$ anos e um rácio homem/mulher de 3:1 (13:4). A endocardite foi a principal manifestação clínica de febre Q crónica em 65% (11/17) dos pacientes, a doença febril recorrente foi também observada em 24% (4/17) e a osteomielite apenas foi observada em duas crianças com idades inferiores a 10 anos, o equivalente a 12% (2/17). Neste grupo, 76% (13/17) dos pacientes apresentavam fatores de predisposição à febre Q crónica, enquanto que, 18% (3/17) não possuía nenhuma condição anterior. As condições pré-existentes incluíam lesão valvular (13/17), febre Q aguda prévia (1/17), doença neoplásica (1/17) e doença renal (2/17).

A febre Q aguda foi diagnosticada pela presença de anticorpos anti-fase II com $IgM \geq 50$ e $IgG \geq 200$ (Bossi *et al.*, 2004) num grupo de 43 indivíduos com uma média de $44,5 \pm 18,3$ anos e com uma proporção entre homens e mulheres de 4:1 (34:9). As manifestações clínicas incluíam doença febril em 77% (33/43), hepatite em 26% (11/43),

doença cardíaca em 9% (4/43) e SDRA em 5%(2/43). O conhecimento de condições pré-existentes para febre Q crónica não estava disponível em 81% (35/43) dos pacientes. Os restantes 19% (8/43) apresentavam quadros clínicos de hepatite, imunodepressão (gravidez e Infeção por HIV) (3/43), insuficiência renal (2/43) e lesão valvular (3/43).

A título comparativo, os estudos CGA também utilizam dados das populações gerais sem critério de exposição que geralmente se encontram na base de dados dos 1000 genomas (The 1000 Genomes Project Consortium, 2015). Como a população portuguesa não se encontra representada (apenas estão representados 107 indivíduos ibéricos de origem espanhola) optámos, por também genotipar uma amostra de conveniência da população portuguesa. Para análise da genética populacional da população geral portuguesa, foram obtidas 34 amostras de conveniência provenientes de uma coleção anónima do Departamento de Genética Humana do Instituto Nacional de Saúde Doutor Ricardo Jorge. Para este grupo de indivíduos não foi possível obter informação sobre a sua exposição à *C. burnetii*. No intuito de evitar possíveis viés na análise da população geral, foi solicitado aos participantes o preenchimento de um questionário de autoavaliação sobre o eventual conhecimento de algum familiar próximo que teria morrido de infeção severa.

3.3. Amostras biológicas

De cada individuo foram colhidos 3 mL de sangue periférico em tubos de EDTA. A extração de DNA genómico foi realizada pelo método automatizado do sistema *MagNA Pure LC* (Roche®). As amostras de sangue anonimizadas foram armazenadas a -20°C e o DNA extraído foi armazenado a 4°C.

3.4. Métodos

A metodologia experimental utilizada para identificar variantes genéticas associadas à Febre Q crónica utilizando estudos populacionais de associação com base em Genes Candidatos (CGAS) foi planeada para permitir dois tipos de abordagem: 1) Estudo exploratório; 2) Estudo réplica.

3.4.1. Estudo exploratório

Um estudo exploratório foi realizado sobre genes candidatos das vias implicadas na resposta imunológica à infeção acima descritos incluindo genes para os quais não se encontrou na literatura estudos de associação com a febre Q, ou se identificou

estudos com pouca expressão (baixo nível de significância estatística, sem replicabilidade, realizados com número reduzido de amostras, etc.), que apontassem para a existência de variantes associadas à febre Q crónica. Assim, este estudo exploratório consistiu na genotipagem dos genes *IL12RB1*, *IFNGR1*, *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* e *TIMP1*. Dado o seu carácter exploratório o estudo foi realizado com base na utilização de uma metodologia de *Next-Generation Sequencing*- NGS (Sequenciação de Nova Geração) pelo facto da NGS permitir a sequenciação nos genes alvo de vastas regiões genómicas codificantes (exões) e não codificantes (regiões flangeadoras, locais de splice na junção exão-intrão e locais reguladores a montante do codão iniciador incluindo na região 5'UTR podendo conter elementos regulatórios essenciais).

3.4.1.1. Preparação das amostras

Os fragmentos a genotipar foram desenhados para incluir os exões dos genes candidatos, regiões intrónicas próximas do local de *splicing* e regiões não codificantes a 2000 pb a montante e na região 5'UTR. Foi utilizado, como referência genómica, o principal transcrito da versão GRCh38.p12 (*Genome Reference Consortium Human Build 38*) disponível no *Ensemble Genome Browser 96* (*Ensembl*, 2019) (<https://www.ensembl.org/index.html>). A Tabela 3.1. apresenta a informação detalhada dos fragmentos em análise.

3.4.1.1.1. Desenho de primers

Os *primers* foram desenhados com recurso à ferramenta informática *Primer-BLAST*, disponibilizada pelo *National Center for Biotechnology Information* (NCBI) em <https://www.ncbi.nlm.nih.gov/tools/primer-blast/> e adquiridos através da Thermo Fisher Scientific®. A Tabela 3.2. apresenta a informação detalhada dos *primers* escolhidos para amplificar os fragmentos em análise.

Tabela 3. 1. Fragmentos a amplificar dos genes *IL12RB1*, *IFNGR1*, *IFNGR2*; *NRAMP1*; *TLR1*, *TLR2* e *TIMP1*

Gene				Fragmento		
Nome	Localização Genômica	Orientação	Coordenadas	Nome	Tamanho (pb)	Coordenadas
<i>TLR1</i>	4p14	Reverse	38790677-38856817	<i>T1A1</i>	12110	38793975-38806088
<i>TLR2</i>	4q31.3	Foward	153684070-153705702	<i>T2A1</i>	6185	153700772-153706956
<i>IL12RB1</i>	19p13.11	Reverse	18058995-18098944	<i>LA</i>	10760	18089867-18100631
				<i>LB</i>	13559	18075127-18088684
				<i>LC1</i>	9748	18065426-18075173
				<i>LC2</i>	7576	18057818-18065393
<i>IFNGR1</i>	6q23.3	Reverse	137197484-137219449	<i>G1A</i>	6069	137214674-137220742
				<i>G1B4</i>	1415	137205750-137207164
				<i>G1B5</i>	2626	137203404-137206039
				<i>G1B7</i>	3206	137200077-137203282
				<i>G1B8</i>	3330	137196715-137200044
<i>IFNGR2</i>	21q22.11	Forward	33402896-33479348	<i>G2A3</i>	3790	33401040-33404829
				<i>G2B</i>	5587	33410124-33415711
				<i>G2C</i>	1821	33420534-33422353
				<i>G2D</i>	1793	33425848-33427642
				<i>G2E</i>	8034	33431071-33439104
				<i>G2F</i>	1636	33478213-33479848
<i>NRAMP1</i>	2q35	Forward	218382029-218396894	<i>RA</i>	8823	218380796-218389008
				<i>RB1</i>	4851	218388989-218393840
				<i>RB2</i>	5217	218393283-218398499
<i>TIMP1</i>	Xp11.3	Forward	47582408-47586789	<i>TIMP1</i>	6726	47581095-47587820

Tabela 3. 2. Primers escolhidos para amplificar os fragmentos dos genes *IL12RB1*, *IFNGR1*, *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* e *TIMP1*

Genes	Fragmentos	Primers foward	Primers reverse
<i>TLR1</i>	T1A1	5'-ACTGGGAGTAGTGGGGATGAG-3'	5'-AAGAAAAGAGACTGGGTTGTCC-3'
<i>TLR2</i>	T2A2	5'-GTGTTGCTGGTATAGGGATAGACA3'	5'-TACTCTTGCCCTCAACACCC3'
<i>IL12RB1</i>	LA	5'-AATAGCACTGACCACTCGGG3'	5'-CCAGCAGGCTATGAGGTCAG3'
	LB	5'-TTTAGAGGCTTTGCCGACAGT3'	5'-CAAGACCGGATGCCACATTG3'
	LC1	5'-AGGGAGAAATGGGTGCTCAA3'	5'-GGTGATCGAGACCTCAAGT3'
	LC2	5'-CCAGCCAAGTGCTTATGGTG3'	5'-CTCCTACTTCTCGCACCCAA3'
<i>IFNGR1</i>	G1A	5'-TGCCTACTCCATAGACGCC3'	5'-AGTTGGGTTTGGTCTGGCAT3'
	G1B4	5'-GGGCAATGTGGCATCTTACAA3'	5'-GGGCTTGAAGAAGGGGAGGA3'
	G1B5	5'-ACAATGCTTAAGAAACACTTGGAGG3'	5'-ACTGCAAATGAGTTTGCCTTCT3'
	G1B7	5'-CCCCTTGACCCAGGAAAGAAA3'	5'-CTAGCGCGTGTAATGGCA3'
	G1B8	5'-ACCAGCAGTAGCATAACGGG3'	5'-TGCTTTTCCCATTCTAATGTGTGC3'
<i>IFNGR2</i>	G2A3	5'-TCTAACTGTATGCCCGTGTCT3'	5'-GTTTGGCTCATAAGCGTGGT3'
	G2B	5'-AGGCAAGAAGGTTTCAGCAAGT3'	5'-CCCATCTCACAAGCACTGGAC3'
	G2C	5'-TGAAAGAACTCCGCACCA3'	5'-GTGTGCATAAACCAGACACGG3'
	G2D	5'-GAGAATCCAGCCCTGCTGTT3'	5'-CCTATACTCCCCGAGGCTGA3'
	G2E	5'-GCAGAGGAGACACAGAACCTT3'	5'-GGGAGTGGGCCTGTTTTTC3'
	G2F	5'-CCCAGTACTTTCAGAGACCGAG3'	5'-TGAGGAAGCTGGGTGAGTCC3'
<i>NRAMP1</i>	RA	5'-TCAAGCATGGTGGCATAGGTC3'	5'-GACACGCTCCTTCACTCACC3'
	RB1	5'-TGCACACCTGTAGTCCTAGTT3'	5'-TAATGAGCCTGGGCAACAACATAA-3'
	RB2	5'-GCAGGAGGGTAGGAATCACC3'	5'-GGGACTGTGGGGACTGAT-3'
<i>TIMP1</i>	TIMP1	5'-GCTGGATACGCAGTAAGCAC3'	5'-CTAGTTCCAAGGGCCTCACT-3'

3.4.1.1.2. Amplificação por PCR

A reação em cadeia da polimerase (*Polymerase Chain Reaction - PCR*) (Saiki *et al.*, 1985; Mullis, 1987) é a técnica de amplificação melhor conhecida e a mais largamente aplicada, quer na investigação quer rotineiramente em laboratórios clínicos (Wittwer e Kusukawa, 2006). A PCR permite a amplificação *in vitro* de uma sequência específica de DNA.

A estratégia de amplificação utilizada neste trabalho envolveu o recurso a pares de *primers* desenhados para cada fragmento e a uma Taq DNA polimerase de alta fidelidade e de longo alcance, a *TaKaRa LA Taq® Hot Start* (Alfagene, Carcavelos, Portugal) conforme o descrito por Jia e colaboradores (Jia *et al.*, 2014).

O termociclador *T Gradient Biometra®* foi usado para executar as amplificações por PCR. Efetuaram-se ensaios sucessivos de otimização de PCR tendo em vista a obtenção dos melhores resultados. Foi utilizado um aditivo de PCR, a betaína, por este composto aumentar a amplificação de sequências ricas em GC devido à sua capacidade

de dissolver a estrutura secundária que bloqueia a ação da polimerase. Foram testadas diferentes concentrações de magnésio, betaína, temperaturas de hibridação e número de ciclos. Deste modo, foram escolhidas misturas de reação (Tabela 3.3.) e condições de PCR específicas para a amplificação de cada fragmento (Tabela 3.4.).

O programa PCR iniciou-se com um passo de desnaturação durante 1 minuto a 94°C, seguido do passo de desnaturação a 98°C durante 10 segundos. A etapa de hibridação decorreu durante 20 minutos com a temperatura de hibridação ótima utilizada a variar segundo o fragmento a amplificar (65°C, 66°C, 67°C, 67,6°C, 68°C ou 69°C). O número de ciclos do PCR variou entre 30, 35 ou 40 ciclos, segundo o fragmento. A extensão decorreu a 72°C durante 10 minutos.

Tabela 3. 3. Misturas de reação para os fragmentos a amplificar dos genes *IL12RB1*, *IFNGR1*, *IFNGR2*; *NRAMPI*; *TLR1*, *TLR2* e *TIMP1*.

Mistura reacional comum a todos os fragmentos¹			
Reagente	Concentração da solução de trabalho	Volume por reação (µL)	Concentração final
<i>TaKaRa LA Buffer II</i> (com MgCl ₂ ²)	10x	2,5	1x
dNTPs	2,5 mM cada	4	0,4mM cada
Betaina	5M	5	1M
<i>Primer forward</i>	25 µM	2,5	2,5 µM
<i>Primer reverse</i>	25 µM	2,5	2,5 µM
DNA	30 -100 ng/µL	1	1,2- 4 ng/µL
Mistura reacional fragmentos LA, LB e T1A1			
Reagente	Concentração da solução de trabalho	Volume por reação (µL)	Concentração final
<i>TaKaRa LA Taq HS</i>	5 U/µL	0,25	1,25 U/ reação
H ₂ O	--	7,25	--
Mistura reacional fragmentos G2C, G1B4, LC1 e G2A3, T2A1 e TIMP1			
Reagente	Concentração da solução de trabalho	Volume por reação (µL)	Concentração final
<i>TaKaRa LA Taq HS</i>	5 U/µL	0,125	0,625 U/ reação
H ₂ O	--	7,375	--
Mistura reacional fragmentos G1B8, G1B7, G2E, G1B5, G2B, G2D, G1A, LC2 G2F, RA, RB1 e RB2			
Reagente	Concentração da solução de trabalho	Volume por reação (µL)	Concentração final
<i>TaKaRa LA Taq HS</i>	5 U/µL	0,125	0,625 U/ reação
H ₂ O	--	6,375	--
MgCl ₂ ³	25 mM	1	1mM

¹ Quanto maior o fragmento a amplificar maior a quantidade de *Taq DNA* polimerase necessária por reação

² O *TaKaRa LA* Buffer II contém $MgCl_2$ numa concentração perfazendo na reação final 2mM de Mg^{++} .

³ Com a adição de mais 1 μ L de uma solução de $MgCl_2$ 25 mM obtemos uma concentração final de 3 mM

Tabela 3. 4. Condições de PCR adequadas para amplificação dos fragmentos pela *TaKaRa LA Taq Hot Start*.

Fragmentos	Etapas do PCR	Temperatura	Tempo	Nº de Ciclos
	Desnaturação Inicial	94°C	1 min.	
	Desnaturação	98°C	10 seg.	
G1B8, G1B7, G2E		65°C		30
G1B5		66°C		30
G2B, G2D, G2C, G1B4, T2A1 e TIMP1		67°C		30
LC1, LA, LB e T1A1	Hibridação	67°C	20 min	40
G2A3		67,6°C		35
RB2		68°C		30
G1A, LC2, RA e RB1		68°C		35
G2F		69,7°C		30
	Extensão	72°C	10 min	

3.4.1.1.3. Eletroforese em gel de agarose

A eletroforese em gel de agarose é um método simples e altamente eficaz para a separação, identificação e purificação de fragmentos de DNA entre os 0,5 e os 25 kb (Voytas, 2001).

A verificação da amplificação dos fragmentos foi efetuada através da realização de eletroforese em gel de agarose (Seakem® *LE Agarose*) a 1% em TBE (tris-borato EDTA) 1x suplementado com brometo de etídio (10 mg/mL, Sigma®). Os produtos de PCR foram carregados juntamente com um tampão de deposição. Utilizou-se como marcador de peso molecular padrão o *1kb DNA Ladder* da New England BioLabs® (Werfen Portugal, Carnaxide, Portugal) (Anexo 2). A corrida de eletroforese decorreu a uma voltagem constante de 80 V durante 90 minutos gerada pela fonte de tensão *BIORAD® PowerPac300*. A visualização do resultado foi feita através de um dispositivo de emissão de raios-ultravioleta acoplado a um sistema de captação de imagem. O tamanho esperado das bandas foi verificado por comparação com as bandas do marcador de peso molecular padrão.

3.4.1.1.4. Quantificação visual dos produtos de PCR

Após o PCR dos diferentes fragmentos, foram obtidas diferentes concentrações de amplicões, o que se verificou pela intensidade das bandas no gel de agarose. O protocolo *Nextera XT* requer uma padronização através da quantificação objetiva das concentrações dos produtos PCR, porém esta estratégia iria esgotar o produto PCR para as próximas fases.

Uma vez que o método de amplificação foi constante (incluindo volume e concentração de reagentes da mistura de reação e volume de amostra depositada no gel de agarose), considera-se que as intensidades das bandas no gel de agarose refletem as concentrações dos amplicões. Deste modo, foi desenvolvida uma estratégia de quantificação visual. A intensidade das bandas no gel foi classificada e anotada entre 1x, 3x, 5x e 10x concentrado (sendo 1 a classificação mais baixa, e 10 a classificação mais elevada, logo a mais concentrada). Através desta classificação, foi adicionado produto de PCR de modo a padronizar as concentrações para a etapa *Pooling* em que os amplicões referentes a uma mesma amostra são agrupados num mesmo tubo.

3.4.1.1.5. *Pooling* das amostras

Perante uma grande quantidade de fragmentos (20 fragmentos amplificados por amostra) de tamanhos distintos, decidiu-se, para poupança de material e de tempo, realizar para cada amostra a verificação da padronização das concentrações por grupos de 4 fragmentos de tamanhos distintos. Assim, para cada amostra, foram constituídos 5 grupos de 4 fragmentos e verificada visualmente a padronização das concentrações dos 20 fragmentos amplificados em apenas 5 corridas de eletroforese (em vez de 20). A eletroforese em gel de agarose decorreu sob as mesmas condições apresentadas anteriormente. Após a verificação visual da homogeneidade das concentrações dos amplicões, foi então efetuada um *pooling* final dos produtos de PCR de cada amostra, isto é, todos os 20 fragmentos pertencentes a uma mesma amostra foram agrupados para que pudessem ser sujeitos às próximas fases do procedimento.

3.4.1.1.6. Purificação dos produtos de PCR

Das amplificações por PCR resultam excedentes como restos de *primers* e reagentes (dNTPs, sais e enzimas) que podem afetar o *base calling*. Para evitar erros de leitura na NGS, os produtos de PCR foram purificados com recurso às esferas magnéticas *AMPure XP* da Beckman Coulter® (*Thermo Fisher Scientific*, Porto Salvo, Portugal) (Anexo 2).

A cada produto de PCR foi adicionado 1,8 µL da solução com esferas magnéticas por cada 1 µL de amostra de DNA, com incubação de 5 minutos à temperatura ambiente (para que as esferas se ligassem ao DNA). A placa com os produtos PCR foi colocada num suporte magnético para permitir a separação magnética das esferas ligadas ao DNA. O sobrenadante foi removido e descartado com o cuidado de não perturbar as esferas. As esferas foram, ainda, lavadas duas vezes com etanol a 70%. No passo final, o DNA purificado foi eluído em 40 µL de tampão de eluição para desagregar o DNA das esferas magnéticas (Beckman Coulter, 2002)

3.4.1.2. Sequenciação NGS

A metodologia de sequenciação de nova geração (*Next Generation Sequencing* - NGS) é tripartida e em cada uma das fases existem etapas específicas de controlo de qualidade. Uma vez que o NGS trabalha com milhares de milhões de dados em simultâneo, existe uma elevada probabilidade que uma base seja lida de forma incorreta. Para ultrapassar esta limitação e garantir resultados fidedignos, existem parâmetros e medidas de qualidade que permitem validar as fases I (*base call*) e II (*variant call*) (Gargis *et al.*, 2016). Para além destas são ainda aplicadas outras medidas de qualidade específicas para CGAS (fase III) que têm por objetivo diminuir a obtenção de erros por falsos positivos na análise de associação (Anderson *et al.*, 2010; David, 2021).

A sequenciação das amostras deste estudo, recorrendo à tecnologia NGS, incluindo a preparação de bibliotecas, *base call* e *variant call*, foi efetuada pela Unidade de Tecnologia e Inovação (UTI) do INSA. Para a sequenciação foi usado o equipamento *MiSeq™* da *Illumina®* (Instrumentos de Laboratório e Científicos – I.L.C., Lisboa Portugal).

3.4.1.2.1. Preparação das bibliotecas NGS

A preparação das bibliotecas foi realizada com o kit *Nextera XT* combinado com o kit indexação *Nextera XT Index v2* da *Illumina®* (Instrumentos de Laboratório e Científicos – I.L.C., Lisboa Portugal), seguindo as instruções do fabricante. Todo o processo foi realizado numa placa de 96 poços com disposição conforme a Figura 3. 1.. Este procedimento inclui as fases: tagmentação, amplificação, limpeza, normalização, diluição e desnaturação de bibliotecas.

3.4.1.2.1.1. Tagmentação

O produto de PCR foi aleatoriamente fragmentado pela enzima *Nextera Transposon* e foram adicionados adaptadores nas extremidades dos fragmentos originados. Os adaptadores servem de ligação entre as bibliotecas geradas e a *flow cell* (superfície de ancoragem onde ocorre a sequenciação). As amostras foram quantificadas e purificadas pela plataforma *Qubit® da Thermo Fisher Scientific Inc* (Thermo Fisher Scientific, Porto Salvo, Portugal), e diluídas para 0,2 ng/μl.

3.4.1.2.1.2. Amplificação de bibliotecas

Para a formação de *Clusters*, sequenciação e identificação da amostra, foram adicionados aos adaptadores, 2 adaptadores *index*, um para distinguir as diferentes bibliotecas (i5) e outro para diferenciar todas as amostras (i7) (Figura 3.1). A organização dos índices foi efetuada com o software *Illumina Experiment Manager* e segundo configuração da Figura 3.1. Foi executado o programa de PCR *Nextera* para amplificação, conforme o protocolo *Nextera XT DNA Library Prep* (Instrumentos de Laboratório e Científicos – I.L.C., Lisboa Portugal) (Illumina, 2016).

3.4.1.2.1.3. Limpeza de bibliotecas

Para purificar as bibliotecas amplificadas, foram utilizadas, novamente, as esferas magnéticas *AMPure XP*, utilizando a metodologia de separação magnética. As esferas purificaram a biblioteca de DNA, além de removerem fragmentos muito pequenos. Foi repetido o procedimento de purificação segundo o protocolo *Instructions For Use AMPure XP* (Beckman Coulter, 2002).

3.4.1.2.1.4. Normalização das bibliotecas

Os volumes de cada biblioteca foram normalizados, garantindo uma representação equitativa das bibliotecas quando agrupadas.

3.4.1.2.1.5. Diluição e desnaturação de bibliotecas

Após o processo de normalização, foi feito um pool das bibliotecas, com 10 μL cada. A execução seguiu o protocolo *MiSeq System - Denature and Dilute Libraries Guide* (Kahl, 2015). Foi utilizado o genoma do fago Phix a 12,5 pM como controlo interno de sequenciação. A biblioteca PhiX fornece um controlo de qualidade para geração de *clusters*, sequenciação, alinhamento e calibração de leituras.

