

Agradecimentos

Esta tese significa para mim, mais do que o rosto do empenho e do esforço neste ano final de Mestrado, uma página de viragem. O final de uma etapa e o início de outra. Como tal, muito tenho a agradecer a todas as pessoas que me apoiaram nesta fase, às que me incentivaram à mudança, e todas aquelas que me deram a mão para o salto.

Em primeiro lugar agradeço ao meu Orientador, Doutor Pedro Esteves, por todo o enriquecimento científico e pessoal que me transmitiu e que foram sem dúvida, a alavanca neste processo. Um obrigada sincero por toda a disponibilidade, ajuda e compreensão, e um muito, muito, muito obrigada por ter acreditado em mim e no meu trabalho e assim me ter apoiado sempre.

Agradeço também à CESPU e à Comissão Coordenadora deste Mestrado, por me terem proporcionado a oportunidade de contactar com diferentes domínios científicos trazidos até nós por Professores e Investigadores de reconhecido mérito, e assim me terem aberto horizontes e perspectivas futuras.

Agradeço ainda ao Centro de Investigação em Biodiversidade e Recursos Genéticos, Centro de Testagem Molecular (CIBIO – CTM), a abertura para a realização do meu trabalho laboratorial, assim como todo o apoio e disponibilidade que sempre me ofereceram. Agradeço especialmente a disponibilidade das amostras utilizadas pois foram uma ajuda fulcral em todo o processo. Um obrigada especial para a Sandra, que desde o início me acompanhou e apoiou em todas as dúvidas e dificuldades que surgiram no laboratório e fora dele. Um obrigada também para a Joana que prontamente me ajudou a dar o primeiro passo no desenho dos *primers* e ao Miguel que na fase final me apoiou na análise de dados.

À Paulinha, agradeço o interesse e prontidão que sempre demonstrou em me ajudar a levar este trabalho a bom porto.

Ao Tiago, pelo companheirismo com que me acompanhou nestes 2 anos de Mestrado. Muito agradeço pelo ombro e pelos abraços, e acima de tudo, pela força de não me deixar assustar com as pedras do caminho.

Aos meus pais, obrigada pelo apoio e até pelas reticências em relação às minhas opções. Foram o empurrão e o travão, que me ajudaram a sonhar sem tirar os pés do chão. Aos meus irmãos, Bea e Raúl, obrigada por partilharem comigo os meus momentos de entusiasmo, loucura e desânimo.

Aos meus primos, obrigada pelos cafezinhos no sofá. Aos tios e avós, agradeço a compreensão pela minha ausência e o carinho com que sempre me receberam.

Resumo

O Receptor *Toll-like* 3 (TLR3) participa na resposta imune inata através do reconhecimento de agentes patogénicos, especialmente de vírus de RNA. Com este estudo, tirando partido da recente sequenciação completa do genoma do Coelho-bravo (*Oryctolagus cuniculus*), sequenciou-se o gene TLR3 e estimou-se a variabilidade genética nas suas populações naturais. O estudo deste gene nesta espécie poderá ser relevante na compreensão da interacção entre os agentes patogénicos e a primeira linha de defesa do organismo. Este mamífero é fortemente atingido por duas doenças virais, a Doença Hemorrágica Viral e a Mixomatose, o que realça a importância deste estudo para a compreensão da interacção vírus-hospedeiro. Em 40 indivíduos estudados, pertencentes às duas subespécies de Coelho-bravo, detectaram-se 41 posições polimórficas, 18 exclusivas das populações de *Oryctolagus cuniculus algirus* e 5 das de *Oryctolagus cuniculus cuniculus*. Estas substituições originaram 15 substituições aminoacídicas, sendo 5 exclusivas de *Oryctolagus cuniculus cuniculus* e 5 exclusivas de *Oryctolagus cuniculus algirus*. A análise das regiões intrónicas deste gene em Coelho-bravo permitiu detectar cinco *C-repeats* que se terão inserido em diferentes momentos da história evolutiva dos Lagomorfos. Adicionalmente, em sequências de TLR3 de mamíferos, detectaram-se 24 codões sob selecção positiva. Os resultados obtidos permitiram concluir que as populações de Coelho-bravo apresentam mais diversidade do que as de outras populações de mamíferos já estudadas. A detecção de uma elevada selecção positiva sugere que este gene terá evoluído com adaptações específicas para cada espécie. Para uma melhor compreensão do papel do TLR3 na resposta imunitária deverão ser realizados estudos de expressão génica e mais populações naturais terão que ser analisadas.

Abstract

The Toll-like receptor 3 (TLR3) participates in the innate immune response by recognizing pathogens, particularly RNA virus. With this study, taking advantage of the recent complete sequencing of the European Rabbit (*Oryctolagus cuniculus*) genome, we sequenced the TLR3 gene and estimated the genetic variability in their natural populations. The study of this gene in this species may be relevant in understanding the interplay between pathogens and the first line of defense. This mammal is strongly affected by two viral diseases, Rabbit Haemorrhagic Disease and Myxomatosis, which underlines the importance of this study for understanding virus-host interaction. In 40 individuals belonging to two subspecies studied, there were 41 polymorphic positions, 18 unique of *Oryctolagus cuniculus algirus* populations and 5 of *Oryctolagus cuniculus cuniculus*. These substitutions led to 15 amino acids changes, 5 unique of *Oryctolagus cuniculus algirus* and 5 of *Oryctolagus cuniculus cuniculus*. The analysis of intronic regions allowed the detection of five *C-repeats* which have been inserted at different times in evolutionary history of Lagomorphs. Additionally, in TLR3 mammalian sequences were detected 24 codons under positive selection. The results showed that the European Rabbit populations have more diversity than those of other mammal populations already studied. The detection of a high positive selection suggests that this gene has evolved with adaptations specific to each species. For a better understanding of the TLR3 role in the immune response, gene expression studies should be carried out and more natural populations have to be analyzed.

Índice

Índice

Lista de Abreviaturas	1
Introdução.....	7
1. Contextualização e organização geral da tese	9
2. Resposta Imunológica.....	11
3. Receptores <i>Toll-like</i>	13
3.1. Caracterização funcional, estrutural e molecular	13
3.2. Sinalização intracelular	17
3.3. TLR3	18
4. <i>Oryctolagus cuniculus</i>	25
4.1. História evolutiva.....	25
4.2. Características genómicas	26
4.3. Principais doenças virais.....	27
Objectivos.....	31
Material e Métodos	35
Resultados.....	43
1. Caracterização da estrutura molecular do gene TLR3 do Coelho-bravo.....	45
2. Caracterização genética do gene TLR3 nas populações de Coelho-bravo.....	52
3. Caracterização dos locais de N-glicosilação do TLR3	56
4. Padrões de selecção natural no gene TLR3 em mamíferos	57
5. Análise filogenética do gene TLR3.....	59
Discussão	61
1. Caracterização dos domínios do gene TLR3.....	63
1.1. Caracterização do domínio LRR-NT.....	64
1.2. Caracterização do domínio LRR	64
1.3. Caracterização do domínio LRR-CT.....	66
1.4. Caracterização do domínio Transmembranar	66
1.5. Caracterização do domínio TIR.....	67

2. Diversidade genética das populações de Coelho-bravo	68
3. Caracterização dos <i>C-repeats</i> do gene TLR3 em Coelho-bravo	69
4. Glicosilação.....	70
Considerações Finais	73
Referências Bibliográficas	77
Anexos	93