	1	2	3	4	5	6	7	8	9	10	11	12
A	SMD1_C1 S502 N716	SMD9_C10 S502 N716	SMD17_C16 S502 N719	SMD25_A15 S502 N720	SMD33_A27 S502 N721	SMD41_A37 S502 N722	SMD49_A45 S502 N723	SMD57_A56 S502 N724	SMD65_T5 S502 N725	SMD73_T13 S502 N727	SMD81_T21 S502 N728	SMD89_T29 S502 N729
B	SMD2_C2 S503 N716	SMD10_C11 S503 N716	SMD18_A1 S503 N719	SMD26_A16 S503 N720	SMD34_A28 S503 N721	SMD42_A38 S503 N722	SMD50_A46 S503 N723	SMD58_A57 S503 N724	SMD66_T6 S503 N725	SMD74_T14 S503 N727	SMD82_T22 S503 N728	SMD90_T30 S503 N729
C	SMD3_C3 S505 N716	SMD11_C12 S505 N716	SMD19_A2 S505 N719	SMD27_A17 S505 N720	SMD35_A29 S505 N721	SMD43_A39 S505 N722	SMD51_A47 S505 N723	SMD59_A58 S505 N724	SMD67_T7 S505 N725	SMD75_T15 S505 N727	SMD83_T23 S505 N728	SMD91_T31 S505 N729
D	SMD4_C4 S506 N716	SMD12_C13 S506 N716	SMD20_A4 S506 N719	SMD28_A19 S506 N720	SMD36_A30 S506 N721	SMD44_A40 S506 N722	SMD52_A51 S506 N723	SMD60_A59 S506 N724	SMD68_T8 S506 N725	SMD76_T16 S506 N727	SMD84_T24 S506 N728	SMD92_T32 S506 N729
E	SMD5_C5 S507 N716	SMD13_C14 S507 N716	SMD21_A8 S507 N719	SMD29_A20 S507 N720	SMD37_A31 S507 N721	SMD45_A41 S507 N722	SMD53_A52 S507 N723	SMD61_T1 S507 N724	SMD69_T9 S507 N725	SMD77_T17 S507 N727	SMD85_T25 S507 N728	SMD93_T34 S507 N729
F	SMD6_C6 S508 N716	SMD14_C15 S508 N716	SMD22_A10 S508 N719	SMD30_A21 S508 N720	SMD38_A34 S508 N721	SMD46_A42 S508 N722	SMD54_A53 S508 N723	SMD62_T2 S508 N724	SMD70_T10 S508 N725	SMD78_T18 S508 N727	SMD86_T26 S508 N728	SMD94_T44 S508 N729
G	SMD7_C7 S510 N716	SMD15_C16 S510 N716	SMD23_A11 S510 N719	SMD31_A24 S510 N720	SMD39_A35 S510 N721	SMD47_A43 S510 N722	SMD55_A54 S510 N723	SMD63_T3 S510 N724	SMD71_T11 S510 N725	SMD79_T19 S510 N727	SMD87_T27 S510 N728	
H	SMD8_C8 S511 N716	SMD16_C17 S511 N716	SMD24_A14 S511 N719	SMD32_A28 S511 N720	SMD40_A36 S511 N721	SMD48_A44 S511 N722	SMD56_A55 S511 N723	SMD64_T4 S511 N724	SMD72_T12 S511 N725	SMD80_T20 S511 N727	SMD88_T28 S511 N728	

Figura 3. 1. Organização dos *index*, visualização superior da placa de 96 poços.

Legenda: S corresponde ao *index* i7; N corresponde ao i5; C – amostra de paciente com febre Q crônica; A – amostra de paciente com febre Q aguda e T – população geral portuguesa

Fonte: Imagem do sistema informático *Illumina experiment manager*TM

3.4.1.2.2. Fase I

A fase I, ou *base call*, iniciou-se com a transferência do produto de desnaturação das bibliotecas para o cartucho do *MiSeq* seguida do carregamento dos reagentes necessários para o aparelho, de acordo com as instruções do fabricante (*Illumina*®, 2019) (Figura 3.2-Fase I: *MiSeq*). A *flow cell* foi limpa para retirar o excesso de tampão de armazenamento e colocada no compartimento respetivo. Todos os passos seguiram as instruções conforme o guia do sistema (*Illumina*®, 2019). Este passo permitiu a sequenciação massiva por síntese para identificação das sequências de DNA. A sequência dos *index* foi reconhecida pelo software do equipamento *MiSeq*, apenas foram consideradas, como leituras, as a sequência entre os *index* i5 e i7. Foram indicados pelo *MiSeq* os parâmetros de controlo de qualidade: dispersão de leituras, *Cluster* e Q30. A dispersão da percentagem de leitura avalia a homogeneidade da sequenciação. *Cluster* é um conjunto de fragmentos clonados na célula de fluxo (*flow cell*) que resultam da amplificação em ponte dos fragmentos de DNA. Quando existe demasiada sobreposição de luminescência dos fragmentos pode gerar erros de leitura. O clustering é a unidade que representa o total número de *clusters*. Uma pontuação de qualidade de 30 (Q30) representa uma taxa de erro de 1 em 1000 bases identificadas, ou seja, com uma precisão de *base call* correspondente de 99,9%. Quando a qualidade da sequenciação atingir Q30, praticamente todas as leituras serão perfeitas, sem erros ou ambiguidades. É por isso que o Q30 é considerado uma referência de qualidade nno NGS.

3.4.1.2.3. Fase II

Durante a fase II do NGS, ou *variant call* ocorreu o reconhecimento do DNA e identificação das variantes genéticas com recursos a metodologias de bioinformáticas. O processo de *variant call* comporta as seguintes etapas de mapeamento, *variant call*, anotação e filtragem (Figura 3.2-Fase II).

3.4.1.2.3.1. Mapeamento

As sequências desmultiplexadas (sequências separadas por amostra), sem adaptadores ou *index*, foram mapeadas no genoma de referência através do *Bowtie 2 v2.3.4* (Langmead *et al.*, 2019) (Figura 3.2. - Fase II: Bowtie). Foi considerado, como referência, o genoma de humano versão GRCh38.p12, disponível em Ensembl.com (*Ensembl*, 2019). As leituras mapeadas foram agrupadas por coordenadas genômicas pelo software *SAMtools v1.9* (Figura 3.2– Fase II: SAM) (Li *et al.*, 2009). As leituras duplicadas foram identificadas e marcadas pela ferramenta GATK v4.1.5.0 com o módulo MarkDuplicates (Figura 3.2 – Fase II: GATK) (Van der Auwera *et al.*, 2013; Broad Institute, 2020).

3.4.1.2.3.2. Variant call

Os valores de qualidade do *base call* foram corrigidos com o software GATK (ferramentas *BaseRecalibrator* e *ApplyRecalibration*) removendo “0”, entradas sem alelos alternativos e alelos alternativos ambíguos ou repetidos (erros sistemáticos) usando a base de dados *Germline variations* (transferida do *Ensembl Release 96*) (Figura 3.2. - Fase II: GATK). O arquivo de alinhamento resultante foi usado para determinação de variantes. Posições variáveis em comparação com a sequência de referência foram inicialmente chamadas para cada amostra usando a ferramenta GATK *HaplotypeCaller* no formato *Genomic Variant Call Format* (GVCF) (Figura 3.2. - Fase II: GATK).

3.4.1.2.3.3. Anotação

Usando a base de dados *Ensembl Release 96 germline variants*, as variantes conhecidas foram anotadas. Para as variantes com *minimum call confidence* de 20 pelo *HaplotypeCaller*, a genotipagem conjunta foi feita com a ferramenta GATK *Genotype GVCFs* (Figura 3.2 - Fase II: GATK).

3.4.1.2.3.4. Filtragem

Os genótipos das amostras e posições variáveis foram anotados e filtrados usando GATK (módulos *SelectVariants* e *VariantFiltration*), software *BCFtools* v1.9 e software PLINK v1.9 (Li, 2011; Chang *et al.*, 2015). Em primeiro lugar, os genótipos das amostras foram filtrados (Figura 3.2. - Fase II: BCF Tools) para valores de cobertura de profundidade inferiores a 10x e quando foi observada uma elevada perda de heterozigotidade alélica. Em seguida, várias posições foram filtradas sequencialmente quando: a variante se encontrava numa região com mais de 8 unidades repetitivas; a variante SNP estava a 2 bp de um indel, e/ou quando não foi possível genotipar com segurança pelo menos 50% das amostras. Finalmente, apenas SNPs bialélicos foram considerados.

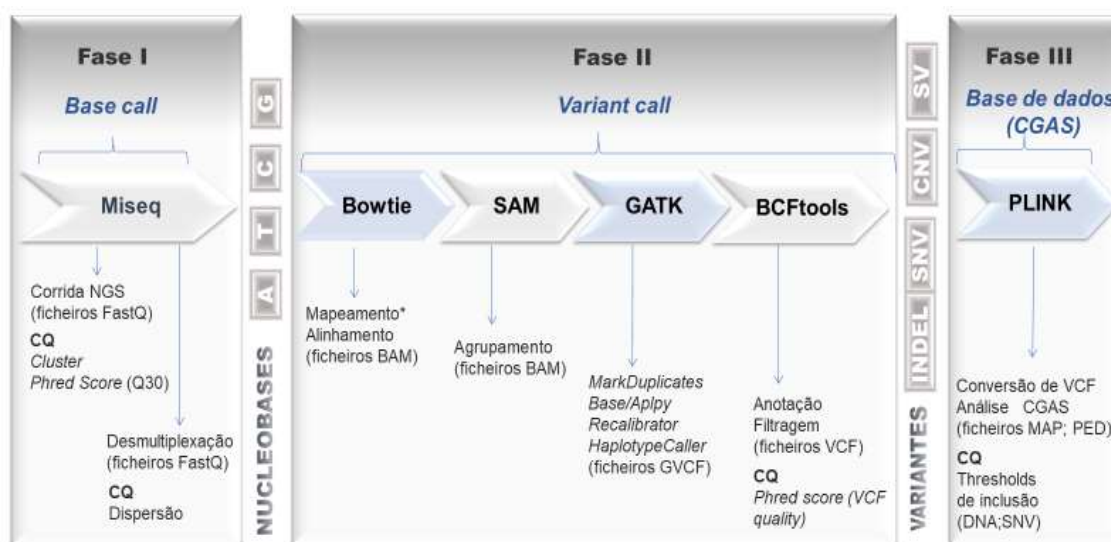


Figura 3. 2. Fluxo da metodologia NGS em fases (I, II, III), indicação das ferramentas bioinformáticas usada em cada passo com os respectivos parâmetros de qualidade.

Legenda: FASTQ- formato baseado em caracteres ASCII; BAM- ficheiro binário de armazenamento de dados; GVCF- *Genomic Variant Call Format*; VCF- *Variant Call Format*; MAP- mapa dos SNPs; PED- *Pedigree Data*; CQ-Controlo de Qualidade. *no genoma de referência

3.4.1.2.4. Fase III

A Fase III do NGS consistiu na análise exploratória dos dados, aplicando os critérios e ajustes específicos do controlo de qualidade para estudos de associação genótipo-fenótipo de uma abordagem por CGAS. Nesta fase foi utilizada a ferramenta PLINK versão 1.07 disponível em <https://zzz.bwh.harvard.edu/plink/>. O PLINK permite realizar uma análise exploratória (*summary statistics*) sobre os dados da genotipagem através da aplicação de níveis de aceitação (*thresholds*) sobre parâmetros de qualidade específicos que incidem, tanto nas amostras como nos SNPs. Os valores *thresholds*

aplicados foram os recomendados para os CGAS (Anderson *et al.*, 2010). Para efetuar o controle de qualidade e a associação dos dados exploratórios foi seguido uma adaptação do tutorial de McDonald *et al.*, 2018 (<https://statgen.us/Tutorials#GWAS: Data Quality Control>).

3.4.1.2.4.1. Controle de Qualidade dos dados

3.4.1.2.4.1.1. Remoção de amostras e SNPs com o Missing Genotypes

Inicialmente foram aplicados os comandos necessários para excluir as amostras com mais de 10% de falha de *genotype call*, ou seja, amostras que foram geradas usando DNA de baixa qualidade e/ou que têm uma taxa de erro de genotipagem superior à média. Segue a visualização do script (onde BIALLELIC é o ficheiro de dados bialélicos):

```
plink --file BIALLELIC --mind 0.10 --recode --out BIALLELIC_clean_mind
```

Em segundo lugar, foram removidos os SNPs com um valor de MAF (frequência do alelo minoritário) superior a 5% e genótipos ausentes (*missing genotypes*) superior a 5%, devido à falha na genotipagem e os SNPs com MAF inferior a 5% e com falha de genotipagem superior a 1%. Os SNPs com falhas de genotipagem podem ter taxas de erro mais altas do que marcadores SNP sem dados ausentes.

```
plink --file BIALLELIC_clean_mind --maf 0.05 --recode --out MAF_greater_5  
plink --file BIALLELIC_clean_mind --exclude MAF_greater_5.map --recode --out  
MAF_less_5  
plink --file MAF_greater_5 --geno 0.05 --recode --out MAF_greater_5_clean  
plink --file MAF_less_5 --geno 0.01 --recode --out MAF_less_5_clean  
plink --file MAF_greater_5_clean --merge MAF_less_5_clean.ped  
MAF_less_5_clean.map --recode --out BIALLELIC_MAF_clean
```

Seguidamente, foi usado um critério mais rigoroso para os dados ausentes, removendo amostras com falha de genotipagem superiores a 3%:

```
plink --file BIALLELIC_MAF_clean --mind 0.03 --recode --out BIALLELIC_clean2
```

3.4.1.2.4.1.2. Teste para o equilíbrio de Hardy-Weinberg (HWE)

Cada SNP foi testado para o desvio em relação ao equilíbrio Hardy-Weinberg (HWE). Quando, verificado apenas na população controlo, o desvio dos pressupostos do HWE é frequentemente utilizado para detetar erros de genotipagem (Anderson *et al.*, 2010). Nesse caso foi utilizado um critério extremo de $p < 10^{-3}$ para rejeitar a hipótese nula de HWE:

```
plink --file BIALLELIC_clean2 --hardy
```

```
plink --file BIALLELIC_clean2 --exclude hardy.txt --recode --out BIALLELIC_clean3
```

3.4.1.2.4.2. Associação Genótipo-Fenótipo

Finalmente, foram executados os comandos para associar os dados que passaram o controlo de qualidade com a febre Q crónica.

```
plink --file BIALLELIC_clean3 --assoc
```

Com este comando é realizado o teste alélico básico de associação que consiste na comparação das frequências alélicas entre os grupos dos casos e dos controlos. O comando ainda fornece a frequência do alelo minoritário nos casos e nos controlos, o qui-quadrado do teste alélico básico com um grau de liberdade (CHISQ 1df), o valor p assintótico para este teste, a razão de chances estimada (odds ratio – OR) (para o alelo minoritário, ou seja, o alelo maioritário é a referência).

As associações genótipo - fenótipo foram consideradas estatisticamente significativas quando o valor $p \leq 0,05$ ($\alpha = 0,05$). No entanto, para minimizar erros do tipo I (falsos positivos), cujo risco aumenta com a realização de comparações múltiplas, foi utilizado o valor p ajustado calculado segundo o critério de Bonferroni ou seja, dividindo o valor de α original pelo numero de comparações realizadas (Bland e Altman, 1995).

3.4.1.2.4.3. Análise dos SNPs em associação

Para os tagSNPs, SNPs em associação estatisticamente significativa com a febre Q crónica, foram analisados na mesma região genómica os padrões de desequilíbrio de ligação (LD) na população com ascendência do norte e oeste da Europa

(CEU) usando o Haploview (Barret *et al.*, 2005). Tanto os tagSNPs como os SNPs em LD com estes foram investigados na literatura científica para identificar publicações onde tenham sido referenciados em associação com doenças infecciosas ou outras patologias.

Variants amontante do início do início da transcrição foram também analisados utilizando a base de dados de genómica funcional Genotype-Tissue Expression (GTEx), que agrupa dados de genotipagem aos padrões de expressão génica sobre todo o genoma (genome-wide gene expression patterns) numa ampla amostragem de tecidos (<https://www.gtexportal.org>) (Lonsdale *et al.*, 2013).

3.4.2. Estudo Réplica

Numa fase posterior ao estudo exploratório foi efetuado um estudo réplica no qual se pretendeu analisar variantes que foram referenciadas na literatura como estando associadas à febre Q crónica. Este novo estudo foi efetuado utilizando as amostras biológicas remanescentes do estudo inicial e recorrendo à sequenciação pelo método Sanger.

3.4.2.1. Preparação das amostras de DNA

3.4.2.1.1. Seleção de variantes

Inicialmente, através da análise da literatura disponível, foram escolhidas 4 variantes que tinham sido associadas à febre Q crónica na população holandesa para efetuar um estudo réplica para a população portuguesa. A informação sobre as variantes escolhidas está resumida na Tabela 3.5.

Tabela 3. 5. Variantes escolhidas para o estudo réplica.

Gene	Localização Genómica	SNP	Coordenada do SNP	HGVS ¹	Referência bibliográfica
<i>P2RX7</i>	12q24.31	rs1718119	121177300	NC_000012.12:g.121177300G>A	Jansen <i>et al.</i> , 2019
				NC_000012.12:g.121177300G>T	
		rs3751143	121184501	NC_000012.12:g.121184501A>C	
				NC_000012.12:g.121184501A>G	
<i>IFNG</i>	12q15	rs1861493	68157416	NC_000012.12:g.68158742T>A	Wielders <i>et al.</i> , 2015
				NC_000012.12:g.68158742T>G	
		rs2430561	68158742	NC_000012.12:g.68157416G>A	
				NC_000012.12:g.68157416G>C	
				NC_000012.12:g.68157416G>T	

¹ Nomenclatura das variantes genéticas de acordo com a *Human Genome Variation Society* (<https://www.hgvs.org/>)

3.4.2.1.2. Desenho de *primers*

Seguidamente, foram desenhados *primers* específicos para os fragmentos a amplificar através da ferramenta *Primer-BLAST* disponibilizada pelo NCBI em <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>. A Tabela 3.7 apresenta os *primers* utilizados para amplificar os fragmentos que contêm as variantes em análise.

Tabela 3. 6. Sequência dos *primers* e tamanho dos fragmentos amplificados dos genes *P2RX7* e *IFNG*

Gene	SNP	Primer forward	Primer reverse	Tamanho do Ampliação (pb)
P2RX7	rs1718119	5'-TCCGCTCCCTGATAGAACCA-3'	5'-CCCAGCGACGTATCCACTAC-3'	511
	rs3751143	5'-GAACCTAGAACCTGAGGGCTT-3'	5'-GTGCCTGGCTTCAGTAAGGA-3'	565
IFNG	rs1861493	5'-TCCATTTCGTGTTTGGGTGAC-3'	5'-GGAGTCTAAGGAGTTCTCCAGC-3'	646
	rs2430561	5'-GGAACCTTCGTTGCTCACTGG-3'	5'-GCTGATACTCCAAAGGTCCCA-3'	468

3.4.2.1.3. Amplificação de DNA por PCR

A amplificação por PCR realizou-se num termociclador *T Gradient Biometra*®. Inicialmente foi efetuado um PCR com um gradiente de temperaturas (65°C a 70°C), com diferentes concentrações de magnésio e betaína, a fim de se obter um protocolo otimizado.

A amplificação foi efetuada com recurso aos *primers forward* e *reverse* descritos na Tabela 3.6. numa mistura de reação com um volume total de 12,5 µL utilizando a Taq DNA polimerase da Fermentas® (Tabela 3.7.)

Tabela 3. 7. Mistura reacional para amplificação dos fragmentos dos genes *P2RX7* e *IFNG*

Reagente	Concentração da solução de trabalho	Volume por reação (µL)	Concentração final
Taq buffer (com (NH ₄) ₂ SO ₄)	10x	1,25	1x
dNTPs	2,5 mM cada	2	0,4mM cada
primer forward	25 µM	1,25	2,5 µM
primer reverse	25 µM	1,25	2,5 µM
Taq DNA polimerase (fermentas®)	5 U/µL	0,125	0,625 U por reação
MgCl ₂	25mM	1,5	3mM
Betaina	5M	2,5	1M
H ₂ O	--	2,125	--
DNA	30 -100 ng/µL	0,5	1,2- 4 ng/µL

O programa de PCR (Tabela 3.8) iniciou-se com um passo de desnaturação inicial a 95°C durante 5 min, seguida de 35 ciclos de desnaturação a 95°C durante 30 segundos, hibridação a 65°C durante 40 segundos e extensão a 72°C durante 40 segundos; A extensão final ocorreu a 72°C durante 7 minutos.

Tabela 3. 8. Condições da amplificação por PCR dos fragmentos dos genes *P2RX7* e *IFNG*

Etapas do PCR	Temperatura	Tempo	Nº de Ciclos
Desnaturaçã o Inicial	95°C	5 min.	-
Desnaturaçã o	95°C	30 seg.	
Hibridaçã o	65°C	40 seg	35
Extensã o	72°C	40 seg	
Extensã o Final	72°C	7 min	-

3.4.2.1.4. Eletroforese em gel de agarose

A verificação da amplificação dos fragmentos foi efetuada, tal como anteriormente, por eletroforese seguida de visualização do resultado através de um dispositivo de emissão de raios-ultravioleta acoplado a um sistema de captação de imagem. A corrida de eletroforese decorreu a uma voltagem constante de 90 V durante 30 minutos gerada pela fonte de tensão *BIORAD® PowerPac300*.

3.4.3. Sequenciação automática pelo método de Sanger

O fluxo de trabalho da sequenciação Sanger por eletroforese capilar envolveu as etapas descritas na Figura 3.3., ou seja: amplificação por PCR; purificação dos produtos de PCR; reação de sequenciação; eletroforese capilar; e análise de dados.



Figura 3. 3. Fluxo de trabalho da sequenciação Sanger por eletroforese capilar.

Fonte: Adaptado de <https://www.thermofisher.com/pt/en/home/life-science/sequencing/sanger-sequencing/sanger-sequencing-workflow.html>

3.4.3.1. Reação de Purificação

Após a amplificação por PCR foi necessário purificar os produtos de PCR, isto é remover os restos de *primers* e dNTPs, para que o próximo passo, a sequenciação cíclica, ocorra sem interferentes. Deste modo, a 2,5 μ L de cada produto de PCR foi adicionado 1 μ L da solução enzimática *ExoProStar* (Alfagene, Carcavelos, Portugal). Esta solução contém Exonuclease I e fosfatase alcalina. Após a homogeneização, procedeu-se à colocação no termociclador *T Gradient Biometra*® com o seguinte perfil térmico: 37°C durante 15 minutos seguida de 80°C durante 15 minutos.

3.4.3.2. Reação de Sequenciação Sanger

A sequenciação pelo método de Sanger foi efetuada no termociclador *T Gradient Biometra*® com o recurso ao kit comercial *BigDye*® *Terminator V 1.1 Cycle Sequencing Kit* (*Applied Biosystems*®). A mistura de reação foi efetuada para um volume total de 10 μ L (Tabela 3.9).

Tabela 3. 9. Mistura da reação de sequenciação Sanger

Mistura de reação	Concentração da solução stock	Volume p/ 1xreação	Concentração final
Tampão Big Dye	5x	3.5 μ L	1,75 X
<i>Big Dye</i>	-	0.5 μ L	na
Primer	2 μ M	1.0 μ L	0,2 μ M
H ₂ O	-	3.5 μ L	-
Produto de reação purificado	-	1.5 μ L	-

As condições da sequenciação cíclica encontram-se descritas na Tabela 3.10. O programa sequenciação iniciou-se com um passo de desnaturação inicial a 96°C durante 1 min, seguida de 25 ciclos de desnaturação a 96°C durante 10 segundos, hibridação a 50°C durante 7 segundos e extensão a 60°C durante 4 minutos.

Tabela 3. 10, Condições da sequenciação cíclica pelo método de Sanger

Etapas da Sequenciação	Temperatura	Tempo	Nº de Ciclos
Desnaturação Inicial	96°C	1 minuto	-
Desnaturação	96°C	10 segundos	
Hibridação	50°C	7 segundos	25
Extensão	60°C	4 minutos	

Após a realização deste procedimento, enviaram-se os produtos para a Unidade de Tecnologia e Inovação (UTI) do Instituto Nacional de Saúde Doutor Ricardo Jorge (INA) para que fossem executadas as etapas seguintes de purificação dos produtos de sequenciação e de eletroforese capilar através do equipamento *3130 Genetic Analyser, Abi Prism (Applied Biosystems)*.

3.4.4. Análise de dados

A visualização das sequências de DNA obtidas e a análise das variantes foram efetuadas com recurso ao programa *Finch TV* versão 1.4.0 disponível para download em <https://finchtv.software.informer.com/1.4/>.

A análise estatística da distribuição alélica e genotípica foi efetuada segundo as recomendações (David, 2021). A análise estatística foi realizada utilizando o teste de independência do qui-quadrado com 1 grau de liberdade no caso de uma análise da distribuição das frequências alélicas de um marcador bialélico. Para este teste utilizamos o calculador disponível na website da VassarStats: (*Web Site for Statistical Computation* <http://vassarstats.net/index.html>). Dado que para as variantes *P2RX7* foi sugerido um modo de hereditariedade recessivo (Tabela 1.1.) (Jansen *et al.*, 2019), considerou-se ainda relevante verificar os modelos de hereditariedade. Para a análise dos modelos de hereditariedade (dominante, recessivo e sobredominante) foi utilizado o teste de independência do qui-quadrado. O *Cochran - Armitage trend test* foi utilizado para análise da distribuição das frequências genotípicas sem pressuposto sobre o modo de hereditariedade. Este teste foi efetuado utilizando o calculador em <https://anikhachatryan.github.io/CA-trend-test-calculator/>.

O equilíbrio de Hardy-Weinberg foi realizado na população controlo para todas as variantes utilizando o teste de independência do qui-quadrado com 1 grau de liberdade.

Os resultados obtidos no estudo exploratório e no estudo réplica encontram-se descritos no Capítulo IV – Resultados.

Capítulo IV – Resultados

4.1. Resultados do Estudo Exploratório

No estudo exploratório, os fragmentos dos genes *IL12RB1*, *IFNGR1* e *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* e *TIMP1* de cada uma das 94 amostras foram amplificados com sucesso. A Figura 4.1. representa as bandas características de cada um dos fragmentos amplificados.