Índice de Figuras

Figura 1. Representação esquemática da estrutura dos receptores <i>Toll-like</i> . Adaptado de Bell et al. (2003). O domínio Transmembranar do TLR atravessa a bicamada fosfolipídica. Na parte superior está localizado o domínio extracelular e na parte inferior o domínio TIR.	14
Figura 2. Vias de sinalização dos receptores <i>Toll-like</i> nas células dendríticas e macrófagos. (Kumar et al. 2009).....	18
Figura 3. Modelo do complexo de sinalização TLR3-dsRNA. O dímero de receptores TLR3 (verde e azul), atravessa a membrana fosfolipídica (branco). Do lado inferior da membrana encontra-se o domínio TIR e do lado superior o domínio extracelular onde ocorre a ligação ao dsRNA (hélice azul e vermelha) (Liu et al. 2008).	21
Figura 4. Vias de sinalização do TLR3 (adaptado de Kumar et al. (2009)).....	22
Figura 5. Organização molecular dos genes TLRs. Adaptado de Matsushima et al. (2007).	23
Figura 6. Localização cromossómica do TLR3 no Homem (4q35). Adaptado de http://atlasgeneticsoncology.org/Indexbyalpha/idxa_T.html	23
Figura 7. Estrutura molecular do gene TLR3 no <i>Oryctolagus cuniculus</i> . A linha superior representa o tamanho total do gene, a inferior corresponde às regiões intrónicas e as caixas azuis representam os exões	24
Figura 8. Separação das populações de <i>Oryctolagus cuniculus</i> . A Nordeste da Península Ibérica, a subespécie de <i>O. c. cuniculus</i> , e a Sudoeste, a subespécie de <i>O. c. algirus</i> (Esteves et al. 2006).....	26
Figura 9. Localização geográfica das populações de Coelho-bravo da Península Ibérica analisadas neste trabalho.	37
Figura 10. Representação esquemática da localização dos <i>primers</i> no gene TLR3 com representação dos exões (E1, E2, E3 e E4), e dos locais de homologia dos <i>primers</i> assinalados pelas setas.....	38
Figura 11. Representação dos domínios conservados no gene TLR3 de Coelho-bravo através do programa SMART (A), SMART + Pfam (B) e LRRfinder (C). No esquema B, os rectângulos a tracejado laranja representam os 8 domínios encontrados através da opção de pesquisa de domínios Pfam existente no programa SMART.	45

Figura 12. Alinhamento da sequência do domínio TIR do Homem e Coelho-bravo. As regiões conservadas estão assinaladas pelas caixas azuis. A vermelho estão assinalados os aminoácidos da sequência de Coelho-bravo que diferem da sequência de Homem (dentro de cada região). A sequência consenso para cada região encontra-se acima de cada caixa azul.	49
Figura 13. Localização dos <i>C-repeats</i> na sequência genómica do TLR3.	50
Figura 14. Alinhamento dos <i>C-repeats</i> encontrados na sequência de TLR3 de Coelho-bravo (CR1 a CR5) com a sequência consenso (ConC) e os <i>C-repeats</i> já classificados (c3, c6, c8, c18 e c19) (Krane et al. 1991). No alinhamento, os pontos significam que a base é a mesma da sequência consenso e os traços são espaços introduzidos nas sequências para se conseguir o alinhamento. A- Adenina, C – Citosina, G- Guanina, T- Timina, S=C+G, R=A+G, Y=C+T, M=A+C, K=T+G. A cinzento estão assinalados os locais diagnósticos identificados por Krane et al. (2001) para a classificação dos <i>C-repeats</i>	51
Figura 15. Árvore filogenética de classificação dos <i>C-repeats</i> com base nas sequências de nucleotídeos. CR1 a CR5 correspondem aos <i>C-repeats</i> identificados na sequência do TLR3 de Coelho-bravo. De acordo com Krane et al. (1991), a sequência c18 pertence ao grupo I, a c6 e a c19 ao grupo II e a c3 e C8 pertencem ao grupo III.	52
Figura 16. Localização das alterações de aminoácidos nos domínios conservados identificados pelo programa LRRfinder na sequência de TLR3 de Coelho-bravo. Correspondência entre os domínios e as regiões codificantes do gene.	56
Figura 17. Árvore filogenética para o gene TLR3 com base na sequência de nucleotídeos do grupo de mamíferos analisado, utilizando o método <i>Neighbour Joining</i> e as opções <i>p-distance</i> e <i>complete deletion</i> . Os valores de <i>Bootstrap</i> obtidos com 1000 réplicas estão presentes nos nodos da árvore filogenética.	60

Índice de Tabelas

Tabela 1. Características dos <i>primers</i> utilizados na amplificação do gene TLR3 no Coelho-bravo: identificação dos <i>primers</i> , sequência, posição (pb), temperatura de <i>annealing</i> (TA) (°C) e tamanho do fragmento amplificado (pb).	38
Tabela 2. Identificação das sequências de mamíferos utilizadas na análise filogenética e na detecção de padrões de selecção natural.	40

Tabela 3. Identificação das sequências de <i>C-repeats</i> utilizadas neste trabalho anteriormente classificadas por Krane et al. (2001).	40
Tabela 4. Especificações usadas na análise com o software PAML	42
Tabela 5. Caracterização dos domínios encontrados através do programa SMART em conjunto com a opção de pesquisa de domínios Pfam apresentados na Figura 11-A,B. .	46
Tabela 6. Caracterização dos domínios encontrados através do programa LRRfinder e identificação das sequências de aminoácidos detectadas em cada domínio.	47
Tabela 7. Caracterização das alterações de nucleotídeos nas 40 amostras analisadas. .	55
Tabela 8. Caracterização dos possíveis locais de N-glicosilação da sequência do gene TLR3 de Coelho-bravo. A negrito estão realçados os locais da sequência de Coelho-bravo que diferem do Homem.....	57
Tabela 9. Sumário dos parâmetros estimados e valores de <i>likelihood</i> dos diferentes modelos de evolução para o TLR3 de Coelho-bravo.	58

Lista de Abreviaturas

- A - Adenina
Å - Ångström
Ala - Alanina (A)
Arg - Arginina (R)
Asn - Asparagina (N)
Asp - Ácido aspártico (D)
C - Citosina
CD – Célula dendrítica
CD14 - *Cluster of differentiation 14*
CDI - Célula dendrítica linfóide
CDm - Célula dendrítica mielóide
CpG ODN – (Ligando do TLR9) Oligodeoxinucleotídeos contendo dinucleotídeos citosina-fosfato-guanina não metilados
CR – *C-repeat*
Cys - Cisteína (C)
DAMP - Padrões moleculares associados ao dano ou ao *stress*
DNA – Ácido desoxirribonucleico
ds (RNA, DNA) – Ácido nucleico (DNA ou RNA) de cadeia dupla
EMCV - Vírus da Encefalomiocardite
FDA - *US Food and Drug Administration*
FEL – (Modelo) *Fixed-effect Likelihood*
G - Guanina
Gln - Glutamina (Q)
Glu - Ácido glutâmico (E)
Gly - Glicina (G)
His - Histidina (H)
HSP – Proteínas de choque térmico
HSV - Vírus herpes simplex
 $\text{I}\kappa\beta$ - Proteína inibitória que participa nas vias de sinalização dos TLRs
ICN - Livro Vermelho dos Vertebrados
IFN - Interferão
IgVH – Região variável das cadeias pesadas das imunoglobulinas
IKK – Inibidor da cinase do factor nuclear kappa B
IL - Interleucina
Ile - Isoleucina (I)
IRAK - Cinase associada ao receptor da IL-1

IRF - Factor de transcrição que participa nas vias de sinalização dos TLRs

Leu - Leucina (L)

LINE – *Long interspersed retrotransposable elements*

LPS - Lipopolissacarídeo

LRR – (Domínio) Rico em repetições de leucina

LRR-CT - (Domínio) Rico em repetições de leucina – C-terminal

LRR-NT - (Domínio) Rico em repetições de leucina – N-terminal

LRT - Teste de *Maximum likelihood*

Lys - Lisina (K)

MAPK – Cinase MAP

Met - Metionina (M)

miRNA - MicroRNA

MV – Mixoma Vírus

MyD88 – Proteína de resposta primária ao gene de diferenciação mielóide (88)

NEMO – Componente regulatório da IKK

NF- κ B – Factor Nuclear kappa B

NLR - Receptores *Nod-like*

O. c. a. – *Oryctolagus cuniculus algirus*

O.c. – *Oryctolagus cuniculus*

O.c.c. – *Oryctolagus cuniculus cuniculus*

PAML - *Phylogenetic analyses by maximum likelihood*

PAMP – Padrões moleculares associados a patogénios

pb – Pares de bases

PCR – Reacção em cadeia da polimerase

Pfam – Base de dados de famílias de proteínas

Phe - Fenilalanina (F)

PMN - Polimorfonucleares

Pro - Prolina (P)