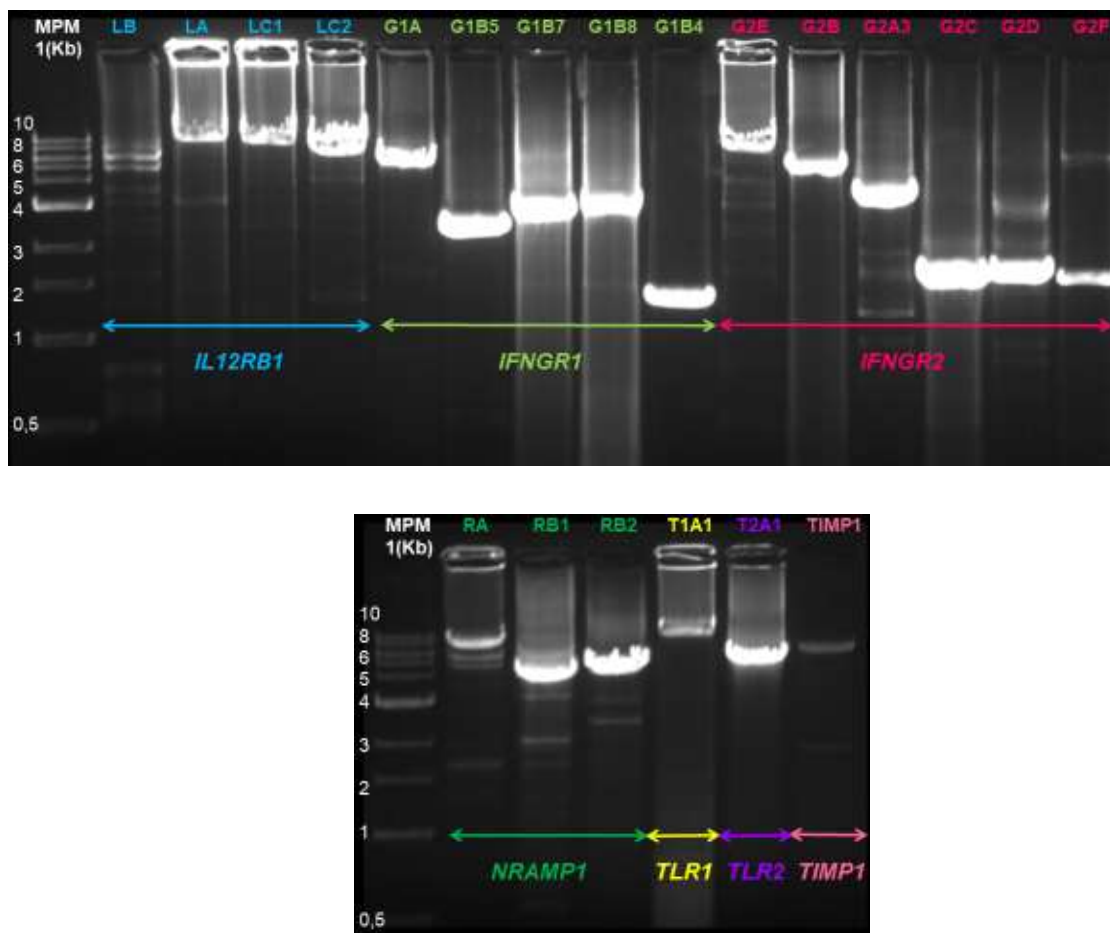


Figura 4. 1. Resultado eletroforese de cada fragmento dos genes *IL12RB1*, *IFNGR1* e *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* e *TIMP1*.

Legenda: MPM - Marcador de peso molecular, Quick-Load® 1 Kb DNA Ladder da *New England BioLabs* ©Inc (Anexo 3).

A intensidade das bandas no gel foi classificada e anotada entre 1x, 3x, 5x e 10x concentrado (sendo 1 a classificação mais baixa, e 10 a classificação mais elevada, logo a mais concentrada) para cada um dos produtos amplificados. A seguir, a visualização da padronização das concentrações, foi realizada na etapa do *pooling* das amostras, efetuada através do agrupamento de 4 fragmentos de tamanhos distintos em 5 grupos.

Nesta etapa, a verificação da padronização das concentrações foi visualizada para todas as amostras através dos 5 géis de agarose (com os 20 fragmentos amplificados por cada amostra), tais como representados na figura 4.2.

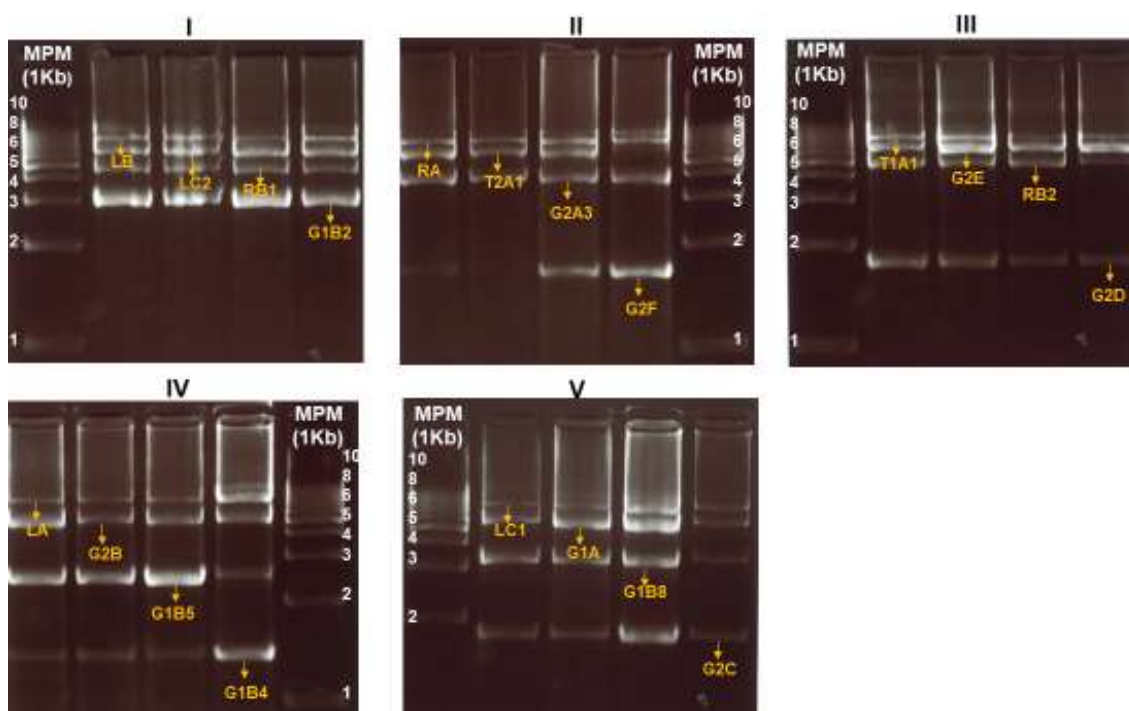


Figura 4. 2. Resultado eletroforese dos 5 agrupamentos de 4 fragmentos cada.

Legenda: I – Gel relativo à amplificação dos fragmentos LB (13559 pb), LC2 (7576 pb), RB1(4851 pb) e G1B7(3206pb); II – Gel relativo à amplificação dos fragmentos RA (8823 pb), T2A1 (6185 pb), G2A3 (3790 pb) e G2F(1636 pb); III – Gel relativo à amplificação dos fragmentos T1A1 (12110 pb), G2E (8034 pb), RB2 (5217 pb) e G2D (1793pb); IV– Gel relativo à amplificação dos fragmentos LA (10760 pb), G2B (5587pb), G1B5 (2626 pb) e G1B4 (1415 pb); V - Gel relativo à amplificação dos fragmentos LC1 (9748 pb), G1A (6069 pb) e G1B8 (3330 pb) e G2C (1821pb); MPM - Marcador de peso molecular, Quick-Load® 1 Kb DNA Ladder da *New England BioLabs* ®Inc (Anexo 3).

A análise exploratória dos dados efetuada na Fase III de NGS, aplicando os critérios e ajustes específicos do controlo de qualidade para estudos de associação genótipo-fenótipo de uma abordagem por CGAS, encontra-se esquematizada na Figura 4.3. São indicados os *thresholds* utilizados e as amostras de DNA e SNPs excluídos em cada uma das etapas.

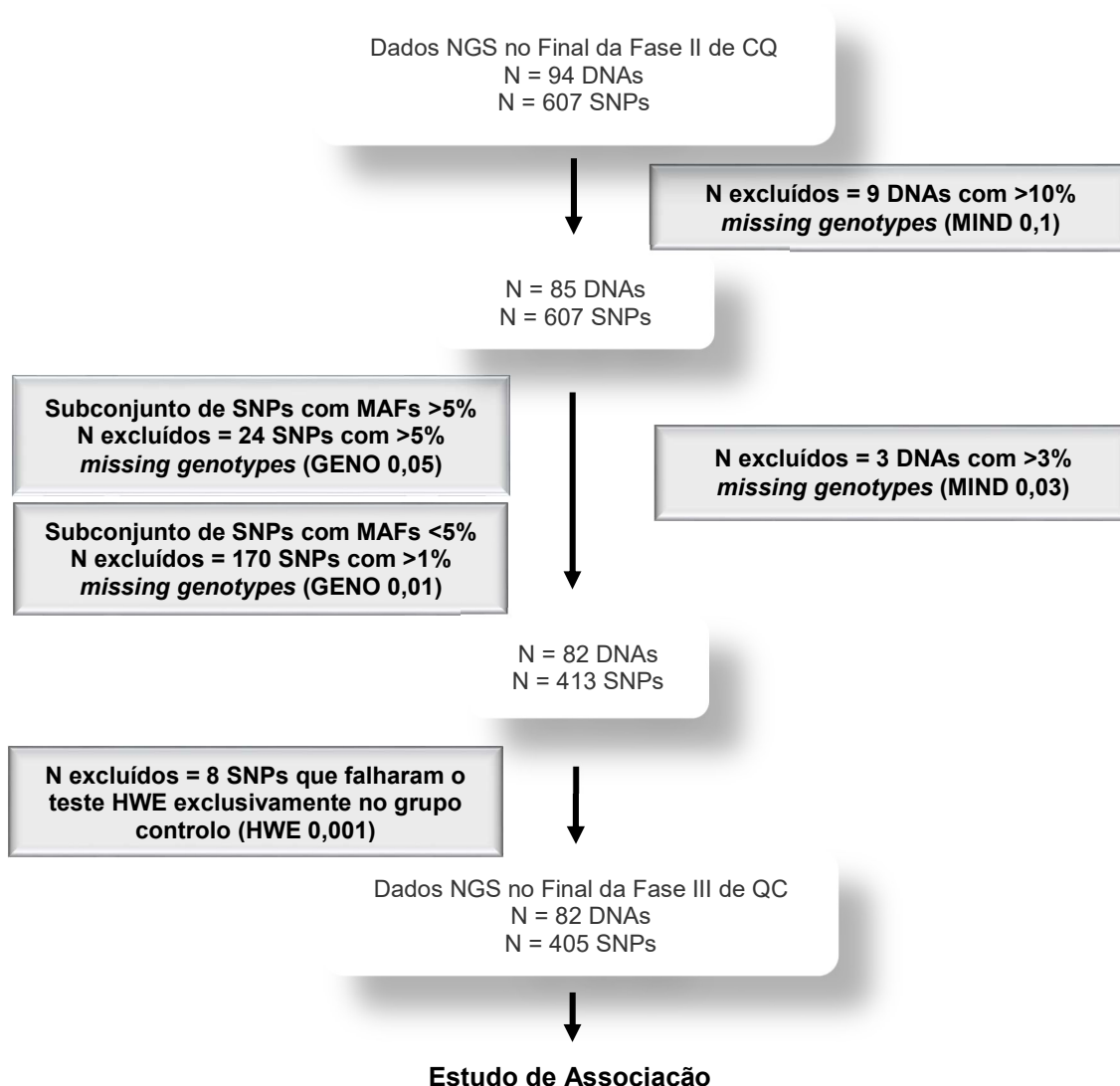


Figura 4. 3. Fase III do Controle de Qualidade (CQ) dos dados de genotipagem

Fonte: Adaptado do curso de comunicação de McDonald e colaboradores (McDonald *et al.*, 2018).

No final da Fase II de NGS, após a anotação e filtração, contabilizaram-se 607 SNPs provenientes de 94 amostras de DNAs. Na primeira etapa de controle de qualidade dos dados genotipados, efetuado com auxílio da ferramenta PLINK, excluíram-se 9 amostras de DNA que continham uma percentagem de genótipos ausentes (*missing genotypes*) superior a 10%. Deste modo, passaram à fase posterior 85 amostras de DNA e 607 SNPs.

Seguidamente, foram excluídos 24 SNPs que possuíam cumulativamente uma MAF (frequência de alelo minoritário) superior a 5% (variantes comuns) e uma percentagem de genótipos ausentes de 5%. Foram também excluídos 170 SNPs que

possuíam cumulativamente uma MAF inferior a 5 % (variantes raras) e uma percentagem de genótipos ausentes de 1%. Ao aplicar um critério mais rigoroso para os dados ausentes, foram, ainda, removidas 3 amostras com falha de genotipagem superior a 3%. Desta forma, foram submetidos ao critério seguinte, 82 amostras de DNA e 413 SNPs.

De seguida, foram excluídos 8 SNPs que possuíam um desvio em relação ao equilíbrio Hardy-Weinberg (HWE) com um valor de $p < 10^{-3}$.

Globalmente, passaram o controlo de qualidade 405 SNPs e 82 amostras de DNA que foram utilizados para efetuar estudos de associação genótipo-fenótipo de uma abordagem por CGAS à febre Q crónica.

Após a aplicação de comandos específicos do software PLINK para associar os SNPs encontrados à febre Q crónica, detetou-se que 324 SNPs apresentavam variações entre os doentes crónicos (casos) e os doentes agudos (controlos). Destes apenas 29 SNPs demonstraram associação estatisticamente significativa com a febre Q crónica ($p < 0,05$) e apenas quatro resistiram à correção de Bonferroni para testes múltiplos (α ajustado = 0,000154321) (Tabela 4.1.).

Tabela 4. 1. Resultados estatisticamente significativos para o teste alélico de associação com a febre Q crônica

SNP	Localização	Alelo	MAF	MAF	Alelo	χ^2	P	OR
		minoritário	Crônicos	Agudos	maioritário			
rs78407108	21:33426675	A	0,20	0,00	G	16,11	0,0000597***	NA
rs17879956	21:33438631	T	0,20	0,00	C	16,11	0,0000597***	NA
rs7277167	21:33401471	T	0,20	0,00	C	15,71	0,0000740***	NA
rs9974603	21:33403994	A	0,23	0,01	C	14,53	0,0001382***	22,22
rs9808753	21:33415005	G	0,20	0,01	A	12,17	0,0004844***	18,75
rs412133	19:18094574	T	0,43	0,18	C	7,032	0,0080070**	3,39
rs374326	19:18090456	C	0,43	0,20	T	6,161	0,0130600*	3,11
rs418892	19:18094975	T	0,43	0,20	C	6,161	0,0130600*	3,11
rs419540	19:18095263	T	0,43	0,20	C	6,161	0,0130600*	3,11
rs845380	19:18070759	A	0,43	0,20	G	5,771	0,0162900*	3,01
rs117612494	19:18059422	T	0,07	0,00	C	5,164	0,0230600*	NA
rs184963792	21:33412855	G	0,07	0,00	A	5,164	0,0230600*	NA
rs121913199	21:33421400	A	0,07	0,00	G	5,164	0,0230600*	NA
rs185459305	21:33427424	T	0,07	0,00	C	5,164	0,0230600*	NA
rs2045386	19:18071456	A	0,18	0,39	G	4,282	0,0385100*	0,33
rs17879591	19:18060585	T	0,30	0,13	C	4,147	0,0417000*	2,83
rs17878594	19:18062703	T	0,30	0,13	C	4,147	0,0417000*	2,83
rs17879124	19:18062748	A	0,30	0,13	G	4,147	0,0417000*	2,83
rs17878265	19:18066763	A	0,30	0,13	G	4,147	0,0417000*	2,83
rs12150884	19:18069163	A	0,07	0,24	G	4,069	0,0436700*	0,23
rs2305739	19:18069384	A	0,07	0,24	G	4,069	0,0436700*	0,23
rs2305740	19:18069426	G	0,07	0,24	A	4,069	0,0436700*	0,23
rs2305741	19:18069455	A	0,07	0,24	G	4,069	0,0436700*	0,23
rs453427	19:18072680	T	0,07	0,24	G	4,069	0,0436700*	0,23
rs447009	19:18074382	A	0,07	0,24	C	4,069	0,0436700*	0,23
rs404068	19:18092159	G	0,07	0,24	A	4,069	0,0436700*	0,23
rs433821	19:18093031	G	0,07	0,24	A	4,069	0,0436700*	0,23
rs845332	19:18093283	G	0,07	0,24	C	4,069	0,0436700*	0,23
rs438421	19:18065276	A	0,37	0,18	G	3,973	0,0462300.	2,56

Legenda: MAF crônicos – frequência do alelo minoritário nos doentes crônicos; MAF agudos – frequência do alelo minoritário nos doentes agudos; χ^2 – Teste alélico básico qui-quadrado; P – valor p assintótico para o teste do qui-quadrado; OR – Odds ratio estimado (para o alelo minoritário, ou seja, o alelo maioritário é a referência); Códigos de significância: 0 **** 0,001 *** 0,01 ** 0,05 * 0,1 ' ' .

Em relação à variante rs78407108, verificou-se, pelo teste alélico do PLINK a indicação de um efeito deletério do alelo minoritário A (MAF crônicos=20%; MAF agudos=0%; χ^2 =16,11; p= 0,0000597).

Quanto à variante rs17879956, verificou-se, pelo teste alélico do PLINK, a indicação de um efeito deletério do alelo minoritário T (MAF crônicos = 20%; MAF agudos = 0%; $\chi^2 = 16,11$; $p = 0,0000597$).

A variante rs7277167, apresentou, pelo teste alélico do PLINK, a indicação de um efeito deletério do alelo minoritário T (MAF crônicos = 20%; MAF agudos = 0%; $\chi^2 = 15,71$; $p = 0,0000740$).

Para as três variantes descritas acima não foi possível calcular o valor de odds ratio, uma vez que os doentes agudos não apresentavam qualquer frequência para os respectivos alelos minoritários.

Em relação à variante rs9974603, através do teste alélico do PLINK, verificou-se a indicação de um efeito deletério do alelo minoritário A (MAF crônicos=23%; MAF agudos= 1%; $\chi^2=14,53$; $p=0,0001382$; OR=22,22). O resultado resiste à correção de Bonferroni para comparações múltiplas ($\alpha = 0,05$ e α ajustado = 0,000154321). A associação foi igualmente significativa aplicando o teste exato de Fisher (*Fisher exact test*), visto existirem valores esperados das frequências das categorias na tabela de contingência inferiores a 5, o que não é preconizado na utilização do teste de qui-quadrado. Com este teste o OR=22,22 e o intervalo de confiança não contem o valor 1 (95 % IC 2,60 – 190, 18; $p=0,0006$).

Na Tabela 4.2 estão indicadas as frequências alélicas e genotípicas das quatro variantes do gene *IFNGR2*, com associação significativa à febre Q crônica que resistiram à correção de Bonferroni, em diferentes populações do Projeto 1000 Genomas (The 1000 Genomes Project Consortium, 2015), CEU (Habitantes do Estado do Utah com ascendência do norte e oeste da Europa) e IBS (Ibérica) e na população portuguesa alvo do presente estudo.

Das populações representadas no Projeto 1000 Genomas, a subpopulação CEU (Habitantes do Estado do Utah com ascendência do norte e oeste da Europa) era a geralmente usada para a comparação com a população portuguesa, uma vez que a população ibérica se encontrava sub-representada (David *et al.*, 2018 a e b). Atualmente essa situação já se encontra corrigida, pelo que a população ibérica de Espanha também foi alvo da nossa análise.

Para uma melhor visualização das frequências alélicas e genotípicas das subpopulações CEU e IBS e dos grupos alvo do presente estudo optou-se por representar esses dados com recurso a gráficos *pie charts* (Figuras 4.4 a Figura 4.11)

Tabela 4. 2. Frequências alélicas e genóticas das variantes do gene IFNGR2 com associação significativa que resistiram ao critério de Bonferroni (0,00016339869) em diferentes populações do Projeto 1000 Genomas e na população portuguesa estudada.

Variante	Tipo de variante	Genótipo	Populações do Projeto 1000 Genomas (2015)				População Estudada			
			Mundial (n=2504)	Europeia (n= 503)	CEU (n=99)	IBS (n=107)	Total (n=82)	Crónicos (n=15)	Agudos (n= 38)	PGP (n=29)
rs78407108 NC_000021.9:g.33426675G>A	Intrónica	GG	0,6890 (1725/2504)	0,8052 (405/503)	0,8283 (82/99)	0,8785 (94/107)	0,8415 (69/82)	0,6000 (9/15)	1,0000 (38/38)	0,7586 (22/29)
		GA	0,2396 (600/2504)	0,1769(89/503)	0,1718 (17/99)	0,1028 (11/107)	0,1585 (13/82)	0,4000 (6/15)	0,0000 (0/38)	0,2414 (7/29)
		AA	0,0715 (179/2504)	0,0179 (9/503)	0,0000 (0/99)	0,0187 (2/107)	0,0000 (0/82)	0,0000 (0/15)	0,0000 (0/38)	0,0000 (0/29)
		Alelo A	0,1913 (958/5008)	0,106 (107/1006)	0,0859 (17/198)	0,0700 (15/214)	0,07927 (13/164)	0,2000 (6/30)	0,0000 (0/76)	0,1207 (7/58)
		Alelo G	0,8087 (4050/5008)	0,8936 (899/1006)	0,9141 (181/198)	0,9299 (199/214)	0,92073 (151/164)	0,8000 (24/30)	1,0000 (76/76)	0,8793 (51/58)
rs17879956 NC_000021.9:g.33438631C>T	Intrónica	CC	0,5735 (1436/2504)	0,7972 (401/503)	0,8081 (80/99)	0,8785 (94/107)	0,8171 (67/82)	0,6000 (9/15)	1,0000 (38/38)	0,6897 (20/29)
		CT	0,3415 (855/2504)	0,1809 (91/503)	0,1919 (19/99)	0,1028 (11/107)	0,1829 (15/82)	0,4000 (6/15)	0,0000 (0/38)	0,3103 (9/29)
		TT	0,0851 (213/2504)	0,0219 (11/503)	0,0000 (0/99)	0,0187 (2/107)	0,0000 (0/82)	0,0000 (0/15)	0,0000 (0/38)	0,0000 (0/29)
		Alelo T	0,256 (1281/5008)	0,1123 (113/1006)	0,0960 (19/198)	0,0701 (15/214)	0,0915 (15/164)	0,2000 (6/30)	0,0000 (0/76)	0,1552 (9/58)
		Alelo C	0,7442 (3727/5008)	0,8877 (893/1006)	0,9040 (179/198)	0,9299 (199/214)	0,9085 (149/164)	0,8000 (24/30)	1,0000 (76/76)	0,8448 (49/58)
rs7277167 NC_000021.9:g.33401471C>T	Variante a montante do gene	CC	0,6270 (1570/2504)	0,8032 (404/503)	0,8283 (82/99)	0,8785 (94/107)	0,8272 (67/81)	0,6000 (9/15)	1,0000 (37/37)	0,7241 (21/29)
		CT	0,2955 (740/2504)	0,1809 (91/503)	0,1717 (17/99)	0,1028 (11/107)	0,1728 (14/81)	0,4000 (6/15)	0,0000 (0/37)	0,2759 (8/29)
		TT	0,0775 (194/2504)	0,0159 (8/503)	0,0000 (0/99)	0,0187 (2/107)	0,0000 (0/81)	0,0000 (0/15)	0,0000 (0/37)	0,0000 (0/29)
		Alelo T	0,2252 (1128/5008)	0,1064 (107/1006)	0,0859 (17/198)	0,0701 (15/214)	0,0864 (14/162)	0,2000 (6/30)	0,0000 (0/74)	0,1379 (8/58)
		Alelo C	0,7748 (3880/5008)	0,8936 (899/1006)	0,9141 (181/198)	0,9299 (199/214)	0,9136 (148/162)	0,8000 (24/30)	1,0000 (74/74)	0,8621 (50/58)
rs9974603 NC_000021.9:g.33403994C>A	Intrónica	CC	0,4884 (1223/2504)	0,7455 (375/503)	0,7374 (73/99)	0,8318 (89/107)	0,8148 (66/81)	0,6000 (9/15)	0,9730 (36/37)	0,7241 (21/29)
		CA	0,3750 (939/2504)	0,2247 (113/503)	0,2424 (24/99)	0,1495 (16/107)	0,1728 (14/81)	0,3333(5/15)	0,0270 (1/37)	0,2759 (8/29)
		AA	0,1366 (342/2504)	0,0298 (15/503)	0,0202 (2/99)	0,0187 (2/107)	0,0123 (1/81)	0,0667 (1/15)	0,0000 (0/37)	0,0000 (0/29)
		Alelo A	0,3241 (1623/5008)	0,1421 (143/1006)	0,1414 (28/198)	0,0935 (20/214)	0,0988(16/162)	0,2333 (7/30)	0,0135 (1/74)	0,1379 (8/58)
		Alelo C	0,6759 (3385/5008)	0,8579 (863/1006)	0,8586 (170/198)	0,9065 (194/214)	0,9012 (146/162)	0,7667 (23/30)	0,9865 (73/74)	0,8621 (50/58)

Legenda: Mundial - População mundial estudada no Projeto 1000 Genomas (versão de 2015); Europeia- População europeia estudada no Projeto 1000 Genomas (versão 2015) que inclui as subpopulações CEU, GBR, FIN, IBS e TSI; CEU- Habitantes do Estado do Utah com ascendência do norte e oeste da Europa; GBR- Britânicos da Inglaterra e Escócia; FIN- Finlandeses na Finlândia; IBS- População Ibérica em Espanha; TSI - Toscanos em Itália; PGP- População Geral Portuguesa analisada neste estudo.