PRR – Receptores de reconhecimento de padrões

REL – (Modelo) *Randon Effect Likelihood*

RHD – Doença hemorrágica viral

RHDV – Vírus da doença hemorrágica viral

RIP – Proteínas envolvida da sinalização dos TLRs

RLR - Receptores *RIG-like*

RNA – Ácido ribonucleico

RSV – Vírus Sincicial Respiratório

SARM – Proteína que contém um motivo tatu e o alfa estéril

Ser - Serina (S)

SINE- *Short interspersed retrotransposable elements*

siRNA - RNA de interferência curto

SLAC – (Modelo) *Single Likelihood Ancestor Sequences*

SMART - *Simple Modular Architecture Research Tool*

ss (RNA, DNA) – Ácido nucleico (DNA ou RNA) de cadeia simples

T - Timina

TA – Temperatura de *annealing*

TAK - Cinases β -activadas do factor de transformação do crescimento

TCR – Receptores dos antigénios das Células T

Thr - Treonina (T)

TIR – (Domínio) Toll/IL-1 e IL-18

TLR – Receptor *Toll-like*

TNF – Factor de Necrose Tumoral

TRAF-6 – Factor 6 associado ao receptor do TNF

TRAM – Molécula adaptadora relacionada com o TRIF

TRIF – Molécula adaptadora do domínio TIR indutora de IFN

tRNA – RNA de transferência

Trp - Triptofano (W)

TYP – (LRR) Típico

Tyr - Tirosina (Y)

Val - Valina (V)

WNV - Vírus do Nilo Ocidental

Introdução

1. Contextualização e organização geral da tese

O Coelho-bravo (*Oryctolagus cuniculus*) foi desde sempre uma espécie usada como modelo em estudos realizados nas áreas da saúde, em particular nas investigações da área da imunologia. De entre os muitos exemplos da utilidade da investigação no Coelho-bravo, para o desenvolvimento científico e cura de doenças com grande impacto nas populações humanas, destacam-se os trabalhos realizados no século XIX por Louis Pasteur, que utilizou a espinal medula do Coelho-bravo para preparar a vacina da raiva (Pasteur 1881), e o papel desta espécie na luta contra a sífilis (Parodi 1907, Wassermann 1906). Na primeira metade do século XX, os centros de investigação em Imunologia Molecular trabalhavam quase exclusivamente com o Coelho-bravo. Contudo, as dificuldades técnicas que o estudo desta espécie apresentou, como por exemplo a impossibilidade de obter anticorpos monoclonais e o aparecimento do ratinho como modelo animal, restringiram a utilização do Coelho-bravo. Ainda assim, este animal mantém uma enorme importância em várias áreas de investigação, tais como a transplantação e a oftalmologia. Hoje em dia, o Coelho-bravo é ainda um modelo de valor inestimável em pesquisa imunológica, incluindo a imunogenética e os estudos de doenças infecciosas. Pelo facto de ser possível comparar a diversidade genética entre populações selvagens e raças domésticas, é também um importante modelo para compreender os processos de especiação e domesticação.

As características únicas do sistema imunitário adaptativo do Coelho-bravo, das quais se destacam a utilização preferencial de um gene IgVH na biossíntese dos anticorpos e a existência de 14 subclasses de imunoglobulinas A, têm merecido um grande interesse e um grande investimento da comunidade científica internacional. Infelizmente, esse interesse não se tem repercutido no estudo do sistema imune inato (tema deste trabalho). Na verdade, com excepção de algum trabalho efectuado no estudo da variabilidade genética de algumas citocinas e dos seus receptores, pouco se sabe acerca do sistema imune desta espécie, particularmente, no que diz respeito ao estudo de receptores do sistema imune inato implicados no reconhecimento de organismos causadores de doença. Assim, de modo a ajudar a colmatar esta lacuna, neste trabalho estudou-se a diversidade genética do receptor *Toll-like 3* (TLR3) em populações de Coelho-bravo, tendo como referência estudos já efectuados para o Homem e outros mamíferos. A escolha do TLR3 deve-se à grande aptidão que esta molécula apresenta para o reconhecimento de ácidos nucleicos de origem viral, pois, como será explicado adiante, o Coelho-bravo é atingido por duas grandes doenças virais, a Doença Hemorrágica Viral (RHD) e a Mixomatose.

Este trabalho está dividido nos capítulos principais da Introdução, Material e Métodos, Resultados, Discussão e Considerações finais. Na Introdução, faz-se uma contextualização da importância dos receptores *Toll-like* no sistema imune inato, focando a sua caracterização estrutural e funcional. É dada uma maior relevância à função, às vias de sinalização e à estrutura molecular do TLR3. Para uma melhor compreensão do trabalho, será revisto o conhecimento actual da diversidade genética e história evolutiva do Coelho-bravo. Para as duas doenças virais já previamente referidas, Doença Hemorrágica Viral e Mixomatose, será feita uma descrição do seu modo de actuação, sintomatologia e evolução molecular. Em todo este capítulo será abordada a selecção natural, nomeadamente os padrões de selecção e a sua importância no âmbito de estudos de genética evolutiva. No capítulo Material e Métodos, serão expostas todas as metodologias adoptadas durante a realização deste trabalho, bem como todas análises de dados efectuadas. Nos capítulos seguintes, Resultados e Discussão, estão descritos os resultados obtidos para o TLR3 e são discutidas as suas implicações na co-evolução dos agentes patogénicos e das espécies de mamíferos, dando especial atenção ao Coelho-bravo. Assim, será feita uma caracterização da estrutura molecular do TLR3 no Coelho-bravo, realçando as alterações de nucleotídeos e de aminoácidos observadas nas amostras analisadas. Proceder-se-á a uma análise evolutiva, onde será apresentada a árvore filogenética do receptor TLR3, bem como os resultados de testes para averiguação da presença de padrões de selecção natural nos mamíferos. Com o capítulo de considerações finais procurar-se-á evidenciar a importância dos resultados obtidos e apontar possíveis linhas de investigação que poderão ser úteis para a compreensão da co-evolução parasita-hospedeiro.

2. Resposta Imunológica

O principal objectivo do sistema imunitário é proteger e garantir a sobrevivência dos indivíduos. Para isso, existem dois grandes tipos de resposta imunológica, a inata e a adaptativa. As respostas do sistema inato são mais simples e estão presentes nos animais vertebrados e nos invertebrados. As respostas adaptativas são mais complexas e estão somente presentes nos vertebrados. No entanto, independentemente do tipo de resposta, o organismo reage e acciona os seus mecanismos de defesa quanto detecta substâncias estranhas, denominadas antigénios. Um dos grandes problemas do sistema imunitário, e que origina o aparecimento de doenças auto-imunes, é a perda da capacidade de detectar moléculas do nosso organismo como próprias e assim identificá-las como antigénios. Essa situação origina doenças graves, como por exemplo a Artrite Reumatóide e o Lúpus.

A resposta imunológica inata é a primeira linha de defesa do nosso organismo e juntamente com a resposta imunológica adaptativa revela-se altamente eficiente nos processos de reconhecimento e destruição de antigénios. A resposta adaptativa apresenta duas grandes funções. A primeira é o reconhecimento dos antigénios usando dois tipos de moléculas, os receptores dos antigénios das células T (TCRs), e as regiões variáveis dos anticorpos produzidas pelas células B. Estas células apresentam grande especificidade no reconhecimento antigénico. Após esse reconhecimento, as células B e as T são estimuladas a proliferar e algumas delas, designadas por células de memória, são mantidas no plasma e formam a memória imunológica para um determinado antigénio. A segunda é a eliminação dos antigénios, que é realizada pelo sistema efector. Existem diversos mecanismos que participam neste processo, tais como: i) neutralização; ii) fagocitose; iii) reacção citotóxica.

A indução dos mecanismos do sistema imune inato não é específica do patogénio em questão. Na verdade independentemente do número de vezes que os constituintes do sistema inato interajam com vários patogénios, a resposta será sempre a mesma. As suas principais funções incluem a opsonização, fagocitose, apoptose, activação do complemento e das cascatas da coagulação, assim como activação das cascatas de sinalização pró-inflamatórias. Para além destes mecanismos de defesa do organismo, este sistema tem a capacidade de activar e modelar a resposta imune adaptativa através da indução de moléculas co-estimuladoras e de citocinas. Ao contrário do que acontece com o sistema imune adaptativo, o sistema imune inato é filogeneticamente conservado e encontra-se presente na grande maioria dos organismos multicelulares (Hoffmann et al. 1999).