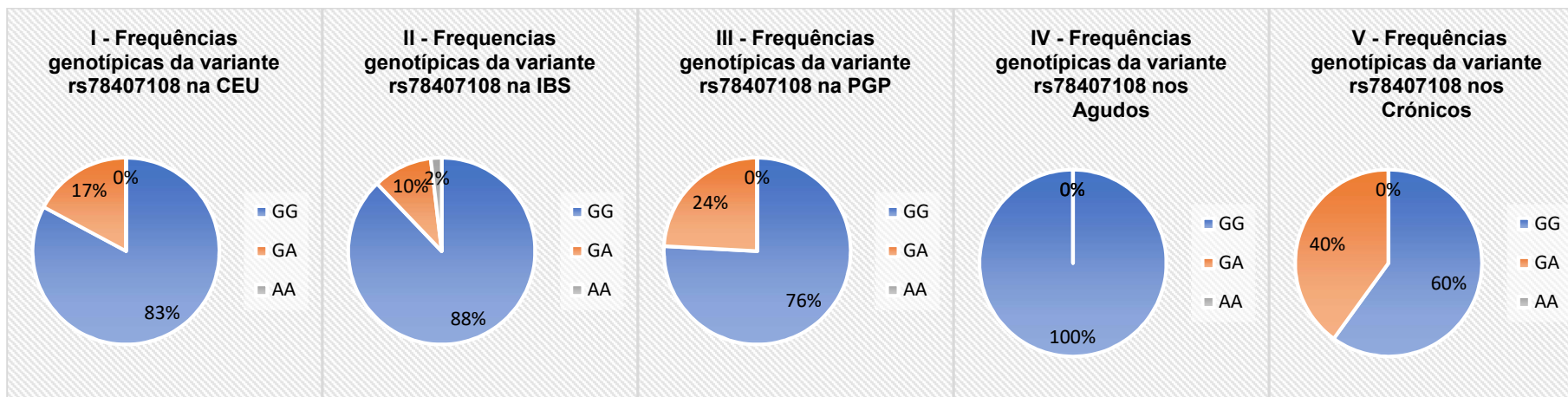


Figura 4. 5. Gráficos *pie chart* relativos às frequências genotípicas da variante rs78407108 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

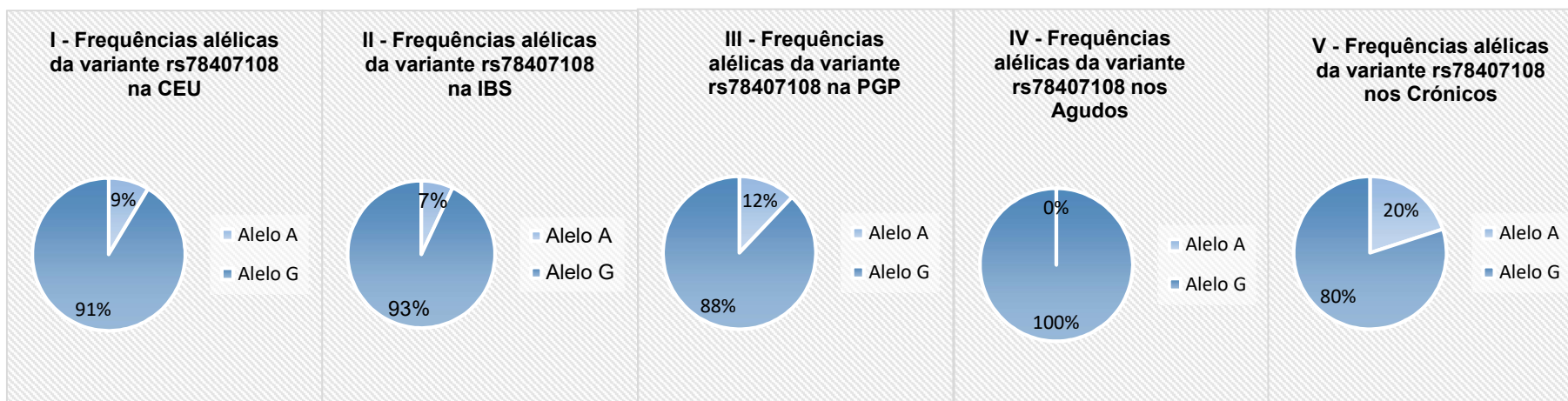


Figura 4. 4. Gráficos *pie chart* relativos às frequências alélicas da variante rs78407108 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

O SNP rs78407108 é uma variante intrónica cuja distribuição genotípica na subpopulação CEU do Projeto 1000 genomas é de, aproximadamente, 83% de homozigóticos GG, 17% de heterozigóticos GA e de 0% de homozigóticos AA (Figura 4.4. I). A frequência alélica do alelo A (minoritário) é de 9% e a frequência alélica do alelo G (maioritário) é de 91% (Figura 4.5. I).

Na subpopulação IBS do Projeto 1000 Genomas, o SNP rs78407108 apresenta uma distribuição genotípica um pouco diferente, sendo de, aproximadamente, 88% de homozigóticos GG, 10% de heterozigóticos GA e de 2% de homozigóticos AA (Figura 4.4. II). A frequências do alelo A é de 7% e a frequência alélica do alelo G (maioritário) é de 93% (Figura 4.5. II).

No grupo de 29 indivíduos analisados como amostra da população geral portuguesa (PGP) verifica-se uma percentagem de heterozigóticos (24%) mais elevada do que as percentagens das subpopulações CEU e IBS e, conseqüentemente, uma percentagem de homozigóticos GG mais baixa de 76% (Figura 4.4. III). Neste grupo, tal como na subpopulação CEU, não verificámos a existência de indivíduos com o genótipo homozigótico AA. A frequência alélica do alelo A foi de 12% e a frequência alélica do alelo G de 88% (Figura 4.5. III), observando-se um valor de MAF ligeiramente superior ao apresentado pelas subpopulações CEU (9%) e IBS (7%).

No grupo de doentes agudos verificou-se que todos os indivíduos possuíam um genótipo GG (Figura 4.4. IV) e, conseqüentemente, a frequência alélica do alelo G é de 100% (Figura 4.5. IV).

Nos doentes crónicos a distribuição genotípica apresentada foi de, aproximadamente, 60% de homozigóticos GG, 40% de heterozigóticos GA e de 0% de homozigóticos AA (Figura 4.4. V), ou seja, a percentagem de heterozigóticos é bastante superior à apresentada nos restantes grupos (CEU, IBS, PGP e agudos). A frequência alélica do alelo A é de 20% e a frequência alélica do alelo G é de 80% (Figura 4.5. V), observando-se um valor de MAF superior em relação aos outros grupos analisados.

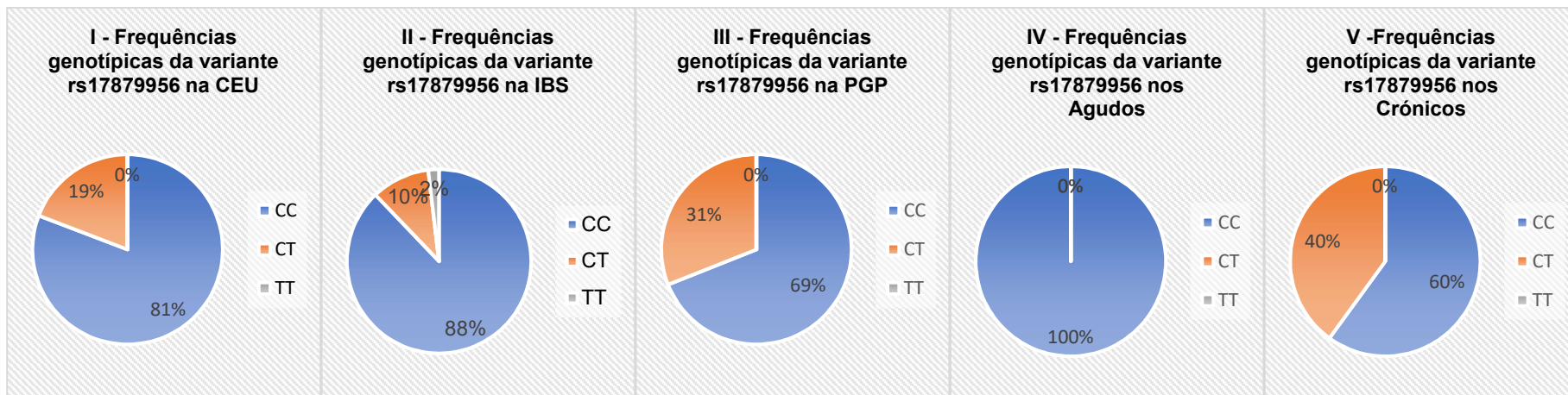


Figura 4. 6. Gráficos *pie chart* relativos às frequências genótípicas da variante rs17879956 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

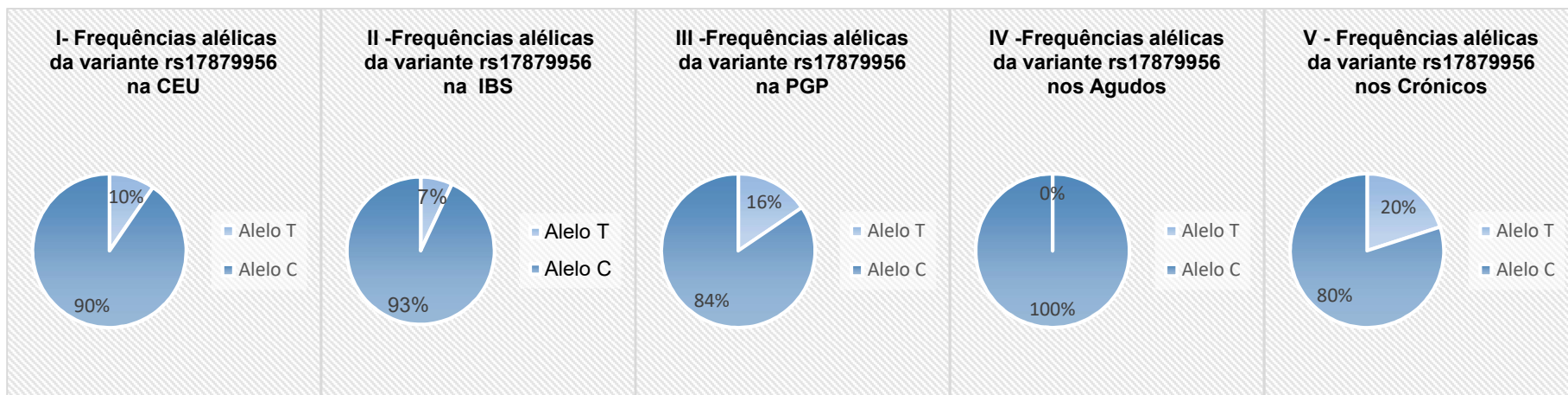


Figura 4. 7. Gráficos *pie chart* relativos às frequências genótípicas da variante rs17879956 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

O SNP rs17879956 é uma variante intrónica cuja distribuição genotípica na subpopulação CEU do Projeto 1000 genomas é de, aproximadamente, 81% de homozigóticos CC, 19% de heterozigóticos CT e de 0% de homozigóticos TT (Figura 4.6 I). A frequência alélica do alelo T (minoritário) é de 10% e a frequência alélica do alelo C (maioritário) é de 90% (Figura 4.7. I).

Na subpopulação IBS do Projeto 1000 Genomas, esta variante apresenta uma distribuição genotípica distinta, sendo de, aproximadamente, 88% de homozigóticos CC, 10% de heterozigóticos CT e de 2% de homozigóticos TT (Figura 4.6.II). A frequências do alelo T (minoritário) é de 7% e a frequência alélica do alelo C (maioritário) é de 93% (Figura 4.7. II).

Na nossa amostra da população geral portuguesa (PGP) verificou-se uma percentagem superior de heterozigóticos (31%) quando comparada com as subpopulações CEU e IBS, e uma percentagem de homozigóticos de 69% (Figura 4.6.III). Neste grupo, tal como na subpopulação CEU, não verificámos a existência de indivíduos com o genótipo homozigótico TT. A frequência alélica do alelo T é de 16 % e a frequência alélica do alelo C é de 84 % (Figura 4.7. III), observando-se um valor de MAF superior ao apresentado quer pela subpopulação CEU quer pela subpopulação IBS.

Nos doentes agudos, verificou-se que todos os indivíduos possuíam um genótipo CC (Figura 4.6. IV). e, conseqüentemente, a frequência alélica do alelo C foi de 100% (Figura 4.7. IV).

Nos doentes crónicos a distribuição genotípica apresentada é de, aproximadamente, 60% de homozigóticos CC, 40% de heterozigóticos CT e de 0% de homozigóticos TT, ou seja, a percentagem de heterozigóticos é superior à apresentada pelos restantes grupos (Figura 4.6.V). A frequência alélica do alelo T é de 20% e a frequência alélica do alelo C é de 80% (Figura 4.7. V), observando-se um valor de MAF superior em relação aos outros grupos analisados.

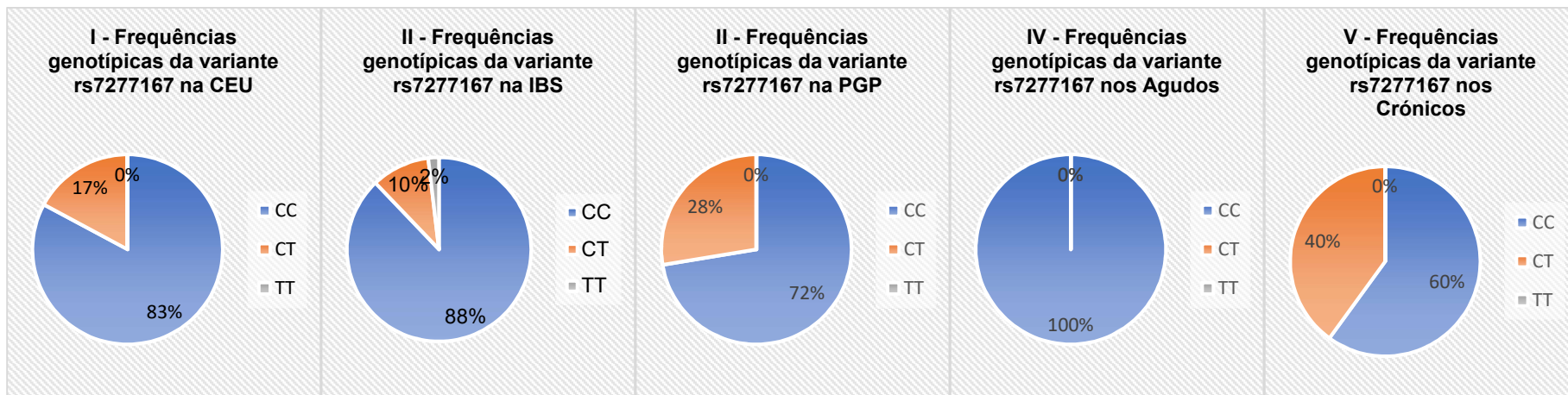


Figura 4. 8. Gráficos *pie chart* relativos às frequências genótípicas da variante rs7277167 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

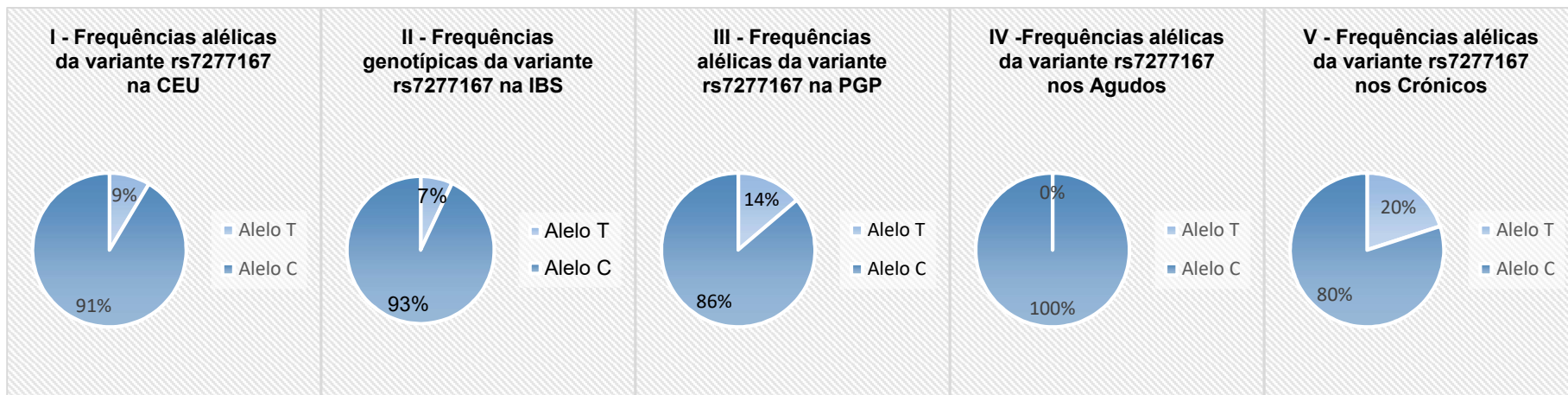


Figura 4. 9. Gráficos *pie chart* relativos às frequências alélicas da variante rs7277167 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

O SNP rs7277167 é uma variante na região 5' UTR, ou seja, a montante do gene, cuja distribuição genotípica na subpopulação CEU do Projeto 1000 genomas é de, aproximadamente, 83% de homozigóticos CC, 17% de heterozigóticos CT e de 0% de homozigóticos TT (Figura 4.8. I). A frequência alélica do alelo T (minoritário) é de 9% e a frequência alélica do alelo C (maioritário) é de 91% (Figura 4.9. I).

Na subpopulação IBS do Projeto 1000 Genomas, esta variante apresenta uma distribuição genotípica distinta, sendo de, aproximadamente, 88% de homozigóticos CC, 10% de heterozigóticos CT e de 2% de homozigóticos TT (Figura 4.8.II). A frequências do alelo T (minoritário) é de 7% e a frequência alélica do alelo C (maioritário) é de 93% (Figura 4.9.II).

Na nossa amostra da população geral portuguesa (PGP), verificou-se uma percentagem superior de heterozigóticos (28%) quando comparada com as subpopulações CEU e IBS, e uma percentagem de homozigóticos de 72% (Figura 4.8. III). Também neste grupo não se verifica a existência de indivíduos com o genótipo homozigótico TT. A frequência alélica do alelo T é de 14% e a frequência alélica do alelo C é de 86% (Figura 4.9 III), observando-se um valor de MAF superior ao apresentado pela subpopulação CEU.

No grupo controlo do nosso estudo, ou seja, nos doentes agudos, verificou-se que todos os indivíduos possuíam um genótipo CC (Figura 4.8. IV) e, conseqüentemente, a frequência alélica do alelo C é de 100% (Figura 4.9. IV).

Nos doentes crónicos a distribuição genotípica apresentada é de, aproximadamente, 60% de homozigóticos CC, 40% de heterozigóticos CT e de 0% de homozigóticos TT, ou seja, a percentagem de heterozigóticos é superior à apresentada nos restantes grupos (Figura 4.8. V). A frequência alélica do alelo T é de 20% e a frequência alélica do alelo C é de 80% (Figura 4.9. V), observando-se um valor de MAF superior em relação aos outros grupos analisados.

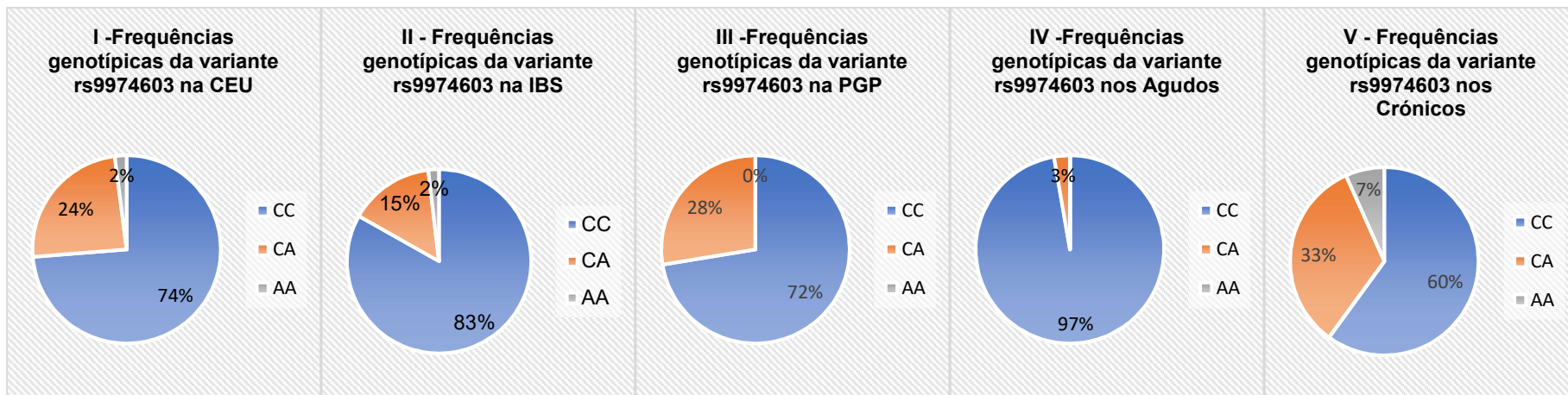


Figura 4. 10. Gráficos *pie chart* relativos às frequências genóticas da variante rs9974603 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

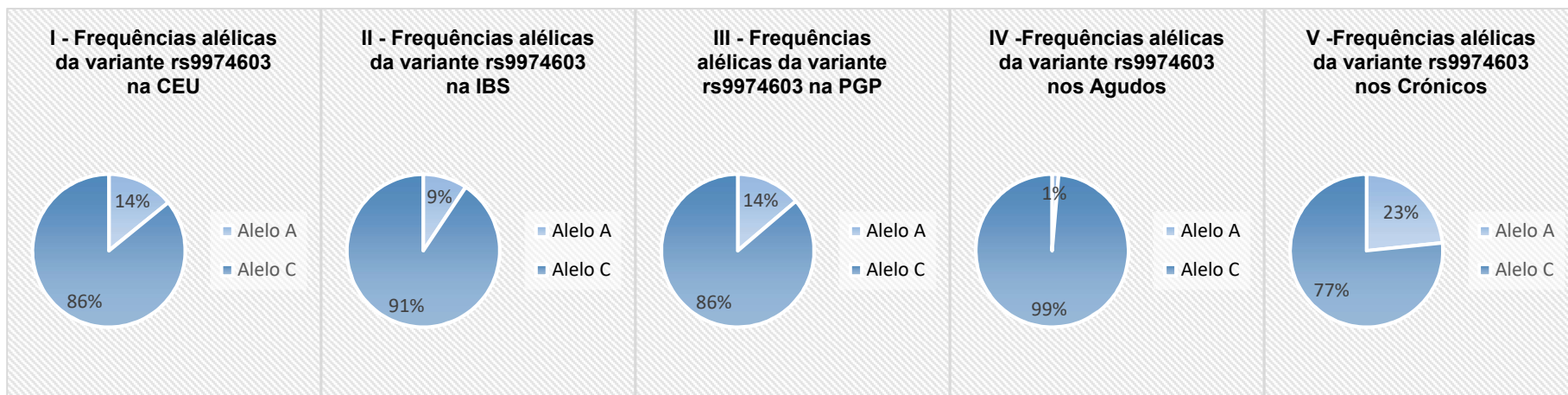


Figura 4. 11. Gráficos *pie chart* relativos às frequências alélicas da variante rs9974603 nas subpopulações CEU e IBS e na população portuguesa analisada no presente estudo.