O reconhecimento das moléculas estranhas através do sistema imune inato é efectuado pelos receptores de reconhecimento de padrões (PRRs) que se ligam a estruturas moleculares conservadas encontradas em grandes grupos de agentes patogénios, denominadas de padrões moleculares associados a patogénios (PAMPs) (Akira et al. 2006, Beutler 2009, Janeway and Medzhitov 2002, Medzhitov 2007, Medzhitov and Janeway 1997). Estes PAMPs são motivos conservados, específicos para cada grupo de microorganismos, que são essenciais para o seu metabolismo e portanto, fundamentais para a sua sobrevivência (Janeway and Medzhitov 2002, Medzhitov 2007). O reconhecimento destes padrões por parte dos PRRs tem algumas vantagens, entre as quais, o facto de os PAMPs serem apresentados apenas por microorganismos e não por células do hospedeiro, permitindo assim ao sistema imune inato distinguir entre o que é do organismo (próprio) e o que é estranho (não próprio). Para além disso, como os PAMPs são essenciais para a sobrevivência do microorganismo, mutações ou perda de padrões podem ser letais e portanto, esses padrões são bastante conservados, não permitindo muitas mutações. Por último, os PAMPs são invariáveis entre os microorganismos da mesma classe, o que implica que apenas é necessário um número limitado de PRRs para detectar a presença de uma infecção (Medzhitov and Janeway 1997). Existem vários grupos de PRRs, que podem ser secretados, expressos na superfície celular ou encontrados em compartimentos intracelulares (Medzhitov 2001), entre os quais os receptores *Nod-like* (NLRs), os receptores *RIG-like* (RLRs) e os receptores *Toll-like* (TLRs), que se destacam como uma das mais importantes famílias de PRRs (Janssens and Beyaert 2003, Kawai and Akira 2009).

3. Receptores *Toll-like*

3.1. Caracterização funcional, estrutural e molecular

O envolvimento dos TLRs na imunidade inata foi pela primeira vez descrito na vulgarmente designada mosca da fruta (*Drosophila melanogaster*), na qual o *Drosophila Toll* foi identificado como um receptor Transmembranar necessário durante o desenvolvimento embrionário, para o estabelecimento da polaridade dorso-ventral (Belvin and Anderson 1996, Hashimoto et al. 1988). A via de sinalização deste receptor apresenta algumas semelhanças com a via do receptor da interleucina-1 (IL-1) e da interleucina-18 (IL-18) nos mamíferos, e de facto, o domínio intracelular do *Drosophila Toll*, do receptor IL-1 e do IL-18 são altamente conservados, sendo denominados de domínios Toll/IL-1 e IL-18 (TIR) (Bell et al. 2003, Belvin and Anderson 1996, Beutler et al. 2006, Gay and Keith 1991). A sua similaridade sustentou a hipótese de a via mediada pelo receptor - *Toll* estar envolvida na regulação da resposta imune (Belvin and Anderson 1996, Lemaitre et al. 1996). Após a confirmação do papel do *Drosophila Toll* na defesa do organismo, nomeadamente contra infecções fúngicas (Lemaitre et al. 1996), vários homólogos deste receptor foram identificados em diversos organismos, desde plantas ao Homem, colectivamente designados de TLRs (Medzhitov et al. 1997).

Nos mamíferos, os TLRs são receptores de membrana tipo I, evolutivamente conservados (Janssens and Beyaert 2003) e com peso molecular que varia entre 90 a 115kD. Caracterizam-se estruturalmente pela presença de um domínio extracelular rico em repetições de leucina (LRRs) (Matsushima et al. 2007), que é responsável pelo reconhecimento dos PAMPs, um domínio Transmembranar e um domínio TIR localizado na sua região intracelular, onde se inicia a sinalização (Akira et al. 2006, Medzhitov et al. 1997) (Figura 1).

No domínio extracelular de um TLR podem existir entre 19 e 25 LRRs que são responsáveis pela interacção proteína-proteína, pela transdução de sinal, processamento de RNA e adesão celular (Matsushima et al. 2007). Este conjunto de LRRs é delimitado por sequências específicas denominadas de domínios N-terminal (NT) e C-terminal (CT) (Bell et al. 2003). Cada LRR possui entre 24 a 29 aminoácidos e pode ser dividido em duas partes, a sequência conservada e o segmento variável (Matsushima et al. 2007). A parte conservada consiste numa sequência xLxxLxLxx, e o segmento variável consiste em xΦxxΦx₄FxXLxx, onde o “x” poderá ser qualquer aminoácido e “Φ” um aminoácido hidrofóbico (alanina, leucina, isoleucina, metionina, fenilalanina, triptofano, tirosina e valina) (Bell et al. 2003, Matsushima et al. 2007). A família mais numerosa de LRRs é

designada de LRRs típicos cuja sequência é LxxLxLxxNxLxxLxxxxFxxLxx (Matsushima et al. 2007).

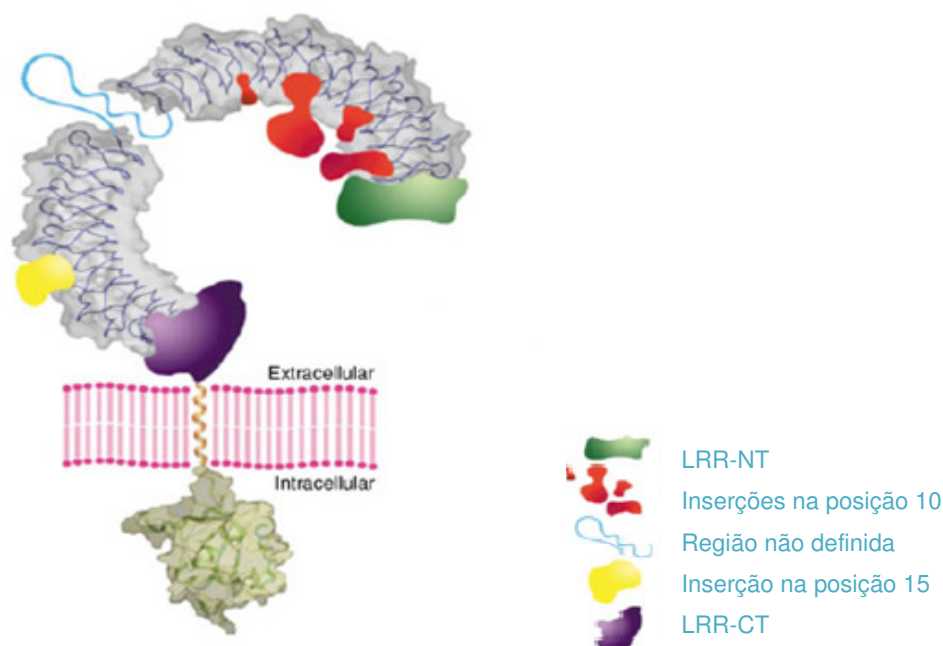


Figura 1. Representação esquemática da estrutura dos receptores *Toll-like*. Adaptado de Bell et al. (2003). O domínio Transmembranar do TLR atravessa a bicamada fosfolipídica. Na parte superior está localizado o domínio extracelular e na parte inferior o domínio TIR.

O domínio extracelular dos TLRs onde estão presentes estes LRRs, tem a forma de arco ou ferradura, constituída por cadeias em espiral, com uma face côncava e uma convexa. A face côncava é formada por uma longa hélice- β e por inserções que permitem a ligação a antígenos (Bell et al. 2003, O'Neill and Bowie 2007). Estas características proporcionam uma área de superfície de ligação 10 vezes maior do que as superfícies de ligação dos anticorpos ou dos TCRs (Bell et al. 2003). Estas repetições presentes no domínio extracelular sofrem frequentemente inserções nas posições 10 e 15 da sequência consenso que poderão facilitar a interação com PAMPs de maiores dimensões, ou com uma molécula assessora, tal como a MD-2 que está associada à molécula de TLR4 (Bell et al. 2003).

Estes receptores são expressos em alguns tipos de células, nomeadamente células dendríticas (CD), com diferentes padrões de expressão dependendo do subconjunto – células dendríticas mielóides (CDm) ou células dendríticas linfóides (CDI) (Liu et al. 2001). As CDm expressam TLR1, 2, 3, 4, 5 e 8, enquanto as CDI expressam exclusivamente TLR7 e TLR9, embora alguns estudos também refiram a expressão de TLR7 em CDm (Jarrossay et al. 2001, Kadowaki et al. 2001). As CD imaturas, iniciam o

seu amadurecimento em resposta a componentes microbianos e a expressão dos TLRs é diferente ao longo dos seus estádios de maturação. As CD imaturas apresentam TLR1, 2, 4 e 5, que desaparecem à medida que as células vão atingindo o estado de maturação (Mellman and Steinman 2001). Por sua vez, o TLR3 é apenas expresso em CD já amadurecidas (Muzio et al. 2000). Um outro tipo de células com funções imunes e que expressam TLRs, são os mastócitos que expressam TLR2, 4, 6 e 8 (Supajatura et al. 2001). Para além das células do sistema imune, estes receptores podem ser expressos em muitos outros tecidos, como por exemplo nas mucosas, tracto respiratório e gastrointestinal (Takeda et al. 2003).

A expressão dos TLRs é modulada por vários factores tais como a invasão microbiana, componentes microbianos e citocinas (Takeda et al. 2003). Por exemplo, a infecção viral induz a expressão do TLR1, 2, 3 e 7 em macrófagos (Miettinen et al. 2001), enquanto os lipopolissacarídeos das bactérias Gram-negativas (LPS) aumentam a expressão de TLR2 nos macrófagos e adipócitos (Matsuguchi et al. 2000). Em termos de localização a nível celular, os TLRs podem ser divididos consoante se localizam à superfície celular ou nos compartimentos endossomais intracelulares (Barton and Kagan 2009, Carty and Bowie 2010, Chaturvedi 2009). O TLR4 e o TLR5 são expressos na superfície celular e detectam LPS e flagelina bacteriana, respectivamente. Por sua vez, os heterodímeros TLR1/2 e TLR2/6 são expressos na superfície celular e detectam lipoproteínas bacterianas triacetiladas e diacetiladas, respectivamente. Por outro lado, os TLR3, 7, 8 e 9, são expressos nos endossomas e detectam ácidos nucleicos de origem microbiana (Alexopoulou et al. 2001, Barton and Kagan 2009, Edelmann et al. 2004, Heil et al. 2004, Kumar et al. 2009, Yoneyama and Fujita 2010). Alguns estudos sugerem que antes de estes TLRs intracelulares responderem a uma infecção viral, precisam de interagir com uma proteína constituída por 598 aminoácidos e com 12 domínios Transmembranares que se localiza no retículo endoplasmático, a UNC-93B (Conley 2007). Esta interacção regula o tráfico de TLRs intracelulares para o endolisossoma quando estimulados por ácido ribonucleico (RNA) ou ácido desoxirribonucleico (DNA), o que revela o papel indutor da UNC-93B na ligação dos TLRs aos seus ligandos (Chaturvedi 2009). Adicionalmente, foi demonstrado que esta proteína tem um papel importante na regulação da interacção entre o retículo endoplasmático e a sinalização dos TLR3, 7 e 9, desconhecendo-se, no entanto, o seu efeito no TLR8 (Akashi-Takamura and Miyake 2008).

Para além do reconhecimento de agentes patogénicos externos, os TLRs têm capacidade de reconhecer moléculas autólogas, tais como: proteínas de choque térmico (HSPs), produtos da matriz intercelular e DNA genómico, o que revela a capacidade destes receptores reconhecerem e actuarem perante sinais de destruição tecidual. Deste

modo, juntamente com o reconhecimento dos PAMPs, os TLRs actuam no reconhecimento de padrões moleculares associados ao dano ou ao *stress* (DAMPs) (Biragyn et al. 2002, Hold and El-Omar 2008, Matzinger 2002, O'Neill et al. 2009, Rakoff-Nahoum and Medzhitov 2009, Roelofs et al. 2006).

Vários estudos têm referido que TLRs funcionais são largamente expressos em células cancerígenas humanas ou em linhas celulares cancerígenas de ratinho e de Homem (He et al. 2007, Huang et al. 2008, Huang et al. 2005, Szajnik et al. 2009, Szczepanski et al. 2009). Desconhece-se o porquê e como essas células controlam e utilizam a activação dos TLRs, sendo sugerido que estes TLRs poderão ter um efeito oposto ao NF- κ B nas células normais e cancerígenas (Balkwill and Coussens 2004). Para além disso, outros estudos mostraram que a activação dos TLRs nas células cancerígenas intervém positivamente na inibição da proliferação do tumor e induz a apoptose. O TLR3-ligando inibe directamente células mamárias cancerígenas de origem humana, enquanto os ligandos do TLR7/8 e 9 são utilizados para tratar Leucemias Linfocíticas Crónicas (Salaun et al. 2006, Spaner and Masellis 2007). Neste sentido, e aproveitando o facto de os TLRs serem expressos em células imunes e iniciarem respostas imunes ou inflamatórias, têm sido feitos esforços no desenvolvimento de uma terapia anti-tumoral utilizando ligandos de TLRs. Como exemplo, o ligando do TLR7, Imiquimod ($C_{14}H_{16}N_4$) que é uma modificação da resposta imune aprovado em 1997 pela *United States Food and Drug Administration* (FDA) para o tratamento das verrugas genitais induzidas pelo vírus do papiloma humano e do carcinoma basocelular (Schon et al. 2003, Smits et al. 2008). O ligando do TLR9, oligodeoxinucleotídeos (ODN), contendo dinucleotídeos citosina-fosfato-guanina não metilados (CpG ODN), tem sido usado em ensaios clínicos de fase I de linfomas não Hodgkin das células B (Leonard et al. 2007).

Além das aplicações terapêuticas, os TLRs estão associados a modulação de outras respostas imunes. É exemplo disso a associação entre os TLRs e os microRNAs (miRNA). Tem sido demonstrado que os miRNA têm um papel importante no desenvolvimento de células imunes, diferenciação e resposta, incluindo as respostas inflamatórias mediadas por TLRs (Androulidaki et al. 2009, Baltimore et al. 2008, Lu and Liston 2009, Taganov et al. 2007). A sua expressão é, por si só, controlada por factores transcripcionais através das vias de sinalização do NF- κ B e MAP cinase (MAPK). Por exemplo, a estimulação por LPS pode regular positivamente a expressão de miRNAs em fagócitos do sistema imune inato do Homem, tais como monócitos e polimorfonucleares (PMN), envolvendo a produção de miRNAs como let-7e, miR-146^a, miR-155, miR-222, sugerindo que estes miRNAs são produzidos por estímulo inflamatório e têm um papel importante no sistema imune (Bazzoni et al. 2009).

3.2. Sinalização intracelular

O domínio TIR dos receptores *Toll-like* associa-se a diferentes moléculas adaptadoras. As mais correntemente conhecidas são a MyD88, TIRAP (Mal), TRIF (TICAM-1) e TRAM (TICAM-2) e o adaptador negativo – SARM, que bloqueia a via TRIF dependente (Jiang et al. 2006, O'Neill and Bowie 2007). O domínio MyD88 é utilizado por todos os TLRs excepto o 3, contudo, o TLR1, 2, 4 e 6 necessitam do TIRAP como molécula de ligação à MyD88. A cascata de sinalização destes últimos receptores prossegue com a participação das cinases associadas ao receptor da IL-1 (IRAKs) e do factor 6 associado ao TNF (TRAF-6). O TRAF-6, por sua vez, conduz à activação das cinases β -activadas do factor de transformação do crescimento (TAKs) que inicia uma cascata de cinases envolvendo o complexo IKK α /IKK β /NEMO e a fosforilação da proteína inibitória I κ B. A I κ B fosforilada, dissocia-se do complexo e é marcada para prosseguir para ubiquitinação e degradação nos proteossomas, levando à activação do factor de transcrição NF- κ B que se desloca para o núcleo, mediando assim a transcrição de alguns genes. Por outro lado, a TAK1, activa o complexo IKK, e também MAPKs como por exemplo ERK1/2, JNK e p38MAPK, que conduzem à activação do AP-1, c-Jun e c-fos e culminam com a transcrição de genes de citocinas inflamatórias (Zhang et al. 2000). Para além da via MyD88 dependente, o TLR4 partilha com o TLR3 a via TRIF-dependente, contudo, necessita do TRAM como molécula de ligação ao TRIF. Nesta situação, a activação do NF- κ B desenvolve-se através de duas cascatas independentes. Por um lado, o domínio N-terminal do TRIF interage com a TRAF-6 e deste modo activa o NF- κ B. Por outro lado, o domínio C-terminal faz interacção com a RIP1 e activa a TAK1. Para além disso, a via TRIF dependente conduz a activação de Interferão tipo I (IFN-I) através do IRF3. O IRF3 é fosforilado e activado pelas cinases IKK-relacionadas, TBK1 e IKKi. Relativamente ao TLR7 e 9, a sua activação conduz à produção de IFN-I (INF- α) através da activação do IRF-7 e também à produção de citocinas inflamatórias pela activação do NF- κ B e AP-1, via Myd88-dependente. Esta molécula interage com IRF-7 através da IRAK1, IRAK4, IKK α , TRAF3 e TRAF6. A IRF-7 é fosforilada pela IAK1/IKK α , translocada para o núcleo e induz à transcrição do IFN- α (Kumagai et al. 2009) (Figura 2).