O SNP rs9974603 é uma variante intrónica cuja distribuição genotípica na população CEU é de 74% de homozigóticos CC, 24% de heterozigóticos CA e 2% de homozigóticos AA (Figura 4.10 I). A frequência alélica do alelo A (minoritário) é de 14% e a frequência alélica do alelo C (maioritário) é de 86% (Figura 4.11 I).

Na subpopulação IBS do Projeto 1000 Genomas, esta variante apresenta uma distribuição genotípica distinta, sendo de, aproximadamente, 83% de homozigóticos CC, 15% de heterozigóticos CA e de 2% de homozigóticos AA (Figura 4.10.II). A frequências do alelo A (minoritário) é de 9% e a frequência alélica do alelo C (maioritário) é de 91% (Figura 4.11.II).

Na amostra da população geral portuguesa (PGP) verificou-se uma distribuição genotípica diferente da apresentada pelas duas subpopulações do projeto 1000 genomas em análise (CEU e IBS), uma vez que, não foi encontrado nenhum individuo com o genótipo homozigótico AA. Desta forma, a distribuição dos genótipos foi de 72% de homozigóticos CC e 28% de heterozigóticos CA (Figura 4.10 III). Quanto às frequências alélicas dos alelos minoritário e maioritário verificou-se que estas são iguais às registadas na subpopulação CEU, ou seja, 14% para o alelo A e 86% do alelo C (Figura 4.11. III).

Nos doentes agudos, a variante em análise, exibiu uma distribuição genotípica de 97% de homozigóticos CC% e de 3% de heterozigóticos CA (Figura 4.10. IV). Neste grupo, tal como na PGP, não verificámos a existência de indivíduos com o genótipo homozigótico AA (Figura 4.10. IV). A frequências do alelo A (minoritário) é de 1% e a frequência alélica do alelo C (maioritário) é de 99% (Figura 4.11.IV).

Nos doentes crónicos a distribuição genotípica apresentada é de, aproximadamente, 60% de homozigóticos CC, 33% de heterozigóticos CA e de 7% de homozigóticos AA, ou seja, a percentagem de heterozigóticos e a percentagem de homozigóticos AA são superiores às apresentadas nos restantes grupos (Figura 4.10. V). A frequência alélica do alelo A é de 23% e a frequência alélica do alelo C é de 77% (Figura 4.11. V), observando-se um valor de MAF superior em relação aos outros grupos analisados.

As variantes rs78407108, rs17879956, rs7277167 e rs9974603 não mostraram ter nenhum significado clínico na pesquisa na base de dados *ClinVar* (<https://www.ncbi.nlm.nih.gov/clinvar/>) (base de dados sobre a relação entre variantes genéticas e a saúde humana) nem estão referidas em literatura relevante para o nosso

estudo. Contudo, cada uma dessas variantes pode estar em desequilíbrio de ligação com outras variantes que tenham significado clínico. Recorrendo ao *Ensemble Genome Browser 96* (Ensembl, 2019) (<https://www.ensembl.org/index.html>) e, através do software *Haploview*, pesquisamos, na população CEU do Projeto 1000 genomas, a existência de variantes que pudessem estar em desequilíbrio de ligação com cada uma das quatro variantes com associação significativa à febre Q crônica. A sequência alvo compreendeu a região 21:33391000 a 21:33450000 onde se encontra o gene *IFNGR2*, de modo a englobar toda a zona sequenciada por NGS. Devido às limitações de alcance do software não foi possível gerar figuras de correlação entre as nossas quatro variantes e todas as variantes com as quais estão em desequilíbrio de ligação. Para gerar imagens que pudessem exemplificar essa correlação restringiu-se a sequência alvo da análise apenas à região que se encontra entre 1000pb a montante e 1000pb a jusante de cada SNP. Nesta região, não foram encontradas variantes em correlação com as variantes rs78407108, rs17879956, rs7277167. As variantes identificadas como estando em desequilíbrio de ligação com a variante rs9974603 na região 1000pb a montante e 1000pb a jusante do SNP, encontram-se identificadas na Figura 4.12.

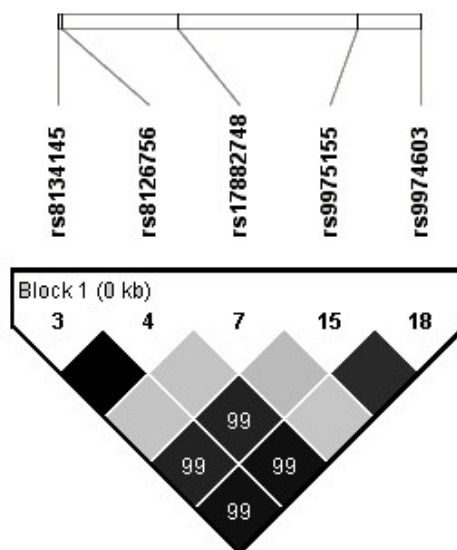


Figura 4. 12. Padrões de desequilíbrio de ligação na região entre 1000pb a montante e 1000pb a jusante da variante rs9974603 de acordo com o HapMap, visualizado usando o software Haploview (Barret et al., 2005) na subpopulação CEU.

O coeficiente de correlação (r^2) entre a variante rs9974603 e a variante rs9975155 é de 0,83, o valor de r^2 entre a variante rs9974603 e a variante rs8126756 e entre a variante rs9974603 e a variante rs8134145 é de 0,99.

As variantes com que mostraram estar em desequilíbrio de ligação com os SNPs rs78407108, rs17879956, rs7277167 e rs9974603 na região 21:33391000 a 21:33450000 encontram-se descritas na tabela 4.3.

Tabela 4. 3. Variantes em desequilíbrio de ligação ($R^2 \geq 0,8$) com os SNPs que mostraram uma relação estatisticamente significativa, resistindo ao critério de Bonferroni, em associação com a febre Q crónica.

Variante	Tipo	Alelos	Gene: Consequência ¹	HGVS ²
rs7277167 ³	SNP	C>A / C>T	IFNGR2: variante 2kb a montante	NC_000021.9:g.33401471C>A / NC_000021.9:g.33401471C>T
rs17880053	Indel	dupG	IFNGR2: variante 2kb a montante	NC_000021.9:g.33401192dup
rs8126756	SNP	T>C / T>G	IFNGR2: variante na região 5' UTR	NC_000021.9:g.33403138T>C / NC_000021.9:g.33403138T>G
rs8134145	SNP	A>C / A>G	IFNGR2: variante na região 5' UTR	NC_000021.9:g.33403128A>C / NC_000021.9:g.33403128A>G
rs9974603 ³	SNP	C>A / C>G	IFNGR2: variante intrónica	NC_000021.9:g.33403994C>A / NC_000021.9:g.33403994C>G
rs78407108 ³	SNP	G>A / G>T	IFNGR2: variante intrónica	NC_000021.9:g.33426675G>A / NC_000021.9:g.33426675G>T
rs2268241	SNP	G>A	IFNGR2: variante intrónica	NC_000021.9:g.33408744G>A
rs2268242	SNP	C>T	IFNGR2: I variante intrónica	NC_000021.9:g.33408798C>T
rs2300373	SNP	A>C / A>G / A>T	IFNGR2: variante intrónica	NC_000021.9:g.33416432A>C / NC_000021.9:g.33416432A>G / NC_000021.9:g.33416432A>T
rs28653198	SNP	A>C	IFNGR2: variante intrónica	NC_000021.9:g.33410089A>C
rs7281873	SNP	A>C / A>G / A>T	IFNGR2: variante intrónica	NC_000021.9:g.33408283A>C / NC_000021.9:g.33408283A>G / NC_000021.9:g.33408283A>T
rs8126735	SNP	G>A / G>C / G>T	IFNGR2: variante intrónica	NC_000021.9:g.33427752G>A / NC_000021.9:g.33427752G>C / NC_000021.9:g.33427752G>T
rs8133431	SNP	C>T	IFNGR2: variante intrónica	NC_000021.9:g.33424249C>T
rs9975155	SNP	T>C / T>G	IFNGR2: variante intrónica	NC_000021.9:g.33403846T>C / NC_000021.9:g.33403846T>G
rs9975473	SNP	T>A / T>G	IFNGR2: variante intrónica	NC_000021.9:g.33404135T>A / NC_000021.9:g.33404135T>G
rs9808753	SNP	A>G	IFNGR2: variante <i>missense</i>	NC_000021.9:g.33415005A>G
rs11088251	SNP	A>C	Nenhuma	NC_000021.9:g.33398504A>C
rs4571135	SNP	A>G	Nenhuma	NC_000021.9:g.33395308A>G
rs7283827	SNP	A>C / A>G / A>T	Nenhuma	NC_000021.9:g.33397598A>C / NC_000021.9:g.33397598A>G / NC_000021.9:g.33397598A>T
rs76661693	Inserção	insA	Nenhuma	NC_000021.9:g.33399283_33399284insA
rs9978166	SNP	A>G / A>T	Nenhuma	NC_000021.9:g.33398210A>G / NC_000021.9:g.33398210A>T
rs9978223	SNP	G>A / G>C	Nenhuma	NC_000021.9:g.33398052G>A / NC_000021.9:g.33398052G>C
rs17879956 ³	SNP	C>T	<i>TMEM50B</i> : variante intrónica	NC_000021.9:g.33438631C>T
rs73195835	SNP	C>G / C>T	<i>TMEM50B</i> : variante intrónica	NC_000021.9:g.33447050C>G / NC_000021.9:g.33447050C>T
rs8129962	SNP	G>A / G>C	<i>TMEM50B</i> : variante intrónica	NC_000021.9:g.33440864G>A / NC_000021.9:g.33440864G>C
rs9974551	SNP	C>G / C>T	<i>TMEM50B</i> : variante intrónica	NC_000021.9:g.33442409C>G / NC_000021.9:g.33442409C>T

Legenda: ¹ De acordo com dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>); ² Nomenclatura das variantes genéticas de acordo com a *Human Genome Variation Society* (<https://www.hgvs.org/>);

³ Variantes identificadas através de análise de associação como associadas à febre Q crónica e apresentando uma relação estatisticamente significativa resistindo ao critério de Bonferroni.

4.2. Resultados do Estudo Réplica

Como foi referido anteriormente, para o estudo réplica foram escolhidas 4 variantes referidas na literatura como estando associadas à febre Q na população holandesa. As variantes rs1718119 e rs3751143 do gene *P2X7R* e as variantes rs2430561 e rs1861493 do gene *INFG* foram sujeitas a amplificação por PCR, eletroforese em gel de agarose para confirmar essa amplificação e posteriormente sequenciadas através de sequenciação Sanger.

O DNA pertencente aos doentes com febre Q crónica (casos) e febre Q aguda (controlos), objeto do estudo exploratório, foi igualmente utilizado no estudo réplica. No entanto, algum DNA encontrava-se degradado/fragmentado e/ou sem a quantidade e qualidade necessárias para uma sequenciação Sanger eficiente. O número de amostras sequenciadas através de sequenciação Sanger é mostrado na tabela 4.4.

Tabela 4. 4. Número de amostras sequenciadas pela Sequenciação de Sanger

Gene	SNP	Nº de crónicos	Nº de agudos	Nº total
<i>P2X7R</i>	rs1718119	14	36	50
	rs3751143	13	29	42
<i>INFG</i>	rs2430561	8	8	16
	rs1861493	7	12	19

As reações de amplificação dos fragmentos em análise, executadas com recurso aos pares de *primers* indicados na tabela 3.6., foram verificadas através de eletroforese em gel de agarose. A figura 4.13. mostra o resultado dos géis de eletroforese dos fragmentos amplificados para cada uma das amostras do grupo de doentes com febre Q crónica (casos). Ao efetuar-se a amplificação das variantes do gene *P2X7R* rs1718119 e rs3751143, obtiveram-se, fragmentos de 511pb (Figura 4.13.I) e 565 pb (Figura 4.13. II), respetivamente. Ao efetuar-se a amplificação das variantes do gene *INFG* rs2430561 e rs1861493, obtiveram-se, fragmentos de 468pb (Figura 4.13. III) e 646pb ((Figura 4.13. IV), respetivamente.

Nos géis apresentados na Figura 4.13. observou-se que para a amostra C11 não se verificou a amplificação de nenhum dos fragmentos em análise. Esta amostra foi, portanto, descartada, assim como todas as amostras (casos e controlos) em que tal se verificou.

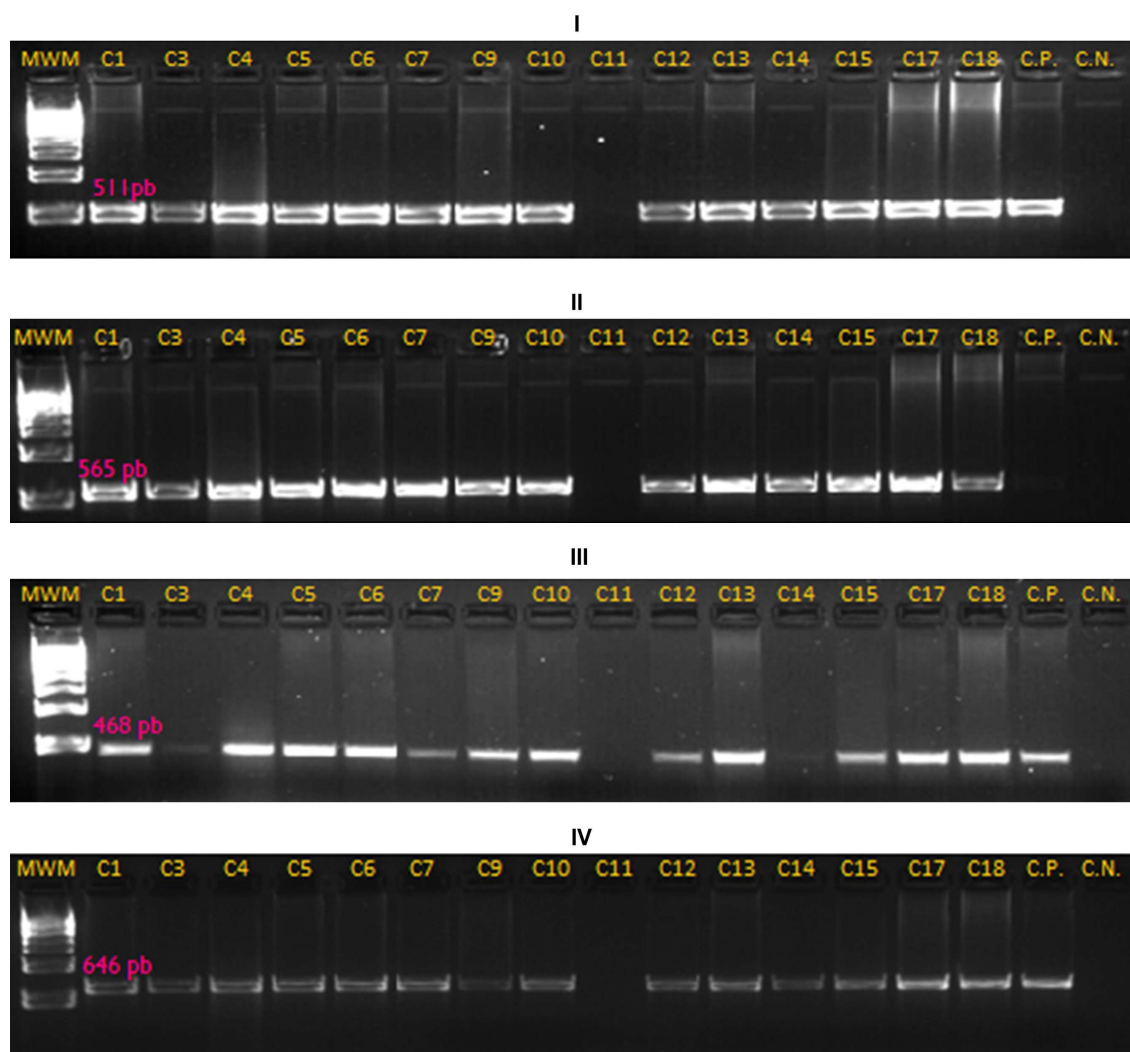


Figura 4. 13. Resultado eletroforese dos fragmentos que contêm as variantes do gene *P2X7R* (rs1718119 e rs3751143) e do gene *INFG* (rs2430561 e rs1861493) para as amostras de doentes com febre Q crónica.

Legenda: I – Gel relativo à amplificação do fragmento de 511pb que contém a variante rs1718119 ; II – Gel relativo à amplificação dos fragmento de 565 pb que contém a varainte rs3751143; III – Gel relativo à amplificação dos fragmento de 468 pb que contém a variante rs2430561 ; IV – Gel relativo à amplificação dos fragmento de 646pb que contém a varainte rs1861493; MWM - Marcador de peso molecular, Quick-Load® 1 Kb DNA Ladder da *New England BioLabs*®Inc (Anexo 3); C1 – amostra doente crónico nº1; C3 – amostra doente crónico nº3; C4 – amostra doente crónico nº4; C5 – amostra doente crónico nº5; C6 – amostra doente crónico nº6; C7 – amostra doente crónico nº7; C9 – amostra doente crónico nº9; C10 – amostra doente crónico nº10; C11 – amostra doente crónico nº11; C12 – amostra doente crónico nº12; C13 – amostra doente crónico nº13; C14 – amostra doente crónico nº14; C15 – amostra doente crónico nº15; C17 – amostra doente crónico nº17; C18 – amostra doente crónico nº18; C.P – controlo positivo; C.N.–controlo negativo.

A sequenciação Sanger por eletroforese capilar de cada um dos fragmentos de DNA amplificados originou eletroferogramas característicos dos diferentes genótipos apresentados pelas amostras (casos e cotrolos).

O ampliação de 511 pb pode apresentar no local da variante rs1718119, o alelo G (alelo ancestral / maioritário) ou o alelo A (alelo minoritário). Desta forma, para esta variante os genótipos são caracterizados como GG, GA e AA conforme os alelos herdados da linha materna e paterna (Figura 4.14.).

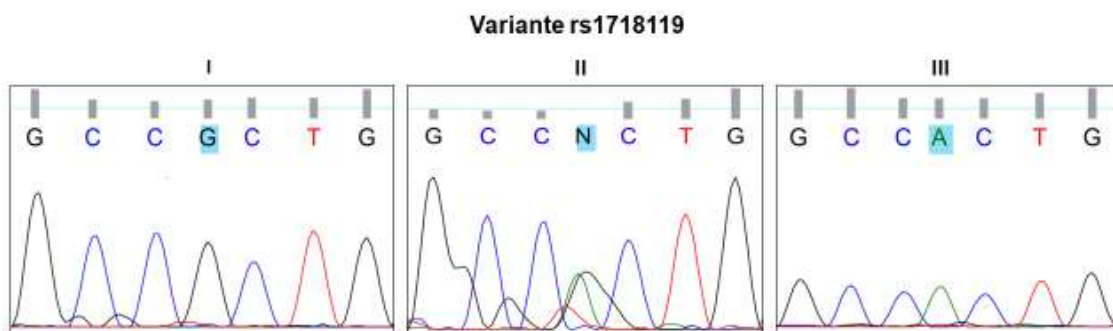


Figura 4. 14. Eletroferograma parcial do gene *P2X7R* mostrando a variante rs1718119 assinalada a azul para os genótipos GG (I), GA (II) e AA (III). N= alelos G e A

O ampliação de 565 pb, pode apresentar no local da variante rs3751143, o alelo A (alelo ancestral / maioritário) ou o alelo C (alelo minoritário). Desta forma, para esta variante os genótipos são caracterizados como AA, AC e CC conforme os alelos herdados da linha materna e paterna (Figura 4.15.).

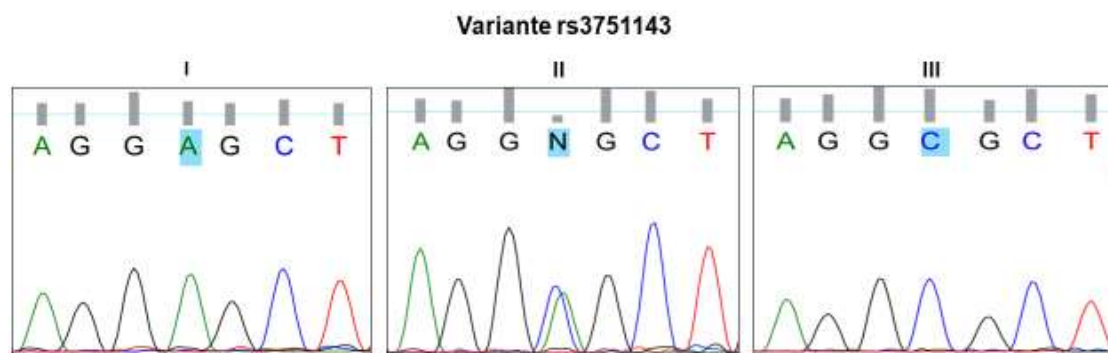


Figura 4. 15. Eletroferograma parcial do gene *P2X7R* mostrando a variante rs3751143 assinalada a azul para os genótipos AA (I), AC (II) e CC (III). N= alelos A e C

O ampliação de 468 pb, pode apresentar no local da variante rs2430561, o alelo T (alelo ancestral / maioritário) ou o alelo A (alelo minoritário). Desta forma, para esta variante os genótipos são caracterizados como TT, TA e AA conforme os alelos herdados da linha materna e paterna. Todavia, no presente estudo, todas as amostras (casos e controles) cuja sequenciação foi bem-sucedida, possuíam o genótipo AA, conforme o apresentado na Figura 4.16.).