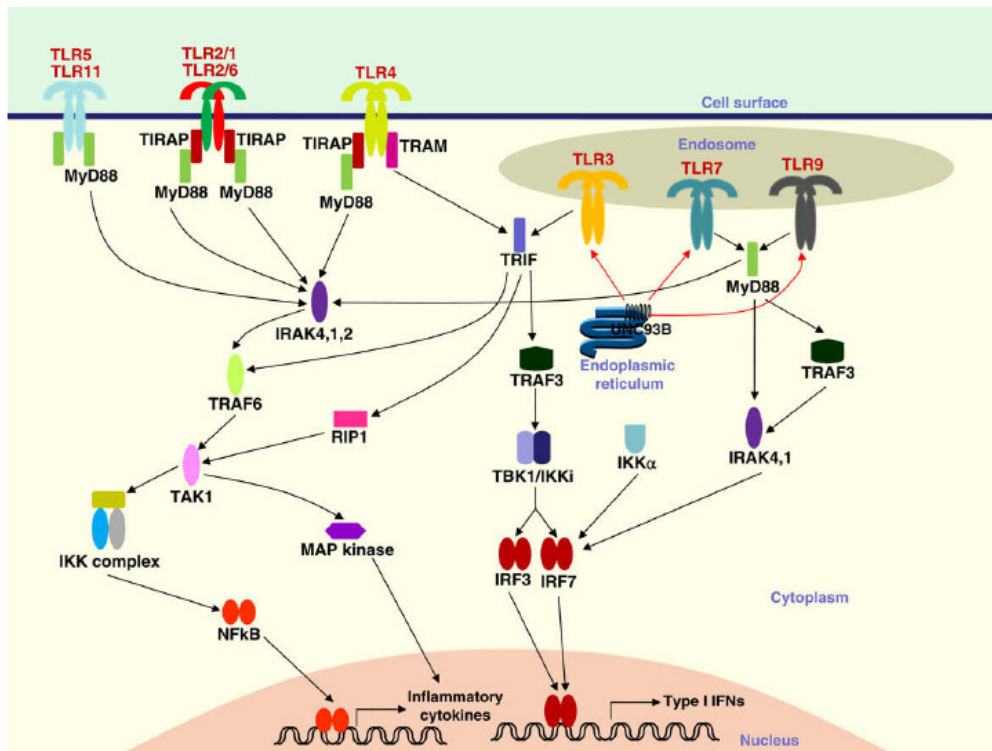


Figura 2. Vias de sinalização dos receptores *Toll-like* nas células dendríticas e macrófagos. (Kumar et al. 2009).

A ligação de um antígeno ao respectivo ectodomínio do receptor *Toll-like* resulta na formação de homo ou heterodímeros que farão o reconhecimento e a inicialização da resposta imune (Bell et al. 2003). Diferentes moléculas adaptadoras são recrutadas e são várias as vias de sinalização activadas que culminam em diferentes respostas imunológicas (Rassa and Ross 2003).

Kurt-Jones et al. (2000) demonstraram pela primeira vez a associação entre os TLRs e a detecção de vírus, ao relacionar a resposta do TLR4 à proteína de fusão (F) do vírus sincicial respiratório (RSV) (Kurt-Jones et al. 2000). Entretanto, vários outros TLRs têm sido associados à infecção viral, tais como o TLR1, TLR2, TLR3, TLR4, TLR6, TLR7, TLR8 e TLR9 (Alexopoulou et al. 2001, Heil et al. 2004, O'Brien et al. 1980, Takeuchi and Akira 2009, Takeuchi et al. 2000).

3.3. TLR3

Funções e importância na sinalização celular

A evidência do papel do TLR3 na resposta viral foi obtida quando Alexopoulou et al. (2001) demonstraram que ratinhos deficientes em macrófagos que expressavam TLR3

apresentavam resposta muito reduzida ao ácido polinosínico-policitídílico (poly I:C), cuja estrutura mimetiza dsRNA de origem viral (Alexopoulou et al. 2001). O poly (I:C) tem sido utilizado para a indução da expressão de IFN α/β , e deste modo, correlacionou-se directamente o TLR3 às citocinas efectoras em infecções virais. O TLR3 para além do reconhecimento do poly (I:C), tem capacidade de reconhecimento de RNA genómico, purificado a partir de vírus RNA de cadeia dupla (dsRNA), como por exemplo retrovírus, e de reconhecer dsRNA produzido durante a replicação viral de vírus de RNA de cadeia simples (ssRNA), tal como o RSV, o vírus da encefalomiocardite (EMCV) e o vírus do Nilo Ocidental (WNV) (Alexopoulou et al. 2001, Wang et al. 2004). Estudos revelam que ratinhos deficientes em TLR3 morrem primeiro do que ratinhos *wild-type* após infecção com citomegalovirus murino. Da mesma forma, a deficiência de TLR3 é associada com a susceptibilidade à infecção por vírus Herpes simplex (HSV) em humanos (Tabeta et al. 2004, Zhang et al. 2007). Por outro lado, apesar de muitos estudos demonstrarem que o TLR3 contribui para a eliminação de alguns vírus, existem estudos que evidenciam que outros vírus beneficiam da activação do TLR3. Foi demonstrado que em ratinhos infectados pelo WNV, o TLR3 facilita a entrada do vírus no cérebro através da produção periférica de citocinas inflamatórias que proporcionam a ruptura da barreira hemato-encefálica (Wang et al. 2004). Para além do reconhecimento viral, o TLR3 está implicado no reconhecimento de RNA de interferência curto (siRNA) independentemente da sequência, o que activa a produção de IL-12 e IFN- γ suprimindo eficientemente a angiogénese no ratinho, o que sugere que o siRNA poderá tratar desordens angiogénicas que afectam cerca de 8% da população mundial (Wang et al. 2004).

No que diz respeito à sua localização, o TLR3 é expresso nos endossomas das CDm, excepto nos fibroblastos e nas células epiteliais onde é expresso à superfície (Matsumoto and Seya 2008). Tal como outros receptores pertencentes a esta família, o TLR3 expressa-se predominantemente no cérebro, coração, pulmões e músculo, sugerindo uma resposta anti-inflamatória nestes tecidos (Carpenter and O'Neill 2007). Como exemplo, nos humanos, o TLR3 expressa-se no sistema nervoso central onde tem um papel de controlo do vírus HSV- tipo 1 (Zhang et al. 2007). Nos mamíferos, o TLR3 também foi descrito como um potente regulador negativo do crescimento axonal (Cameron et al. 2007). Adicionalmente, o ligando deste receptor juntamente com o IFN-I pode directamente intervir na inibição da proliferação e induzir a apoptose celular nas células do melanoma humano (Salaun et al. 2007).

Em termos estruturais, o TLR3 no Homem consiste num domínio extracelular constituído por 23 LRRs e por um domínio NT e um CT em cada extremidade respectivamente (Bell et al. 2003). Para além disso, possui um domínio Transmembranar

que se caracteriza por uma única hélice que faz a ligação ao domínio intracelular TIR (Bell et al. 2003). No domínio extracelular, o conjunto de LRRs formam uma estrutura com forma de arco ou ferradura com cadeias em espiral, onde uma face é preenchida de carboidratos e outra é livre de glicosilação. Existem 15 potenciais locais de ligação a carboidratos (Bell et al. 2005, Choe et al. 2005). As moléculas de TLR3 estão dispostas em dímeros e os resíduos de superfície altamente conservados no C-terminal, juntamente com a inserção específica no LRR20, formam a interface do homodímero (Choe et al. 2005) (Figura 3).

O domínio extracelular do TLR3 tem um diâmetro interno de aproximadamente 42 Ångström (Å), um diâmetro externo de 90Å e uma largura de 35Å (Bell et al. 2003, Botos et al. 2009). As cadeias laterais dos LRRs consenso situam-se no interior do ectodomínio e formam o core hidrofóbico. Os primeiros 8 resíduos de cada motivo LRR contribuem para a cadeia- β que se estende numa hélice- β e, que por sua vez, forma a face côncava e interna da molécula (Bell et al. 2003, Botos et al. 2009). O motivo LRR-NT contém uma estrutura em forma de gancho (*hairpin loop*) estabilizado por uma única ligação dissulfito. O motivo LRR-CT tem uma estrutura mais globular e contém uma α -hélice interna e duas ligações dissulfídicas que são necessárias para a funcionalidade do TLR3 (Ranjith-Kumar et al. 2007). A interacção entre os dois domínios LRR-CT é a única força que leva à dimerização das moléculas de TLR3 através de pontes de hidrogénio e interacções electrostáticas (Figura 3). Possivelmente, a proximidade destes domínios no lúmen do endossoma, poderá juntar os dois domínios TIR no lado citoplasmático numa posição capaz de recrutar as moléculas adaptadoras e de iniciar a sinalização (Botos et al. 2009). Assim, para que a resposta se inicie é necessário um estímulo antigénico que se irá ligar ao receptor. O TLR3 tem um local único para a ligação ao dsRNA, contudo, não é muito específico, pois não distingue as diferenças entre as sequências de base dos seus ligandos. Este receptor não se liga ao dsRNA em meio com pH neutro (de Bouteiller et al. 2005, Leonard et al. 2008), mas a ligação é forte quando ocorre a pH abaixo de 6,5, que pode ser encontrado em compartimentos endocíticos (de Bouteiller et al. 2005, Leonard et al. 2008). Para além disso, a ligação só acontece se o ligando tiver um comprimento superior a 40-45pb, e quanto maior for o comprimento, maior será a afinidade da ligação. A ligação de um dímero de TLR3 a um ligando de 40-45pb tem capacidade para produzir a activação de sinal (Figura 3). Se aumentar o comprimento do ligando, mais dímeros de TLR3 irão ligar-se. Por exemplo, se a molécula de dsRNA tiver um comprimento de 90pb, a ligação irá efectuar-se a 4 moléculas de TLR3, e a 6 se o ligando se estender a 139pb (Botos et al. 2009). A ligação do domínio externo do TLR3 ao dsRNA imobilizado é

saturada e reversível e pode ser inibida por dsRNA livre ou seu análogo, mas não por dsDNA, ssRNA ou ssDNA.

A interacção do TLR3 com o seu ligando acontece em dois locais específicos, separados aproximadamente por 55Å. Um dos locais de ligação situa-se em direcção ao C-terminal e inclui os resíduos LRR19 e LRR20. O segundo local de ligação situa-se próximo do N-terminal e inclui os resíduos LRR-NT, LRR1, LRR2 e LRR3 (Botos et al. 2009). Os locais de ligação interagem apenas com o esqueleto ribose-fosfato, o que explica o não reconhecimento da sequência de bases de ácidos nucleicos (Botos et al. 2009).

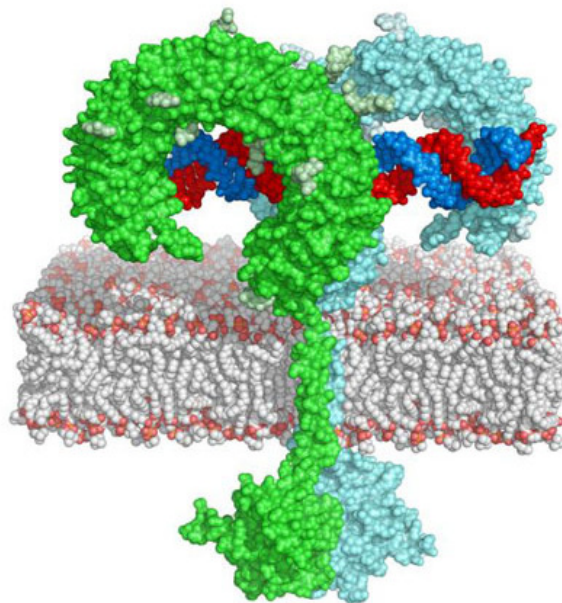


Figura 3. Modelo do complexo de sinalização TLR3-dsRNA. O dímero de receptores TLR3 (verde e azul), atravessa a membrana fosfolipídica (branco). Do lado inferior da membrana encontra-se o domínio TIR e do lado superior o domínio extracelular onde ocorre a ligação ao dsRNA (hélice azul e vermelha) (Liu et al. 2008).

Alguns estudos demonstraram que o CD14 potencia a activação do TLR3 por dsRNA através da ligação directa ao poly (I:C) e intervém na sua internalização celular (Lee et al. 2006). Como o domínio extracelular do CD14 consiste em LRRs, o poly (I:C) pode ser transferido do CD14 para o TLR3 (Kim et al. 2005). Assim, nos fibroblastos o TLR3 expresso na superfície celular está envolvido no reconhecimento de dsRNA (Matsumoto et al. 2002). Como o CD14 também é expresso à superfície destas células, pode cooperar com o TLR3 na internalização do dsRNA. Por outro lado, nas células dendríticas mielóides CD14 negativas, a internalização do dsRNA extracelular deverá acontecer com outro receptor (Okahira et al. 2005).

A interacção do domínio externo do TLR3 com dsRNA, leva à dimerização do receptor e ao recrutamento da molécula adaptadora onde se iniciará a cascata de sinalização. O TLR3 recruta a via que utiliza o TRIF como adaptador essencial e assim denominada via TRIF-dependente (Oshiumi et al. 2003, Yamamoto et al. 2002). Esta via culmina com a activação do NF- κ B, MAPKs e do factor de transcrição IRF3, que por sua vez é essencial à indução do IFN-I, particularmente o IFN- β (Oshiumi et al. 2003, Yamamoto et al. 2002), tal como se pode visualizar na Figura 4.