Tabela 4. 5. Associação da variante rs1718119 com o aumento do risco de febre Q crônica sob diferentes modelos de hereditariedade

Variante rs1718119							
Modelo de Hereditariedade	Tabela de Contingência	Genótipos			P (χ^2 de Pearson)	P (teste de Fisher) ¹	OR ¹
NA	Crônicos Agudos	GG	GA	AA	0,62261 ²	–	–
		10	3	1			
	Crônicos Agudos	21	14	1	0,6315 ¹	0,7868	0,7609 (0,2494 - 2,3212)
		G	A				
Contagem Alélica	Crônicos Agudos	23	5		0,3929 ¹	0,5218	0,5600 (0,1473-2,1291)
		56	16				
Modelo Dominante	Crônicos Agudos	GG	GA+AA		–	0,0000	2,6923 (0,1567-46,2667)
		10	4				
	Crônicos Agudos	21	15		–	0,3274	0,4286 (0,1014-1,8121)
		GA+GG	AA				
Modelo Recessivo	Crônicos Agudos	13	1		–	0,0000	0,4286 (0,1014-1,8121)
		35	1				
Modelo Sobredominante	Crônicos Agudos	AA+GG	AG		–	0,3274	0,4286 (0,1014-1,8121)
		11	3				
	Crônicos Agudos	22	14		–	0,3274	0,4286 (0,1014-1,8121)

¹ Calculado através do VassarStats: (Web Site for Statistical Computation <http://vassarstats.net/index.html>)

² Calculado através do "Cochran - Armitage trend test calculator" <https://anikhachatryan.github.io/CA-trend-test-calculator/>

Tabela 4. 6. Associação da variante rs3751143 com o aumento do risco de febre Q crônica sob diferentes modelos de hereditariedade

Variante rs3751143							
Modelo de Hereditariedade	Tabela de Contingência	Genótipos			P (χ^2 de Pearson)	P (teste de Fisher) ¹	OR ¹
—	Crônicos Agudos	AA	AC	CC	0,5077 ²	–	–
		7	5	1			
	Crônicos Agudos	13	12	4	0,4930 ¹	0,6160	0,7000 (0,2520 - 1,9448)
		A	C				
Contagem Alélica	Crônicos Agudos	19	7		0,5902 ¹	0,7410	0,6964 (0,1873-2,5895)
		38	20				
Modelo Dominante	Crônicos Agudos	AA	AC+CC		–	1,0000	0,5208 (0,5208- 5,1789)
		7	6				
	Crônicos Agudos	13	16		–	1,0000	0,8854 (0,2319 - 3,3802)
		AC+AA	CC				
Modelo Recessivo	Crônicos Agudos	12	1		0,8625 ¹	1,0000	0,8854 (0,2319 - 3,3802)
		25	4				
Modelo Sobredominante	Crônicos Agudos	AA+CC	AC		0,8625 ¹	1,0000	0,8854 (0,2319 - 3,3802)
		8	5				
	Crônicos Agudos	17	12		0,8625 ¹	1,0000	0,8854 (0,2319 - 3,3802)

¹ Calculado através do VassarStats: (Web Site for Statistical Computation <http://vassarstats.net/index.html>)

² Calculado através do "Cochran - Armitage trend test calculator" <https://anikhachatryan.github.io/CA-trend-test-calculator/>

TABELA 4. 7. Associação da variante rs2430561 com o aumento do risco de febre Q crônica sob diferentes modelos de hereditariedade

		Variante rs2430561					
Modelo de Hereditariedade	Tabela de Contingência	Genótipos			P (χ^2 de Pearson)	P (teste de Fisher) ¹	OR ¹
Modelo Aditivo	Crônicos	TT	TA	AA	NaN ²	-	-
	Agudos	0	0	8			
Contagem Alélica	Crônicos	T	A		-	1,0000	NaN
	Agudos	0	16				
Modelo Dominante	Crônicos	TT	TA+AA		-	1,0000	NaN (NaN-NaN)
	Agudos	0	8				
Modelo Recessivo	Crônicos	TA+TT	AA		-	1,0000	NaN (NaN-NaN)
	Agudos	0	8				
Modelo Sobredominante	Crônicos	AA+TT	TA		-	1,0000	NaN (NaN-NaN)
	Agudos	8	0				

¹ Calculado através do VassarStats: (Web Site for Statistical Computation <http://vassarstats.net/index.html>)

² Calculado através do "Cochran - Armitage trend test calculator" <https://anikhachatryan.github.io/CA-trend-test-calculator/>

Tabela 4. 8. Associação das variantes rs1861493 com o aumento do risco de febre Q crônica sob diferentes modelos de hereditariedade

		Variante rs1861493					
Modelo de Hereditariedade	Tabela de Contingência	Genótipos			P (χ^2 de Pearson)	P (teste de Fisher) ¹	OR ¹
Modelo Aditivo	Crônicos	AA	AG	GG	0,3100 ²	-	-
	Agudos	5	0	2			
Contagem alélica	Crônicos	A	G		-	1	1,4000 (0,2009 - 9,7544)
	Agudos	10	2				
Modelo Dominante	Crônicos	AA	AG+GG		-	1	1,2000 (0,1474 - 9,7679)
	Agudos	5	2				
Modelo Recessivo	Crônicos	AG+GG	GG		-	NaN	NaN
	Agudos	2	2				
Modelo Sobredominante	Crônicos	AA+GG	AG		-	0,2632	0 (0- NaN)
	Agudos	7	0				

¹ Calculado através do VassarStats: (Web Site for Statistical Computation <http://vassarstats.net/index.html>)

² Calculado através do "Cochran - Armitage trend test calculator" <https://anikhachatryan.github.io/CA-trend-test-calculator/>

Capítulo V – Discussão

No presente estudo, procurou-se identificar, na população portuguesa, os determinantes genéticos da infeção na febre Q crónica, ou seja, as variantes genéticas dos principais genes envolvidos na resposta imune que pudessem estar associadas a um desenvolvimento de febre Q crónica no caso de infeção pelo agente patogénico *C. burnetii*. Com o intuito de verificar a frequência com que ocorrem certas variantes genética nos diferentes grupos, recorreremos à análise retrospectiva caso-controlo usando métodos de associação genótipo-fenótipo com base em genes candidatos (*Candidate gene Association Studies – CGA*) através de duas abordagens distintas: um estudo exploratório, no qual se procurou identificar novas variantes associadas a fenótipos clínicos da doença, e um estudo réplica com vista a confirmar a associação de variantes previamente identificadas.

Uma resposta imunitária deficitária poderá contribuir para o desenvolvimento da febre Q crónica (Jansen *et al.*, 2019). Desta forma, as variações genéticas em genes das vias de resposta imunitária poderão ser determinantes na progressão para o fenótipo clínico crónico desta doença.

De entre os genes candidatos analisados no estudo exploratório por NGS (*IL12RB1*, *IFNGR1*, *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* *TIMP1*), SNPs com significância estatística resistindo aos critérios de Bonferroni para testes múltiplos foram identificados apenas na sequência do gene *IFNGR2*. Dos quatro SNPs identificados, três eram variantes intrónicas, rs78407108, rs17879956, rs9974603, e a quarta uma variante 2kb a montante do gene, rs7277167 (<https://www.ncbi.nlm.nih.gov/snp/>). Na variante intrónica *IFNGR2* rs78407108, o alelo A, associado a casos de febre Q crónica, é um alelo raro (frequência < 0,05). O alelo T da variante *IFNGR2* rs17879956 foi associado à febre Q crónica. Esta é uma variante intrónica do gene *TMEM50B* (gene da proteína transmembranar 50B) comum na população global. O gene *TMEM50B* é um gene reverso que está parcialmente sobreposto com o gene *IFNGR2* que se encontra na cadeia *forward*, ou seja, estes dois genes são transcritos em direções opostas no mesmo locus genómico. O alelo A do SNP rs9974603, associado à febre Q crónica, é uma variante intrónica do gene *IFNGR2* comum na população global. O alelo T do SNP rs7277167, associado à febre Q crónica é uma variante 2kb a montante de *IFNGR2* comum na população global. A variante é um eQTL (*expression quantitative trait loci*) onde o alelo T foi associado à maior expressão do gene *IFNGR2* e menor expressão do *TMEM50B*, no sistema circulatório (artéria coronária, sangue total) (<https://qtexportal.org/home/>). O alelo A do SNP rs9974603, associado à febre Q

crónica, é uma variante intrónica do gene *IFNGR2* comum na população global. Para esta variante foi calculado um valor de *odds ratio* (OR) de 22,22 (95 % IC 2,60 – 190,18; p=0,0006). Para um valor de OR > 1, o alelo minoritário (o alelo A, neste caso) estará associado a uma maior probabilidade de desenvolvimento de febre Q crónica. Indivíduos que possuam no seu genótipo o alelo A terão uma maior probabilidade de desenvolvimento de febre Q crónica no caso de exposição à *C. burnetii*. O OR sendo de 22,22 significa que o risco de ocorrência de febre Q crónica nos indivíduos que possuem o alelo A é 22,22 vezes superior em relação aos indivíduos que não possuem este alelo. O risco de desenvolver febre Q crónica é 22,22 vezes superior na presença do alelo A.

Para as variantes associadas, observámos, na população portuguesa, uma frequência do alelo minoritário mais elevada do que a relatada para as subpopulações CEU e IBS do Projeto 1000 genomas para o conjunto das variantes rs78407108, rs17879956, rs7277167 e rs9974603. Este resultado vai ao encontro das nossas observações em estudos anteriores onde se verificou que a heterogeneidade genética na população portuguesa é superior à da população europeia quando representada pela subpopulação CEU (David *et al.*, 2018 a e b). No presente estudo confirmamos esta tendência também relativamente à população de Espanha aqui representada pela subpopulação IBS. Salienta-se uma frequência do alelo minoritário no grupo controlo (doentes com febre Q aguda) inferior à observada na população geral portuguesa, em oposição a uma frequência deste alelo superior à da população geral portuguesa no grupo dos casos (doentes com febre Q crónica). O enriquecimento do alelo minoritário na população com febre Q crónica em relação à população geral é um resultado que vai ao encontro da associação do alelo minoritário nestas variantes com o fenótipo clinicamente grave da febre Q.

No nosso conhecimento, os SNPs rs78407108, rs17879956, rs7277167 e rs9974603 não foram citados anteriormente na literatura internacional, nem relatados no ClinVar (<https://www.ncbi.nlm.nih.gov/snp/>). Portanto, para explorar esses tagSNPs investigamos toda a região da sequência *IFNGR2*, entre as posições 33390000-33450000, para padrões de desequilíbrio de ligação (LD) na subpopulação com ascendência do norte e oeste da Europa (CEU) usando o Haploview (Barret *et al.* 2005). As variantes em LD com as tagSNPs foram identificadas como variantes dos genes *IFNGR2* e/ou *TMEM50B*. Nenhuma foi referida na base de dados ClinVar, no entanto, cinco dessas variantes foram citadas na literatura científica: rs9808753 uma variante missense do gene *IFNGR2* (também referida como Q64R), o indel rs17880053, uma variante 2kb a montante do gene *IFNGR2*, os SNPs rs8126756 e rs8134145, duas

variantes 5'UTR do gene *IFNGR2* e rs9978223, uma variante não classificada. Pesquisamos as variantes rs17880053, rs8126756, rs8134145 no portal *Genotype-Tissue Expression* (GTEx) (<https://gtexportal.org/home/>). Nem a rs8126756 nem a rs8134145 foram referenciadas no portal GTEx. No entanto, o alelo alterado de rs17880053 parece estar associado à maior expressão do *IFNGR2* e menor expressão dos genes *TMEM50B*.

De acordo com a informação do *dbSNP* (<https://www.ncbi.nlm.nih.gov/snp/>), a variante rs9808753 é uma variante benigna, que apresenta os alelos A>G na posição 21:33415005. Para a substituição de aminoácidos Q64R da variante rs9808753, o efeito previsto da alteração na função da proteína foi classificado como “tolerado” por *SIFT* e “benigno” por *PolyPhen-2*, duas ferramentas do *Ensembl Genome Browser* que preveem o efeito de uma substituição de aminoácidos na estrutura e função da proteína (https://www.ensembl.org/info/genome/variation/prediction/protein_function.html). Esta variante é referida em 31 publicações na base de dados *LitVar* (<https://www.ncbi.nlm.nih.gov/CBBresearch/Lu/Demo/LitVar/#!?query=>). Embora o SNP rs9808753 leve a uma substituição de aminoácidos da glutamina por arginina (Q64R ou Gln64Arg) no recetor *IFNGR2*, atualmente não está claro se essa alteração interfere na capacidade de ligação com *IFN γ* ou alterara outras funções biológicas (Chen *et al.*, 2011a). De Paus e colaboradores consideraram esta variante um polimorfismo funcional (De Paus *et al.* 2011). Outros estudos demonstraram a associação desta variante a doenças relacionadas com o sistema imunitário (Xiang, *et al.*, 2014). Por exemplo, o aminoácido R64 foi referido como estando correlacionado com um início progressivo de esclerose múltipla (Schrijver *et al.*, 2004), enquanto que o aminoácido Q64 foi descrito como um fator de risco para lúpus eritematoso sistémico (Nakashima *et al.*, 1999), ambas as condições associadas a alterações no equilíbrio imunológico (Xiang, *et al.*, 2014). Este SNP foi, ainda, associado com a virémia da Hepatite B após correção para testes múltiplos ($p < 0,02$) (Huang *et al.*, 2010). Chen e colaboradores relataram que a variante *IFNGR2* rs9808753 alterou a relação do índice de massa corporal (IMC) e o risco de linfoma non-Hodgkin (LNH) para linfoma de células B e linfoma de células T, sugerindo que a via de transdução de *IFN γ* pode desempenhar um papel na relação entre IMC e risco de LNH (Chen *et al.*, 2011b).

De acordo com a informação do *dbSNP* (<https://www.ncbi.nlm.nih.gov/snp/>), a variante rs17880053 é um indel 2Kb a montante do gene *IFNGR2*, localizada na posição 21: 33401192 onde ocorre uma duplicação do alelo G (dupG). Esta variante é referida em 3 citações no *LitVar*, contudo, até agora, não foi associada a nenhuma doença. Na

nossa consulta da base de dados GTEX, o alelo G duplicado da variante rs17880053, associado à febre Q crónica no nosso estudo, parece induzir uma expressão maior do gene *IFNGR2*. Este resultado, ao ser confirmado, poderá ir ao encontro às observações de Schoffelen e colaboradores que mostraram que a produção e a resposta mediada por IFN γ estão intactas na febre Q crónica e, desta forma, uma deficiência nas respostas de IFN γ não explica a falha na eliminação da infeção e o desenvolvimento de febre Q crónica, mas, ao contrário ao que acontece com a tuberculose, observa-se na febre Q crónica, uma sobre-expressão do IFN γ associada à doença grave (Schoffelen *et al.* 2017). Os indivíduos com endocardite e doença vascular devida a febre Q exibem uma elevada produção de IFN γ no sangue total específica para *C. burnetii*, enquanto anteriormente era assumido que a cronicidade da febre Q era devida à falta de resposta dos linfócitos T e produção prejudicada de IFN γ (Ferwerda *et al.*, 2008; Q-VAX 2008; Schoffelen *et al.*, 2017).

De acordo com a informação do *dbSNP* (<https://www.ncbi.nlm.nih.gov/snp/>), a variante rs9978223 é uma variante, até à data, sem classificação, que apresenta os alelos G>A/G>C na posição 21:33398052. Esta variante foi citada apenas uma vez na base de dados *LitVar*. A variante rs9978223 mostrou uma associação significativa ($p=0,00021$) com o enfarte agudo do miocárdio, na população indiana, abaixo do valor p corrigido de Bonferroni ($p = 0,00061$) (Shalia *et al.*, 2017). Apesar de pouco documentado na literatura existem referências que correlacionaram o enfarte do miocárdio à endocardite infecciosa (Hunter *et al.*, 2001; Overend e Rose, 2012; Calero-Núñez *et al.*, 2018). Como a endocardite é a principal consequência da febre Q crónica, a hipótese levantada da implicação indireta da variante rs9978223 com este fenótipo merece ser aprofundada.

Segundo a informação do *dbSNP* (<https://www.ncbi.nlm.nih.gov/snp/>), a variante rs8126756, uma variante 5'UTR do gene *IFNGR2*, está localizada na posição 21:33403138, apresentando os alelos T>C,G. Esta variante aparece citada em 4 publicações no *LitVar*. Em haplótipo, a variante rs8126756 foi responsável por uma elevada expressão do gene *IFNGR2 in vitro* associada a resistência à tuberculose na população Vietnamita (Hijikata *et al.*, 2012). A variante genotípica homozigótica *IFNGR2* T>C (rs8126756) foi associada com a não resposta entre os pacientes com doença de Cochr à terapia anti-TNF (OR: 0.09, 95% CI: 0.01–0.65, $P = 0.02$) (Bank *et al.*, 2018).

Segundo a informação do *dbSNP* (<https://www.ncbi.nlm.nih.gov/snp/>), a variante rs8134145, uma variante 5'UTR do gene *IFNGR2*, está localizada na posição

21:33403128, apresentando os alelos A>C,G. É citada na plataforma LitVar em 3 artigos, contudo não foi encontrada referenciada como estando associada a qualquer doença.

Em relação aos estudos de replica para as variantes rs1861493 e rs2430561 do gene *IFNG* e rs1718119 e rs3751143 do gene *P2RX7* não foi identificada uma relação de associação genótipo – fenótipo com a febre Q estaticamente significativa, contrariamente aos resultados de Wielders e colaboradores, (Wielders *et al.*, 2015) e Jansen e colaboradores (Jansen *et al.*, 2019), respetivamente. Este resultado poderá ser devido a uma redução do número de amostras analisadas pela baixa qualidade do DNA por degradação do mesmo visto este estudo ter sido realizado meses depois do estudo exploratório inicial.

Uma das críticas do nosso trabalho é o facto da nossa amostragem ter sido bastante reduzida, resultando num poder estatístico menor e elevado risco de falsos positivos. Devido a este constrangimento, foi seleccionada uma estratégia no controlo de qualidade dos dados NGS para estudos de associação genótipo - fenótipo que nos pareceu ser a mais adequada ao nosso trabalho por separar as variantes com uma frequência do alelo minoritário menor (MAF) < 5% dos > 5%. Isto permitiu aplicar um critério de *missing genotypes* mais restrigente para as variantes raras (remoção de amostras com mais de 1% de genótipos ausentes) do que para as variantes mais comuns (remoção de amostras com mais de 5% de genótipos ausentes) (McDonald *et al.*, 2018 (<https://statgen.us/Tutorials#GWAS: Data Quality Control>)). De facto, as variantes raras não poderiam estar bem representadas na nossa base de dados devido ao número reduzido de amostras, podendo enviesar os resultados de associação.

Outro aspeto que poderá melhorar com o aumento do número de amostras será a obtenção de um valor do OR referente às variantes rs78407108, rs17879956 e rs7277167, para as quais não existiram na nossa amostra controlo (indivíduos com febre Q aguda), indivíduos heterozigóticos. Dado que, na análise da população geral portuguesa, o alelo minoritário estava representado em heterozigotia, seria de esperar que um aumento do número de amostras contribuisse para a obtenção de um valor do OR a representar o tamanho do efeito da associação.

Finalmente, contrariamente aos estudos já efetuados nesta área, cujos grupos experimentais possuíam fatores de predisposição para o desenvolvimento de febre Q crónica (Schoffelen *et al.*, 2015; Jansen *et al.*, 2017; Jansen *et al.*, 2019;) independente do seu fenótipo clínico, no nosso estudo, apenas os casos, ou seja, o grupo de doentes

com febre Q crónica, possuíam fatores de predisposição. Esta falha poderá introduzir viés nos resultados dados ao facto dos grupos de casos e controlo não serem o mais idênticos possível exceto em relação ao fenótipo da doença tal com preconizado para estes estudos caso controlo (David, 2021).

Capítulo VI – Conclusão

Vários estudos sugerem que a via IL12-IFN γ é crucial para a defesa imunitária contra a *C. burnetii* (Andoh *et al.*, 2007; Capo e Mege, 2012; Wielders *et al.*, 2015; Schoffelen *et al.*, 2017) bem como contra outras bactérias intracelulares como *Mycobacterium spp* (Rossuw *et al.*, 2003; Tso *et al.*, 2005, Bulat-Kardum *et al.*, 2006; Pacheco *et al.*, 2008; Vallinoto *et al.*, 2010; Ghanavi *et al.*, 2021) e vírus, como os vírus da hepatite C (Huang *et al.*, 2007) e B (HBV) (Huang *et al.*, 2014; Heidari *et al.*, 2020). Variantes em genes importantes desta via, nomeadamente, no gene do interferão gama (*IFNG*) e no seu recetor (*IFNGR1*) modulam o progresso de IFN γ na interação patógeno-hospedeiro (Wielders *et al.*, 2015; Schoffelen *et al.*, 2017; Heidari *et al.*, 2020; Ghanavi *et al.*, 2021). Variantes nos genes dos recetores do interferão foram associadas a doenças crónicas (Jülicher *et al.*, 2002; Kerner *et al.*, 2020).

O recetor IFN γ (*IFNGR*) é um recetor heterodimérico composto pelas cadeias *IFNGR-1* (cadeia α) e *IFNGR-2* (cadeia β). A subunidade *IFNGR-1* liga-se ao IFN γ com a maior afinidade. A subunidade *IFNGR-2* não se liga ativamente ao IFN γ ; no entanto, é necessário para a ativação específica de IFN γ da via de transdução de sinal intracelular. O gene *IFNGR2* codifica a cadeia β não ligante do IFN γ localizada no cromossoma 21. A iniciação da cascata de transdução de sinal IFN γ serve para inibir diretamente a replicação do microrganismo infetante e tanto estimula quanto modula o sistema imunológico. Bem como o *IFNGR2* parece estar associado à resposta celular ao IFN γ (Kerner *et al.*, 2020), estudos recentes levantam a hipótese de que possam existir outros agonistas do recetor do IFN γ (Kerner *et al.*, 2020).

Desta forma e de acordo com os nossos resultados, é possível concluir que, de uma forma geral, os objetivos aos quais nos propusemos com a realização da presente dissertação de mestrado foram atingidos. A estratégia adotada, usando métodos de associação genótipo-fenótipo com base em CGAS, bem como a utilização dos métodos de sequenciação NGS e Sanger revelaram-se adequados para a identificação de determinantes genéticos da infeção na febre Q crónica. Conclui-se, ainda, que as variantes genéticas da via IL12-IFN γ , nomeadamente, as variantes do gene *IFNGR2*, podem estar associadas à progressão da febre Q para uma fase crónica em pacientes com fatores de predisposição para o desenvolvimento de febre Q crónica.

Capítulo VII – Perspetivas futuras

Os nossos resultados não apontam para nenhuma explicação clara do papel do gene *IFNGR2* e suas variantes na febre Q crónica. Por esta razão, será imprescindível avançar com estudos funcionais de modo a esclarecer o papel das variantes associadas com a febre Q crónica, seja através de estudos de expressão génica *ex-vivo*, verificando o aumento da expressão dos recetores no caso de variantes associadas nas regiões reguladoras, ou de estudos de sobrevivência e multiplicação bacteriana em modelos de infecção *ex-vivo* previligiando as variantes associadas.

Raramente as variantes são herdadas isoladamente, existindo combinações alélicas que ocorrem com mais frequência do que seria esperado pela viabilidade aleatória (Regateiro, 2007; David, 2021). A associação de haplótipos poderá ser mais marcante do que a associação individual das variantes (David *et al.*, 2018a). Os haplótipos associados poderão também ser objeto de estudos funcionais. Uma abordagem de grande interesse seria a realização de estudos longitudinais com vista ao seguimento de doentes agudos ao longo do tempo para avaliar a propensão ao desenvolvimento da febre Q crónica segundo o genótipo.

Referências bibliográficas:

- Abel, L., El-Baghdadi, J., Bousfiha, A. A., Casanova, J. L., & Schurr, E. (2014). Human genetics of tuberculosis: a long and winding road. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1645), 20130428.
- Abnave, P., Muracciole, X., & Ghigo, E. (2017). Coxiella burnetii lipopolysaccharide: what do we know?. *International journal of molecular sciences*, 18(12), 2509.
- Alcaïs, A., Abel, L., & Casanova, J. L. (2009). Human genetics of infectious diseases: between proof of principle and paradigm. *The Journal of clinical investigation*, 119(9), 2506-2514.
- Alter, A., Grant, A., Abel, L., Alcaïs, A., & Schurr, E. (2011). Leprosy as a genetic disease. *Mammalian genome*, 22(1), 19-31.
- Amitai, Z., Bromberg, M., Bernstein, M., Raveh, D., Keysary, A., David, D., ... Shohat, T. (2010). A Large Q Fever Outbreak in an Urban School in Central Israel. *Clinical Infectious Diseases*, 50(11), 1433–1438. doi:10.1086/652442
- Ammerdorffer, A., Stappers, M. H., Oosting, M., Schoffelen, T., Hagenaars, J. C., Bleeker-Rovers, C. P., ... & Joosten, L. A. (2016). Genetic variation in TLR10 is not associated with chronic Q fever, despite the inhibitory effect of TLR10 on Coxiella burnetii-induced cytokines in vitro. *Cytokine*, 77, 196-202.
- Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature Protocols*, 5(9), 1564–1573. doi:10.1038/nprot.2010.11
- Andoh, M., Zhang, G., Russell-Lodrigue, K. E., Shive, H. R., Weeks, B. R., & Samuel, J. E. (2007). T cells are essential for bacterial clearance, and gamma interferon, tumor necrosis factor alpha, and B cells are crucial for disease development in Coxiella burnetii infection in mice. *Infection and immunity*, 75(7), 3245-3255.
- Angelakis, E., & Raoult, D. (2010). Q fever. *Veterinary microbiology*, 140(3-4), 297-309.
- Angelakis, E., Mediannikov, O., Socolovschi, C., Mouffok, N., Bassene, H., Tall, A., ... & Raoult, D. (2014). Coxiella burnetii-positive PCR in febrile patients in rural and urban Africa. *International Journal of Infectious Diseases*, 28, 107-110.
- Archer, N. S., Nassif, N. T., & O'Brien, B. A. (2015). Genetic variants of SLC11A1 are associated with both autoimmune and infectious diseases: systematic review and meta-

- analysis. *Genes & Immunity*, 16(4), 275-283.
- Arricau-Bouvery, N., & Rodolakis, A. (2005). Is Q fever an emerging or re-emerging zoonosis?. *Veterinary research*, 36(3), 327-349.
- Asner, S. A., Morre, S. A., Bochud, P. Y., & Greub, G. (2014). Host factors and genetic susceptibility to infections due to intracellular bacteria and fastidious organisms. *Clinical Microbiology and Infection*, 20(12), 1246-1253. doi: 10.1111/1469-0691.12806.
- Bank, S., Andersen, P. S., Burisch, J., Pedersen, N., Roug, S., Galsgaard, J., ... & Andersen, V. (2018). Genetically determined high activity of IL-12 and IL-18 in ulcerative colitis and TLR5 in Crohns disease were associated with non-response to anti-TNF therapy. *The pharmacogenomics journal*, 18(1), 87-97.
- Barrett, J. C., Fry, B., Maller, J. D. M. J., & Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2), 263-265.
- Barry, A. O., Mege, J. L., & Ghigo, E. (2011). Hijacked phagosomes and leukocyte activation: an intimate relationship. *Journal of leukocyte biology*, 89(3), 373-382.
- Beckman Coulter. (2002). Agencourt AMPure XP PCR Purification. AMPure XP. Retrieved from <http://autodiscover.josam.com/pdf/josam/catalog/Instructions.pdf>
- Bellamy, R., Ruwende, C., Corrah, T., McAdam, K. P., Whittle, H. C., & Hill, A. V. (1998). Variations in the NRAMP1 gene and susceptibility to tuberculosis in West Africans. *New England Journal of Medicine*, 338(10), 640-644.
- Benoit, M., Barbarat, B., Bernard, A., Olive, D., & Mege, J. L. (2008). *Coxiella burnetii*, the agent of Q fever, stimulates an atypical M2 activation program in human macrophages. *European journal of immunology*, 38(4), 1065-1070.
- Bland JM, Altman DG (1995) Multiple significance tests: the Bonferroni method. *BMJ* 310(6973):170
- Boisson-Dupuis, S., Bustamante, J., El-Baghdadi, J., Camcioglu, Y., Parvaneh, N., El Azbaoui, S., ... & Abel, L. (2015). Inherited and acquired immunodeficiencies underlying tuberculosis in childhood. *Immunological reviews*, 264(1), 103-120.
- Boisson-Dupuis, S., El Baghdadi, J., Parvaneh, N., Bousfiha, A., Bustamante, J., Feinberg, J., ... & Casanova, J. L. (2011). IL-12R β 1 deficiency in two of fifty children with severe tuberculosis from Iran, Morocco, and Turkey. *PloS one*, 6(4), e18524.

- Bossi, P., Van Loock, F., Tegnell, A., & Gouvras, G. (2004). Bichat clinical guidelines for bioterrorist agents. *Eurosurveillance*, 9(12), 19-20.
- Boucherit, N., Barry, A. O., Mottola, G., Trouplin, V., Capo, C., Mege, J. L., & Ghigo, E. (2012). Effects of *Coxiella burnetii* on MAPKinases phosphorylation. *FEMS Immunology & Medical Microbiology*, 64(1), 101-103..
- Brennan, R. E., Russell, K., Zhang, G., & Samuel, J. E. (2004). Both inducible nitric oxide synthase and NADPH oxidase contribute to the control of virulent phase I *Coxiella burnetii* infections. *Infection and immunity*, 72(11), 6666-6675.
- Broad Institute. (2020). Somatic short variant discovery (SNPs + Indels). Retrieved 13 August 2020, from <https://gatk.broadinstitute.org/hc/enus/articles/360035535932-Germlineshort-variant-discovery-SNPs-Indels>
- Buijs, S. B., Jansen, A. F. M., Oosterheert, J. J., Schoffelen, T., Wever, P. C., Hoepelman, A. I. M., ... & Bleeker-Rovers, C. P. (2021). Single nucleotide polymorphism (SNP) rs3751143 in P2RX7 is associated with therapy failure in chronic Q fever while rs7125062 in MMP1 is associated with fewer complications. *Clinical Microbiology and Infection*, 27(5), 786-e1.
- Bulat-Kardum, L., Etokebe, G. E., Knezevic, J., Balen, S., Matakovic-Mileusnic, N., Zaputovic, L., ... & Dembic, Z. (2006). Interferon- γ Receptor-1 Gene Promoter Polymorphisms (G-611A; T-56C) and Susceptibility to Tuberculosis. *Scandinavian journal of immunology*, 63(2), 142-150.
- Burnet, F. M., & Freeman, M. (1937). Experimental studies on the virus of "Q" fever. *Medical Journal of Australia*, 2(8), 299-305.
- Bustamante, J., Boisson-Dupuis, S., Abel, L., & Casanova, J. L. (2014, December). Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN- γ immunity. In *Seminars in immunology* (Vol. 26, No. 6, pp. 454-470). Academic Press.
- Calero-Núñez, S., Ferrer Bleda, V., Corbí-Pascual, M., Córdoba-Soriano, J. G., Fuentes-Manso, R., Tercero-Martínez, A., ... & Barrionuevo Sánchez, M. I. (2018). Myocardial infarction associated with infective endocarditis: a case series. *European Heart Journal-Case Reports*, 2(1), yty032..
- Capo, C., Lindberg, F. P., Meconi, S., Zaffran, Y., Tardei, G., Brown, E. J., ... & Mege, J. L.

- (1999). Subversion of monocyte functions by *Coxiella burnetii*: impairment of the cross-talk between $\alpha\beta 3$ integrin and CR3. *The Journal of Immunology*, 163(11), 6078-6085..
- Capo, C., & Mege, J. L. (2012). Role of innate and adaptive immunity in the control of Q fever. *Coxiella burnetii: Recent Advances and New Perspectives in Research of the Q Fever Bacterium*, 273-286.
- Capo, C., Moynault, A., Collette, Y., Olive, D., Brown, E. J., Raoult, D., & Mege, J. L. (2003). *Coxiella burnetii* avoids macrophage phagocytosis by interfering with spatial distribution of complement receptor 3. *The journal of immunology*, 170(8), 4217-4225.
- Capo, C., Zugun, F., Stein, A., Tardei, G., Lepidi, H., Raoult, D., & Mege, J. L. (1996). Upregulation of tumor necrosis factor alpha and interleukin-1 beta in Q fever endocarditis. *Infection and Immunity*, 64(5), 1638-1642.
- Carcopino X, Raoult D, Bretelle F, On Boubli L, Stein A. Managing Q Fever during Pregnancy: The Benefits of Long-Term Cotrimoxazole Therapy. *Clin Infect Dis*, 2007; 45: 548 555
- Casanova, J. L., & Abel, L. (2002). Genetic dissection of immunity to mycobacteria: the human model. *Annual review of immunology*, 20(1), 581-620.
- Centers for Disease Control and Prevention (CDC), (2011). Emergency Preparedness and Response: Bioterrorism Agents/Diseases. URL <http://www.bt.cdc.gov/agent/agentlist-category.asp>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4(1), s13742-015.
- Chen, C., Dow, C., Wang, P., Sidney, J., Read, A., Harmsen, A., ... & Peters, B. (2011a). Identification of CD4+ T cell epitopes in *C. burnetii* antigens targeted by antibody responses. *PLoS One*, 6(3), e17712.
- Chen, Y., Wang, J., Ge, P., Cao, D., Miao, B., Robertson, I., ... & Guo, A. (2017). Tissue inhibitor of metalloproteinases 1, a novel biomarker of tuberculosis. *Molecular medicine reports*, 15(1), 483-487..
- Chen, Y., Zheng, T., Lan, Q., Foss, F., Kim, C., Chen, X., ... & Zhang, Y. (2011b). Cytokine polymorphisms in Th1/Th2 pathway genes, body mass index, and risk of non-Hodgkin lymphoma. *Blood, The Journal of the American Society of Hematology*, 117(2), 585-590.

- Chmielewski, T., & Tylewska-Wierzbanowska, S. (2013). Q fever outbreaks in Poland during 2005–2011. *Medical science monitor: international medical journal of experimental and clinical research*, 19, 1073..
- Chorev, M., & Carmel, L. (2012). The function of introns. *Frontiers in genetics*, 3, 55..
- Cox, H. R., & Beli, E. J. (1939). The cultivation of *Rickettsia diaporica* in tissue culture and in the tissues of developing chick embryos. *Public Health Reports (1896-1970)*, 2171-2178..
- Cunha, L. D., Ribeiro, J. M., Fernandes, T. D., Massis, L. M., Khoo, C. A., Moffatt, J. H., ... & Zamboni, D. S. (2015). Inhibition of inflammasome activation by *Coxiella burnetii* type IV secretion system effector IcaA. *Nature communications*, 6(1), 1-13.
- David, S. (2020). Métodos para estudos populacionais de associação genótipo-fenótipo com base em genes candidatos, (7), 33–37.
- David, S. (2021). A current guide to candidate gene association studies. *Trends in Genetics*, 37(12), 1056-1059.
- David, S., Aguiar, P., Antunes, L., Dias, A., Morais, A., Sakuntabhai, A., & Lavinha, J. (2018a). Variants in the non-coding region of the TLR2 gene associated with infectious subphenotypes in pediatric sickle cell anemia. *Immunogenetics*, 70(1), 37-51.
- David, S., Correia, V., Antunes, L., Faria, R., Ferrão, J., Faustino, P., ... & Rebelo de Andrade, H. (2018b). Population genetics of IFITM3 in Portugal and Central Africa reveals a potential modifier of influenza severity. *Immunogenetics*, 70(3), 169-177.
- De Paus, R. A., Kilic, S. S., van Dissel, J. T., & van de Vosse, E. (2011). Effect of amino acid substitutions in the human IFN- γ R2 on IFN- γ responsiveness. *Genes & Immunity*, 12(2), 136-144.
- Dellacasagrande, J., Capo, C., Raoult, D., & Mege, J. L. (1999). IFN- γ -mediated control of *Coxiella burnetii* survival in monocytes: the role of cell apoptosis and TNF. *The Journal of Immunology*, 162(4), 2259-2265.
- Dellacasagrande, J., Ghigo, E., Machergui-El, S., Hammami, Toman, R., Raoult, D., ... & Mege, J. L. (2000). $\alpha\beta 3$ integrin and bacterial lipopolysaccharide are involved in *Coxiella burnetii*-stimulated production of tumor necrosis factor by human monocytes. *Infection and immunity*, 68(10), 5673-5678.

- Derrick, E. H. (1937). "Q" fever, a new fever entity: clinical features, diagnosis and laboratory investigation. *Medical Journal of Australia*, 2(8), 281-299..
- DGS (Direc o-Geral da Sa de), (2010). Febre Q, In: Doenas de Declara o Obrigat ria 2004– 2008, Publica o do Minist rio da Sa de, pp. 61–62. Dispon vel em: <http://www.dgs.pt>
- Doherty, T. M., & Andersen, P. (2005). Vaccines for tuberculosis: novel concepts and recent progress. *Clinical microbiology reviews*, 18(4), 687-702.
- Dupuis, G., Peter, O., Peacock, M., Burgdorfer, W., & Haller, E. (1985). Immunoglobulin responses in acute Q fever. *Journal of Clinical Microbiology*, 22(4), 484-487.
- Dupuis G, Petite J, Peter O, & Vouilloz M. (1987). An important outbreak of human Q fever in a Swiss alpine valley. *International journal of epidemiology*, 16(2), 282-287.
- Eldin, C., M lenotte, C., Mediannikov, O., Ghigo, E., Million, M., Edouard, S., ... & Raoult, D. (2017). From Q fever to *Coxiella burnetii* infection: a paradigm change. *Clinical microbiology reviews*, 30(1), 115-190.
- Elliott, A., Schoenlaub, L., Freches, D., Mitchell, W., & Zhang, G. (2015). Neutrophils play an important role in protective immunity against *Coxiella burnetii* infection. *Infection and immunity*, 83(8), 3104-3113.
- Errasti, C., Baranda, M., Almaraz, J., de la Hoz, C., Gutierrez, E., Villate, J., & Pe a, V. (1984). An outbreak of Q fever in the Basque country. *Canadian Medical Association Journal*, 131, 48–49.
- Everett, B., Cameron, B., Li, H., Vollmer-Conna, U., Davenport, T., Hickie, I., ... & Lloyd, A. R. (2007). Polymorphisms in Toll-like receptors-2 and-4 are not associated with disease manifestations in acute Q fever. *Genes & Immunity*, 8(8), 699-702.
- Ferwerda, G., Meyer-Wentrup, F., Kullberg, B. J., Netea, M. G., & Adema, G. J. (2008). Dectin-1 synergizes with TLR2 and TLR4 for cytokine production in human primary monocytes and macrophages. *Cellular microbiology*, 10(10), 2058-2066.
- Flynn, J. L., Chan, J., & Lin, P. L. (2011). Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal immunology*, 4(3), 271-278.
- Fournier, P. E., Marrie, T. J., & Raoult, D. (1998). Diagnosis of Q fever. *Journal of clinical microbiology*, 36(7), 1823-1834..

- Fournier, P. E., & Raoult, D. (1999). Predominant immunoglobulin A response to phase II antigen of *Coxiella burnetii* in acute Q fever. *Clinical Diagnostic Laboratory Immunology*, 6(2), 173-177.
- Gami, A. S., Antonios, V. S., Thompson, R. L., Chaliki, H. P., & Ammash, N. M. (2004, February). Q fever endocarditis in the United States. In *Mayo Clinic Proceedings* (Vol. 79, No. 2, pp. 253-257). Elsevier.
- Gargis AS, Kalman L, Lubin IM. Assuring the Quality of Next-Generation Sequencing in Clinical Microbiology and Public Health Laboratories. *J Clin Microbiol*. 2016 Dec;54(12):2857-2865. doi: 10.1128/JCM.00949-16. Epub 2016 Aug 10. PMID: 27510831; PMCID: PMC5121372.
- Ghanavi, J., Farnia, P., Farnia, P., & Velayati, A. A. (2021). The role of interferon-gamma and interferon-gamma receptor in tuberculosis and nontuberculous mycobacterial infections. *International Journal of Mycobacteriology*, 10(4), 349..
- Ghaoui, H., Bitam, I., Ait-Oudhia, K., Achour, N., SaadDjaballah, A., Saadnia, F.Z., *et al.* (2018) *Coxiella burnetii* infection with women's febrile spontaneous abortion reported in Algiers. *New Microbes New Infect* 26: 8–14
- Ghigo, E., Capo, C., Raoult, D., & Mege, J. L. (2001). Interleukin-10 stimulates *Coxiella burnetii* replication in human monocytes through tumor necrosis factor down-modulation: role in microbicidal defect of Q fever. *Infection and immunity*, 69(4), 2345-2352.
- Ghigo, E., Capo, C., Tung, C. H., Raoult, D., Gorvel, J. P., & Mege, J. L. (2002). *Coxiella burnetii* survival in THP-1 monocytes involves the impairment of phagosome maturation: IFN- γ mediates its restoration and bacterial killing. *The Journal of Immunology*, 169(8), 4488-4495.
- Ghigo, E., Imbert, G., Capo, C., Raoult, D., & MEGE, J. L. (2003). Interleukin-4 induces *Coxiella burnetii* replication in human monocytes but not in macrophages. *Annals of the New York Academy of Sciences*, 990(1), 450-459.
- Ghigo, E., Honstettre, A., Capo, C., Gorvel, J. P., Raoult, D., & Mege, J. L. (2004). Link between impaired maturation of phagosomes and defective *Coxiella burnetii* killing in patients with chronic Q fever. *The Journal of Infectious Diseases*, 190(10), 1767-1772.
- Ghigo, E., Pretat, L., Desnues, B., Capo, C., Raoult, D., & Mege, J. L. (2009). Intracellular life of *coxiella burnetii* in macrophages: An update. *Annals of the New York Academy of*

Sciences, 1166, 55–66. doi:10.1111/j.1749- 6632.2009.04515.x

- Gideon, H. P., & Flynn, J. L. (2011). Latent tuberculosis: what the host “sees”?. *Immunologic research*, 50(2-3), 202-212.
- Giri P., Mohapatra B. (2017) Candidate Gene. In: Vonk J., Shackelford T. (eds) *Encyclopedia of Animal Cognition and Behavior*. Springer, Cham. https://doi.org/10.1007/978-3-319-47829-6_1-1
- Gordon, S., & Taylor, P. R. (2005). Monocyte and macrophage heterogeneity. *Nature reviews immunology*, 5(12), 953-964..
- Graham, J. G., MacDonald, L. J., Hussain, S. K., Sharma, U. M., Kurten, R. C., & Voth, D. E. (2013). Virulent *Coxiella burnetii* pathotypes productively infect primary human alveolar macrophages. *Cellular microbiology*, 15(6), 1012-1025..
- Grant, A. V., El Baghdadi, J., Sabri, A., El Azbaoui, S., Alaoui-Tahiri, K., Rhorfi, I. A., ... & Abel, L. (2013). Age-dependent association between pulmonary tuberculosis and common TOX variants in the 8q12–13 linkage region. *The American Journal of Human Genetics*, 92(3), 407-414.
- Graves, S., & Stenos, J. (2009). Rickettsioses in Australia. *Annals of the New York Academy of Sciences*, 1166(1), 151-15
- Graves, S. R., & Islam, A. (2016). Endemic Q fever in New South Wales, Australia: a case series (2005–2013). *The American Journal of Tropical Medicine and Hygiene*, 95(1), 55.
- Greenwood, C. M., Fujiwara, T. M., Boothroyd, L. J., Miller, M. A., Frappier, D., Fanning, E. A., ... & Morgan, K. (2000). Linkage of tuberculosis to chromosome 2q35 loci, including NRAMP1, in a large aboriginal Canadian family. *The American Journal of Human Genetics*, 67(2), 405-416.
- Guatteo, R. (2013). Q Fever an emerging disease. *Servet Editorial*.
- Guigno, D., Coupland, B., Smith, E. G., Farrell, I. D., Desselberger, U., & Caul, E. O. (1992). Primary humoral antibody response to *Coxiella burnetii*, the causative agent of Q fever. *Journal of Clinical Microbiology*, 30(8), 1958 LP – 1967.
- Hackstadt, T. (1986). Antigenic variation in the phase I lipopolysaccharide of *Coxiella burnetii* isolates. *Infection and immunity*, 52(1), 337-340.

- Harris, R. J., Storm, P. A., Lloyd, A., Arens, M., & Marmion, B. P. (2000). Long-term persistence of *Coxiella burnetii* in the host after primary Q fever. *Epidemiology & Infection*, 124(3), 543-549.
- He, S., Wang, B., Zhu, X., Chen, Z., Chen, J., Hua, D., ... & Jin, T. (2017). Association of IFNGR1 and IFNG genetic polymorphisms with the risk for pulmonary tuberculosis in the Chinese Tibetan population. *Oncotarget*, 8(58), 98417.
- Heidari, Z., Moudi, B., & Mahmoudzadeh-Sagheb, H. (2020). Interferon gamma gene polymorphisms and chronic hepatitis B infections in an Iranian population. *The Turkish journal of gastroenterology: the official journal of Turkish Society of Gastroenterology*, 31(7), 515–521. <https://doi.org/10.5152/tjg.2020.181024>
- Helbig, K., Harris, R., Ayres, J., Dunckley, H., Lloyd, A., Robson, J., & Marmion, B. P. (2005). Immune response genes in the post-Q-fever fatigue syndrome, Q fever endocarditis and uncomplicated acute primary Q fever. *Qjm*, 98(8), 565-574..
- Helbig, K. J., Heatley, S. L., Harris, R. J., Mullighan, C. G., Bardy, P. G., & Marmion, B. P. (2003). Variation in immune response genes and chronic Q fever. Concepts: preliminary test with post-Q fever fatigue syndrome. *Genes & Immunity*, 4(1), 82-85.
- Herrington,, D. M., Howard T. D, Hawkins, G. A, Reboussin, D.M., Xu, J., Zheng, S. L, Brosnihan, K. B, Meyers, D. A, Bleecker, E. R. (2002 a). Estrogen-receptor polymorphisms and effects of estrogen replacement on high-density lipoprotein cholesterol in women with coronary disease. *N Engl J Med* 346:967–974.
- Hijikata, M., Shojima, J., Matsushita, I., Tokunaga, K., Ohashi, J., Hang, N. T., ... & Keicho, N. (2012). Association of IFNGR2 gene polymorphisms with pulmonary tuberculosis among the Vietnamese. *Human genetics*, 131(5), 675-682.
- Hilbink, F., Penrose, M., Kovacova, E., & Kazar, J. (1993). Q fever is absent from New Zealand. *International journal of epidemiology*, 22(5), 945-949.
- Holland, S. M. (2001, March). Immunotherapy of mycobacterial infections. In *Seminars in respiratory infections* (Vol. 16, No. 1, pp. 47-59).
- Holland, S. M., Eisenstein, E. M., Kuhns, D. B., Turner, M. L., Fleisher, T. A., Strober, W., & Gallin, J. I. (1994). Treatment of refractory disseminated nontuberculous mycobacterial infection with interferon gamma: a preliminary report. *New England Journal of Medicine*, 330(19), 1348-1355.

- Honstetter, A., Imbert, G., Ghigo, E., Gouriet, F., Capo, C., Raoult, D., & Mege, J. L. (2003). Dysregulation of cytokines in acute Q fever: role of interleukin-10 and tumor necrosis factor in chronic evolution of Q fever. *The Journal of infectious diseases*, 187(6), 956-962..
- Honstetter, A., Ghigo, E., Moynault, A., Capo, C., Toman, R., Akira, S., ... & Mege, J. L. (2004). Lipopolysaccharide from *Coxiella burnetii* is involved in bacterial phagocytosis, filamentous actin reorganization, and inflammatory responses through Toll-like receptor 4. *The Journal of Immunology*, 172(6), 3695-3703.
- Huang, H. H., Shih, W. L., Li, Y. H., Wu, C. F., Chen, P. J., Lin, C. L., ... & Yu, M. W. (2011). Hepatitis B viraemia: its heritability and association with common genetic variation in the interferon γ signalling pathway. *Gut*, 60(1), 99-107.
- Huang, Y., Yang, H., Borg, B. B., Su, X., Rhodes, S. L., Yang, K., ... & Liang, T. J. (2007). A functional SNP of interferon- γ gene is important for interferon- α -induced and spontaneous recovery from hepatitis C virus infection. *Proceedings of the National Academy of Sciences*, 104(3), 985-990.
- Hunter, A. J., & Girard, D. E. (2001). Thrombolytics in infectious endocarditis associated myocardial infarction. *The Journal of emergency medicine*, 21(4), 401-406. Illumina. (2016). Nextera XT DNA Library Prep Reference Guide (15031942).
- International Livestock Research Institute (ILRI). (2012) Mapping of Poverty and Likely Zoonoses Hotspots. UK: Report to Department for International Development.
- Jansen, A. F., Schoffelen, T., Bleeker-Rovers, C. P., Wever, P. C., Jaeger, M., Oosting, M., ... & van de Vosse, E. (2019). Genetic variations in innate immunity genes affect response to *Coxiella burnetii* and are associated with susceptibility to chronic Q fever. *Clinical Microbiology and Infection*, 25(5), 631-e11..
- Jansen, A. F., Schoffelen, T., Textoris, J., Mege, J. L., Bleeker-Rovers, C. P., Roest, H. I., ... & van Deuren, M. (2017). Involvement of matrix metalloproteinases in chronic Q fever. *Clinical Microbiology and Infection*, 23(7), 487-e7.
- Jülicher, S., Bongartz, M., Luty, A. J., Kremsner, P. G., & Kun, J. F. (2003). Functional analysis of a promoter variant of the gene encoding the interferon-gamma receptor chain I. *Immunogenetics*, 54(10), 675-680.
- Kahl, G. (2015). MiSeq™ system. The Dictionary of Genomics, Transcriptomics and

Proteomics. <https://doi.org/10.1002/9783527678679.dg07722>

- Kerner, G., Rosain, J., Guérin, A., Al-Khabaz, A., Oleaga-Quintas, C., Rapaport, F., ... & Bustamante, J. (2020). Inherited human IFN γ deficiency underlies mycobacterial disease. *The Journal of clinical investigation*, 130(6), 3158-3171.
- Kováčová, E., Kazár, J., & Šimková, A. (1998). Clinical and Serological Analysis of a Q Fever Outbreak in Western Slovakia with Four-Year Follow-Up. *European Journal of Clinical Microbiology and Infectious Diseases*, 17(12), 867–869. doi:10.1007/s100960050209
- Langley, J. M., Marrie, T. J., LeBlanc, J. C., Almudevar, A., Resch, L., & Raoult, D. (2003). *Coxiella burnetii* seropositivity in parturient women is associated with adverse pregnancy outcomes. *American journal of obstetrics and gynecology*, 189(1), 228-232.
- Lei nº 12/2005 de 26 de janeiro, Diário da República n.º 18/2005, Série I-A de 2005-01-26, páginas 606 - 611
- Li, H. T., Zhang, T. T., Zhou, Y. Q., Huang, Q. H., & Huang, J. (2006). SLC11A1 (formerly NRAMP1) gene polymorphisms and tuberculosis susceptibility: a meta-analysis. *The International Journal of Tuberculosis and Lung Disease*, 10(1), 3-12.
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987-2993.
- Limonard, G. J., Thijsen, S. F., Bossink, A. W., Asscheman, A., & Bouwman, J. J. (2012). Developing a new clinical tool for diagnosing chronic Q fever: the *Coxiella* ELISPOT. *FEMS Immunology & Medical Microbiology*, 64(1), 57-60.
- Lonsdale J, Thomas J, Salvatore M *et al* (2013) The genotype-tissue expression (GTEx) project. *Nat Genet* 45(6):580–585
- Lukáčová, M., Barak, I., & Kazar, J. (2008). Role of structural variations of polysaccharide antigens in the pathogenicity of Gram-negative bacteria. *Clinical microbiology and infection*, 14(3), 200-206.
- MacArthur, D. G., Manolio, T. A., Dimmock, D. P., Rehm, H. L., Shendure, J., Abecasis, G. R., ... & Gunter, C. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature*, 508(7497), 469-476.
- Malik, S., Abel, L., Tooker, H., Poon, A., Simkin, L., Girard, M., ... & Schurr, E. (2005). Alleles of

the NRAMP1 gene are risk factors for pediatric tuberculosis disease. *Proceedings of the National Academy of Sciences*, 102(34), 12183-12188.

Maltezou H, Raoult D. Q fever in children. (2002). *Lancet Infect Dis*; 2(11): 686-691.

Marino, S., Cilfone, N. A., Mattila, J. T., Linderman, J. J., Flynn, J. L., & Kirschner, D. E. (2015). Macrophage polarization drives granuloma outcome during *Mycobacterium tuberculosis* infection. *Infection and immunity*, 83(1), 324-338.

Marrie, T. J. (1990). Q fever in octogenarians. *Annals of the New York Academy of Sciences*, 590, 266-270..

Maurin, M., & Raoult, D. F. (1999). Q fever. *Clinical microbiology reviews*, 12(4), 518-553.

McDonald, M-L., Schrauwen I., Leal S.,(2018). Genome-Wide Association Exercise - Data Quality Control disponível no endereço https://github.com/statgenetics/statgen-courses/blob/master/handout/PLINK_data_QC.pdf (ver também: [https://statgen.us/Tutorials#GWAS: Data Quality Control](https://statgen.us/Tutorials#GWAS:DataQualityControl))

Million, M., and Raoult, D. (2015) Recent advances in the study of Q fever epidemiology, diagnosis and management. *J Infect* 71: S2–S9.

Mills, C. D. (2015). Anatomy of a discovery: m1 and m2 macrophages. *Frontiers in immunology*, 6, 212.

Mills, C. D., & Ley, K. (2014). M1 and M2 macrophages: the chicken and the egg of immunity. *Journal of innate immunity*, 6(6), 716-726.

Mori, M., Mertens, K., Cutler, S. J., & Santos, A. S. (2017). Critical aspects for detection of *Coxiella burnetii*. *Vector-Borne and Zoonotic Diseases*, 17(1), 33-41..

Mullis K.B. (1987) Process for amplifying nucleic acid sequences. U S Patent 4.683.202.

Nakashima, H., Inoue, H., Akahoshi, M., Tanaka, Y., Yamaoka, K., Ogami, E., ... & Niho, Y. (1999). The combination of polymorphisms within interferon- γ receptor 1 and receptor 2 associated with the risk of systemic lupus erythematosus. *FEBS letters*, 453(1-2), 187-190..

Narasaki, C. T., & Toman, R. (2012). Lipopolysaccharide of *Coxiella burnetii*. *Coxiella burnetii: recent advances and new perspectives in research of the Q fever bacterium*, 65-90.

- Neupane, K., & Kaswan, D. (2020). *Coxiella Burnetii*. In *StatPearls [Internet]*. StatPearls Publishing.
- OIE Terrestrial Manual. (2018) Q fever. Paris, France: World Organization of Animal Health.
- Overend, L., & Rose, E. (2012). Uncertainties in managing myocardial infarction associated with infective endocarditis. *Experimental and clinical cardiology*, 17(3), 144–145.
- Oyston, P.C.F., Davies C. (2011). Q fever: the neglected biothreat agente. *Journal of Medical Microbiology*. 60 (1), 9–21. <https://www.ncbi.nlm.nih.gov/pubmed/2103050>
- Pacheco AG, Cardoso CC, Moraes MO (2008) IFNG +874T/A, IL10-1082G/A and TNF-308G/A polymorphisms in association with tuberculosis susceptibility: a meta-analysis study. *Hum Genet* 123(5): 477–484
- Peacock, M. G., Philip, R. N., Williams, J. C., & Faulkner, R. S. (1983). Serological evaluation of Q fever in humans: enhanced phase I titers of immunoglobulins G and A are diagnostic for Q fever endocarditis. *Infection and immunity*, 41(3), 1089-1098.
- Philip, C. B. (1948). Comments on the name of the Q fever organism. *Public Health Rep.*, 63, 58.
- Q-VAX. Q fever vaccine and Q-VAX skin test. Product Information e TGA approved. Australia: CSL Biotherapies; 2008.
- Ramakrishnan, L. (2012). Revisiting the role of the granuloma in tuberculosis. *Nature Reviews Immunology*, 12(5), 352-366.
- Raoult, D., Marrie, T. J., & Mege, J. L. (2005). Natural history and pathophysiology of Q fever. *The Lancet infectious diseases*, 5(4), 219-226.
- Raoult, D., Tissot-Dupont, H., Foucault, C., & Q fever 1985–1998. (2000). Clinical and epidemiologic features of 1,383 infections. *Medicine (Baltimore)*, 79, 109-123..
- Regateiro, F. (2007). *Manual de Genética Médica*. 1ª Edição. 2ª Reimpressão. Coimbra: Imprensa da Universidade de Coimbra.
- Roest, H. I. J., Tilburg, J. J. H. C., Van der Hoek, W., Vellema, P., Van Zijderveld, F. G., Klaassen, C. H. W., & Raoult, D. (2011). The Q fever epidemic in The Netherlands: history, onset, response and reflection. *Epidemiology & Infection*, 139(1), 1-12..

- Roest, H. I., van Solt, C. B., Tilburg, J. J., Klaassen, C. H., Hovius, E. K., Roest, F. T., ... & van Zijderveld, F. G. (2013). Search for possible additional reservoirs for human Q fever, The Netherlands. *Emerging Infectious Diseases*, 19(5), 834.
- Rossouw M, Nel HJ, Cooke GS, van Helden PD, Hoal EG (2003) Association between tuberculosis and a polymorphic NFkappaB binding site in the interferon gamma gene. *Lancet* 361(9372):1871– 1872
- Rousset É, Cremoux R de, Bronner A, Jourdain E, Touratier A, Sidi-Boumedine K. La fièvre Q. *Bull des GTV - Hors Série “Zoonose Mal bactériennes.”* 2012; 53–67.
- Sabatier, F., Dignat-George, F., Mege, J. L., Brunet, C., Raoult, D., & Sampil, J. (1997). CD4+ T-cell lymphopenia in Q fever endocarditis. *Clinical Diagnostic Laboratory Immunology*, 4(1), 89-92.
- Sahu, R., Rawool, D. B., Dhaka, P., Yadav, J. P., Mishra, S. P., Kumar, M., ... & Barbuddhe, S. B. (2021). Current perspectives on the occurrence of Q fever: highlighting the need for systematic surveillance for a neglected zoonotic disease in Indian subcontinent. *Environmental Microbiology Reports*, 13(2), 138-158.
- Saiki, R. K., Scharf, S., Faloona, F. Mullis, K. B, Horn, G. T., Erlich, H. A., Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230(4732):1350-1354.
- Sandoz, K. M., Popham, D. L., Beare, P. A., Sturdevant, D. E., Hansen, B., Nair, V., & Heinzen, R. A. (2016). Transcriptional profiling of *Coxiella burnetii* reveals extensive cell wall remodeling in the small cell variant developmental form. *PLoS One*, 11(2), e0149957..
- Santos A. S., Bacellar F, & França A. (2007) Febre Q: revisão de conceitos. *Revista Portuguesa de Medicina Interna*; 12(2):90-9.
- Santos, A. S., De Sousa, R., Alves, F., Proença, P., Núncio, M. S., Dumler, J. S., & Bacellar, F. (2009). Isolation of *Coxiella burnetii* from the blood of a patient with positive *Anaplasma phagocytophilum* serological results. *Clinical Microbiology and Infection*, 15, 192-193..
- Santos, A. S., Tilburg, J. J., Botelho, A., Barahona, M. J., Núncio, M. S., Nabuurs-Franssen, M. H., & Klaassen, C. H. (2012). Genotypic diversity of clinical *Coxiella burnetii* isolates from Portugal based on MST and MLVA typing. *International Journal of Medical Microbiology*, 302(6), 253-256.

- Santos, A. S. (2015). Febre Q: do diagnóstico à investigação ecoepidemiológica de *Coxiella burnetii* no contexto da infeção humana. *Boletim Epidemiológico Observações*, 4, 20-24.
- Santos, A. S. (2017). Human diagnosis and epidemiology of Q fever in Portugal. *The principles and Practice of Q fever*, 135.
- Santos, A. S. (2019). Febre Q. EM *Doenças associadas a artrópodes vetores e roedores*. 135-140. 2ª ed. Instituto Nacional de Saúde Doutor Ricardo Jorge. Lisboa.
- Shalia, K., Saranath, D., Rayar, J., Shah, V. K., Mashru, M. R., & Soneji, S. L. (2017). Identification of a single nucleotide polymorphism indicative of high risk in acute myocardial infarction. *The Indian Journal of Medical Research*, 146(4), 505.
- Schneider, T., Jahn, H., Steinhoff, D., Guschoreck, H., Liesenfeld, O., Mäter-Böhm, H., ... Dissmann, T. (1993). [A Q fever epidemic in Berlin. The epidemiological and clinical aspects]. *Deutsche Medizinische Wochenschrift (1946)*, 118, 689–695. doi:10.1055/s-2008-1059379
- Schoffelen, T., Joosten, L. A., Herremans, T., de Haan, A. F., Ammerdorffer, A., Rümke, H. C., ... & van Deuren, M. (2013). Specific interferon γ detection for the diagnosis of previous Q fever. *Clinical infectious diseases*, 56(12), 1742-1751.
- Schoffelen, T., Ammerdorffer, A., Hagenaars, J.C.J.P., Bleeker-Rovers, C.P., Wegdam-Blans, M.C., Wever, P.C., *et al.* (2015) Genetic variation in pattern recognition receptors and adaptor proteins associated with development of chronic Q fever. *J Infect Dis* 212: 818–829.
- Schoffelen, T., Textoris, J., Bleeker-Rovers, C. P., Amara, A. B., van der Meer, J. W. M., Netea, M. G., ... & van de Vosse, E. (2017). Intact interferon- γ response against *Coxiella burnetii* by peripheral blood mononuclear cells in chronic Q fever. *Clinical Microbiology and Infection*, 23(3), 209-e9.
- Schrijver, H. M., Hooper-Van Veen, T., Van Belzen, M. J., Crusius, J. B. A., Pena, A. S., Barkhof, F., ... & Uitdehaag, B. M. J. (2004). Polymorphisms in the genes encoding interferon- γ and interferon- γ receptors in multiple sclerosis. *European Journal of immunogenetics*, 31(3), 133-140.
- Shannon, J. G., Howe, D., & Heinzen, R. A. (2005). Virulent *Coxiella burnetii* does not activate human dendritic cells: role of lipopolysaccharide as a shielding molecule. *Proceedings*

of the National Academy of Sciences, 102(24), 8722-8727..

- Shannon, J. G., & Heinzen, R. A. (2009). Adaptive immunity to the obligate intracellular pathogen *Coxiella burnetii*. *Immunologic research*, 43(1), 138-148..
- Shastri, B. S. (2009). SNPs: impact on gene function and phenotype. *Single Nucleotide Polymorphisms*, 3-22..
- Silva Miranda, M., Breiman, A., Allain, S., Deknuydt, F., & Altare, F. (2012). The tuberculous granuloma: an unsuccessful host defence mechanism providing a safety shelter for the bacteria?. *Clinical and Developmental Immunology*, 2012.
- Sireci, G., Badami, G. D., Di Liberto, D., Blanda, V., Grippi, F., Di Paola, L., ... & Torina, A. (2021). Recent Advances on the Innate Immune Response to *Coxiella burnetii*. *Frontiers in cellular and infection microbiology*, 1069.
- Sobotta, K., Hillarius, K., Mager, M., Kerner, K., Heydel, C., & Menge, C. (2016). *Coxiella burnetii* infects primary bovine macrophages and limits their host cell response. *Infection and immunity*, 84(6), 1722-1734.
- Teunis, P. F. M., Schimmer, B., Notermans, D. W., Leenders, A. C. A. P., Wever, P. C., Kretzschmar, M. E. E., & Schneeberger, P. M. (2013). Time-course of antibody responses against *Coxiella burnetii* following acute Q fever. *Epidemiology & Infection*, 141(1), 62-73..
- The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* 526,68–74 (2015). <https://doi.org/10.1038/nature15393>
- Thiriou, J. D., Martinez-Martinez, Y. B., Endsley, J. J., & Torres, A. G. (2020). Hacking the host: exploitation of macrophage polarization by intracellular bacterial pathogens. *Pathogens and disease*, 78(1), ftaa009.
- Tissot-Dupont, H., Amadei, M. A., Nezri, M., & Raoult, D. (2004). Wind in November, Q fever in December. *Emerging infectious diseases*, 10(7), 1264.
- Toman, R., Skultety, L., & Ihnatko, R. (2009). *Coxiella burnetii* glycomics and proteomics—tools for linking structure to function. *Annals of the New York Academy of Sciences*, 1166(1), 67-78.
- Trinchieri, G. (2003). Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nature Reviews Immunology*, 3(2), 133-146..

- Tso HW, Ip WK, Chong WP, Tam CM, Chiang AKS, Lau YL (2005) Association of interferon gamma and interleukin 10 genes with tuberculosis in Hong Kong Chinese. *Genes Immun* 6(4):358–363
- Vallinoto ACR, Graça ES, Araújo MS, Azevedo VN, CayresVallinoto I, Machado LFA, Ishak MOG, Ishak R (2010) IFNG + 874T/A polymorphism and cytokine plasma levels are associated with susceptibility to Mycobacterium tuberculosis infection and clinical manifestation of tuberculosis. *Hum Immunol* 71(7):692–696
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., LevyMoonshine, A., ... DePristo, M. A. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, 43(1110), 11.10.1-11.10.33. doi:10.1002/0471250953.bi1110s43
- Van der Hoek, W., Schneeberger, P. M., Oomen, T., Wegdam-Blans, M. C., Dijkstra, F., Notermans, D. W., ... & van Duynhoven, Y. (2012). Shifting priorities in the aftermath of a Q fever epidemic in 2007 to 2009 in the Netherlands: from acute to chronic infection. *Eurosurveillance*, 17(3), 20059.
- Van der Hoek, W. V. D., Morroy, G., Renders, N. H., Wever, P. C., Hermans, M. H., Leenders, A. C., & Schneeberger, P. M. (2012). Epidemic Q fever in humans in the Netherlands. *Coxiella burnetii: Recent Advances and New Perspectives in Research of the Q Fever Bacterium*, 329-364.
- Van Schaik, E. J., Chen, C., Mertens, K., Weber, M. M., & Samuel, J. E. (2013). Molecular pathogenesis of the obligate intracellular bacterium *Coxiella burnetii*. *Nature Reviews Microbiology*, 11(8), 561-573.
- Vidal, S. M., Malo, D., Vogan, K., Skamene, E., & Gros, P. (1993). Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell*, 73(3), 469-485.
- Waag, D. M. (2007). *Coxiella burnetii*: host and bacterial responses to infection. *Vaccine*, 25(42), 7288-7295.
- White, B., Brooks, T., & Seaton, R. A. (2013). Q fever in military and paramilitary personnel in conflict zones: Case report and review. *Travel Medicine and Infectious Disease*, 11(2), 134–137. doi:10.1016/j.tmaid.2012.11.001
- Wielders, C. C. H., Hackert, V. H., Schimmer, B., Hodemaekers, H. M., de Klerk, A., Hoebe, C. J. P. A., ... & Janssen, R. (2015). Single nucleotide polymorphisms in immune response

genes in acute Q fever cases with differences in self-reported symptoms. *European Journal of Clinical Microbiology & Infectious Diseases*, 34(5), 943-950.

Williams, J. C., Peacock, M. G., & McCaul, T. F. (1981). Immunological and biological characterization of *Coxiella burnetii*, phases I and II, separated from host components. *Infection and Immunity*, 32(2), 840-851.

Wittwer, C. T., Kuskawa, N. (2006). Chap. 37 Nucleic Acid Techniques in: Burtis C.A., Ashwood E. R., Bruns D. E., Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, 4th ed. St Louis: Elsevier Saunders


Xiang, L., Elci, O. U., Rehm, K. E., & Marshall Jr, G. D. (2014). Associations between cytokine receptor polymorphisms and variability in laboratory immune parameters in normal humans. *Human immunology*, 75(1), 91-97.

Zamboni, D. S., Campos, M. A., Torrecilhas, A. C., Kiss, K., Samuel, J. E., Golenbock, D. T., ... & Gazzinelli, R. T. (2004). Stimulation of toll-like receptor 2 by *Coxiella burnetii* is required for macrophage production of pro-inflammatory cytokines and resistance to infection. *Journal of Biological Chemistry*, 279(52), 54405-54415.

Zamboni D. S., McGrath S, Rabinovitch M, Roy CR. *Coxiella burnetii* express type IV secretion system proteins that function similarly to components of the *Legionella pneumophila* Dot/Icm system. *Mol Microbiol*, (2003); 49(4): 965-976.

Anexos

Anexo 1 – Poster



Genetic determinants of infection in chronic Q fever

S David¹, L Castro, M Machado¹, U Gaspar, JC Robles-Fuentes, J Mendonça¹, J Ferr o¹, L Vieira¹, AS Santos¹,
¹ INSA– Instituto Nacional de Sa de Doutor Ricardo Jorge

Introduction

Q fever, caused by *Coxiella burnetii*, is a zoonotic disease with worldwide distribution. The most of affected patients develop acute Q fever, usually presenting as a self-limited febrile illness with varying degrees of hepatitis and/or pneumonia. However, 1-5% of *C. burnetii*-infected individuals develop chronic Q fever, a persistent infection mainly manifested as endocarditis. Genes involved in human immune response are well recognized to influence the clinical course of infection. Despite the above mentioned, the association between genetic variants in immune response genes and chronic Q fever remains practically unexplored. Therefore, we sought to identify variants in immune response genes associated with chronic Q fever resorting to a Candidate Gene Association Studies (CGAS) approach.

Materials and Methods

The workflow performed is outlined in Figure 1.

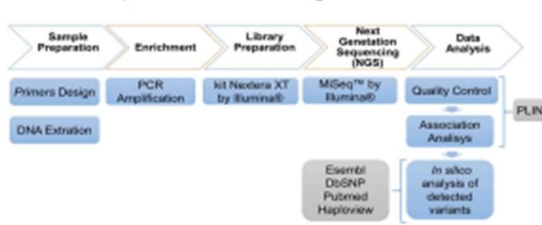


Figure 1. Study Workflow.

Candidate genes were selected and a Next Generation Sequencing (NGS) strategy was chosen for the genotyping of the 5'UTR region and exons of *IL12RB1*, *IFNGR1*, *IFNGR2*, *NRAMP1*, *TLR1*, *TLR2* and *TIMP1*. After informed consent, 34 healthy individuals and 60 Q fever patients (43 acute and 17 chronic forms) were recruited. The genes were amplified, in 2kb to 14kb fragments, with Takara LA Taq  DNA polymerase for Nextera library construction and Illumina  sequencing. Since NGS works simultaneously with billions of data points, it requires the application of several quality control metrics in order to guarantee reliable results (Figure 2). The data exploratory analysis (quality control) and the genotype-phenotype association were performed using the PLINK toolset. Specific metrics for CGAS were implemented.

Results

NGS has provided 607 SNPs from 94 DNA samples. In quality control of the genotyped data, 9 DNA samples with more than 10% of missing genotypes were excluded. Then, 24 SNPs that cumulatively had a MAF (minority allele frequency) greater than 5% and 5% of absent genotypes were excluded. We also excluded 170 SNPs that cumulatively had a MAF of less than 5% and 1% of absent genotypes. Three samples with more than 3% genotyping failure were also removed. Next, 8 SNPs that had a deviation from the Hardy-Weinberg equilibrium (HWE) with a p-value <10⁻³ were excluded. Overall, 405 SNPs and 82 DNA samples passed quality control. These were used to perform genotype-phenotype association studies in a CGAS approach to chronic Q fever. 324 SNPs showed variations between chronic patients and acute patients. Of these, 29 SNPs proved to be significant (p<0.05) and 4 passed the Bonferroni correction criteria (p<0.000154321).

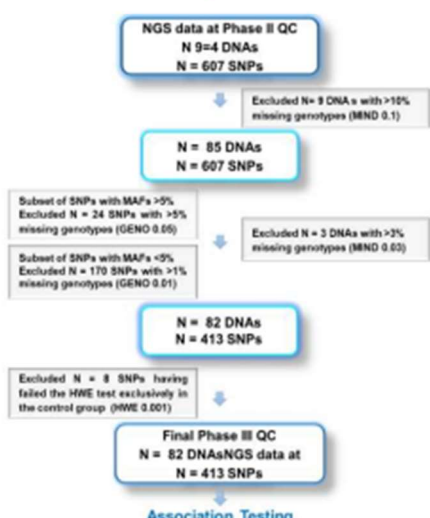


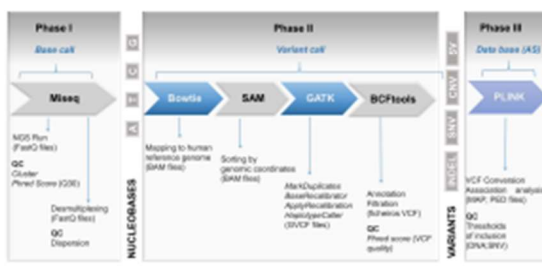
Figure 3. Phase III QC of genotyping data, adapted from the course communication of McDonald and collaborators (McDonald et al. 2018).

Conclusions

Several SNPs were identified passing Bonferroni criteria for statistical significance. The strategy appears promising for identifying genetic determinants of chronic Q fever.

Figure 2. Flow of the NGS methodology in phases (I, II, III), indication of the bioinformatics tools used in each step with the respective quality parameters.

Definitions: AS - Association studies; NGS - Next generation sequencing; QC - quality control.

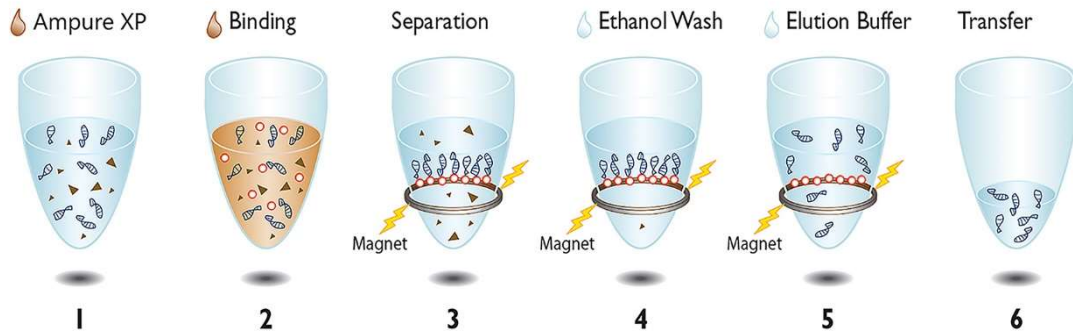


L Castro's contribution was in the context of preparation for his Master's degree in Molecular Technologies in Health, Escola Superior de Tecnologia da Sa de de Lisboa, Lisboa, Portugal.

U Gaspar's contribution was in the context of preparation for his Master's degree in Biomedical Sciences, Universidade da Beira Interior, Covilh , Portugal.

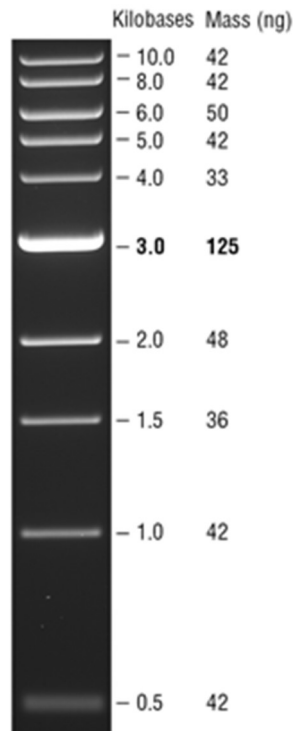
JC Robles-Fuentes's contribution was in the context of an ERASMUS program at the INSA while in preparation for his degree as an Advanced Expert in Biomedical and Clinical Laboratory, Atl ntida CIDEP, Granada, Spain.

Anexo 2 – Protocolo AMPure XP



Protocolo AMPure XP. Em cada poço da placa, com produtos PCR (1), adicionam-se as esferas magnéticas (2), os ADNs ligam às esferas magnéticas sendo estas ancoradas no poço através por magnetismo (3), são feitas lavagens (4) e separadas as esferas dos ADNs (4 e 5) (imagem retirada do protocolo Instructions For Use AMPure XP).

Anexo 3 – Marcador de Peso Molecular



Marcador de peso molecular padrão “1kb DNA Ladder” da New England BioLabs® (Werfen Portugal, Carnaxide, Portugal).