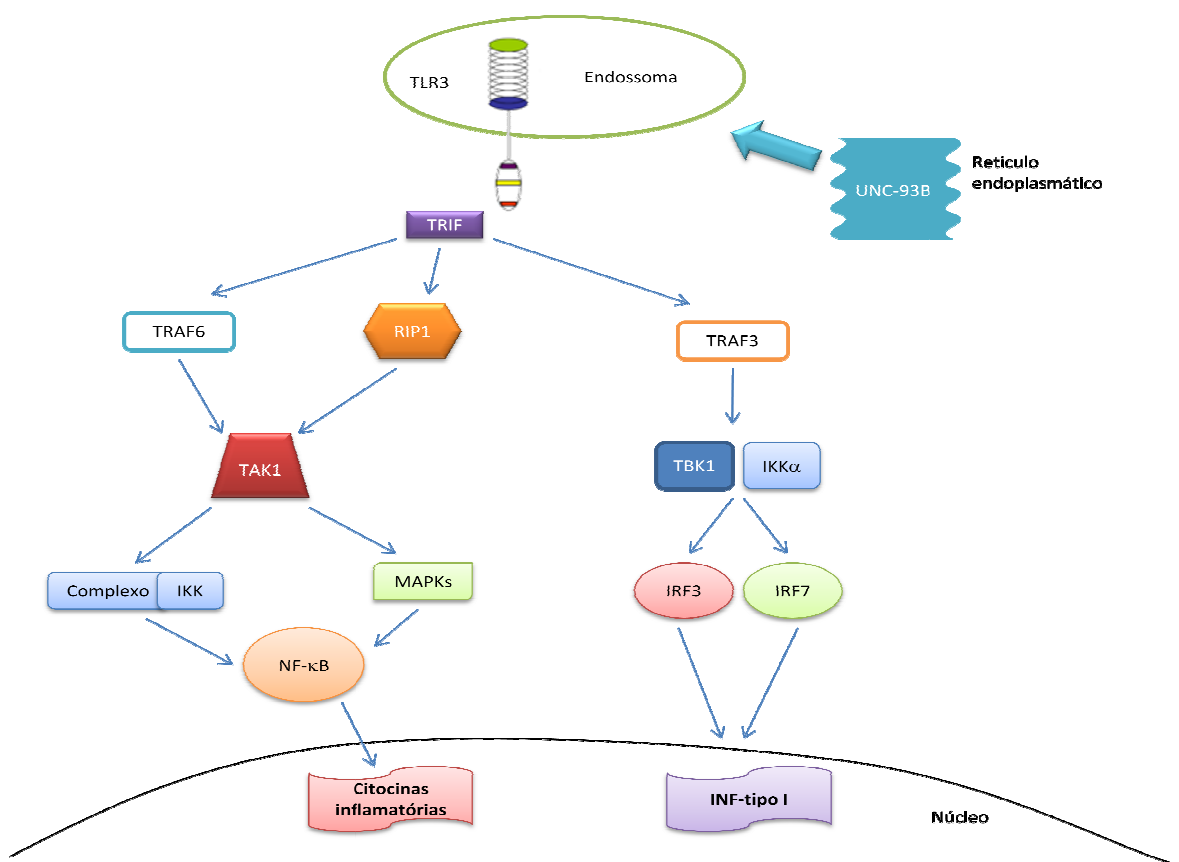


Figura 4. Vias de sinalização do TLR3 (adaptado de Kumar et al. (2009)).

Caracterização molecular e evolução do TLR3

Um grande número de TLRs tem sido caracterizado em vários organismos e a análise desta família de receptores, em particular nos vertebrados, permitiu estabelecer uma organização molecular comum (Figura 5) (Matsushima et al. 2007). Foi proposto que os TLRs se dividem em três regiões principais, o domínio externo ou ectodomínio, a região Transmembranar e o domínio TIR. O ectodomínio engloba o domínio LRR-NT, os

LRRs e o LRR-CT, os quais estão principalmente relacionados com o reconhecimento dos ligandos.

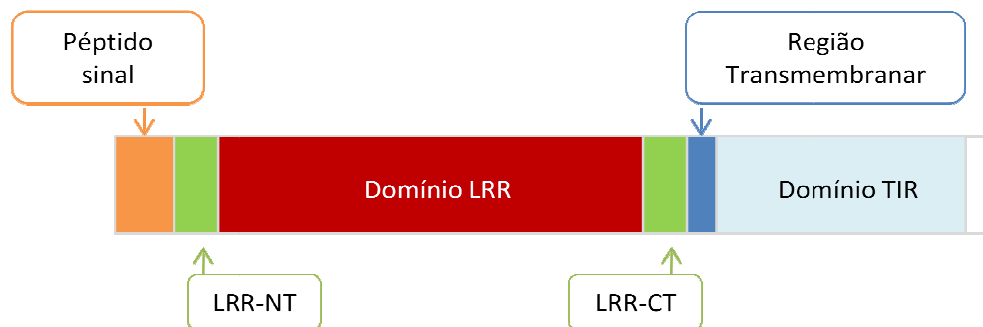


Figura 5. Organização molecular dos genes TLRs. Adaptado de Matsushima et al. (2007).

Estudos conduzidos em alguns mamíferos confirmam esta estrutura (Boraschi and Tagliabue 2006, Chang et al. 2009, Dhara et al. 2007, McGettrick and O'Neill 2004, West et al. 2006). Particularmente um estudo realizado no TLR3 de Búfalo e Nilgó (ou Antílope-azul) (Dhara et al. 2007), relativamente à sua estrutura molecular, indica que o domínio extracelular é composto por 675 aminoácidos (posição 23 à 697), o domínio Transmembranar é composto de 23 aminoácidos (posição 703 à 725), e o domínio citoplasmático TIR consiste em 246 aminoácidos (posição 755 a 900). Os primeiros 26 aminoácidos serão possivelmente a sequência sinal em cada espécie.

No Homem o gene TLR3 está localizado no *locus* 35 do braço longo do cromossoma 4 (4q35) (Genbank, www.ncbi.nlm.nih.gov) (Figura 6). Este possui uma sequência codificante de 2715 pares de bases (pb) organizada em 4 exões de comprimento 441pb, 192pb, 1853pb e 229pb, respectivamente (Genbank, www.ncbi.nlm.nih.gov, Astakhova et al. (2009)).

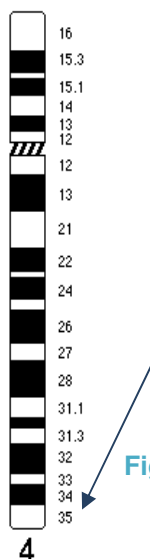


Figura 6. Localização cromossômica do TLR3 no Homem (4q35). Adaptado de http://atlasgeneticsoncology.org/Indexbyalpha/idxa_T.html.

No Coelho-bravo (*Oryctolagus cuniculus*) este gene está localizado no cromossoma 2, com um comprimento total de 11762pb (Genbank, www.ncbi.nlm.nih.gov). É constituído por quatro exões, com um comprimento de 441, 192, 1853 e 229pb (Astakhova et al. 2009) (Figura 7). Estes 2718pb originam 906 codões (Genbank, www.ncbi.nlm.nih.gov). Nesta espécie está descrita a existência de uma região rica em LRRs do aminoácido 42 ao 163, e do 433 ao 648. Para além disso, existe um domínio LRR-CT das posições 646 à 686. O domínio TIR está localizado entre os aminoácidos 759 e 897. Nas posições 275 a 544 foi ainda descrito provisoriamente um receptor de proteína cinase LRR-like (Genbank, www.ncbi.nlm.nih.gov).

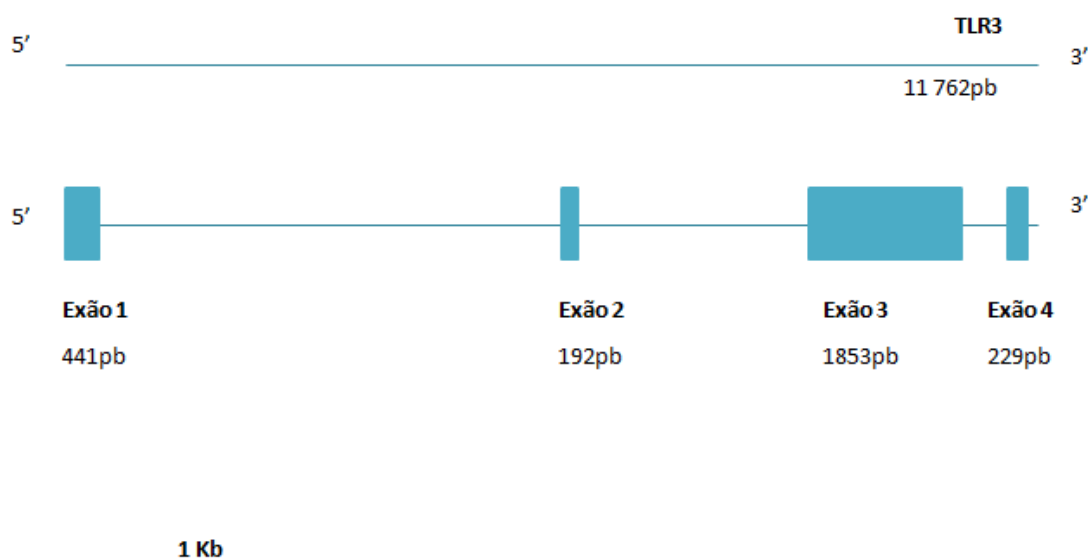


Figura 7. Estrutura molecular do gene TLR3 no *Oryctolagus cuniculus*. A linha superior representa o tamanho total do gene, a inferior corresponde às regiões intrónicas e as caixas azuis representam os exões

Uma das principais funções dos receptores *Toll-like* é reconhecer agentes patogénicos, estando alguns dos TLRs particularmente associados a respostas virais, como é o caso dos TLR3, 7 e 8. É interessante referir que no genoma do Coelho-bravo o gene TLR7 está ausente e o TLR8 é um pseudogene (Astakhova et al. 2009). O estudo do reconhecimento viral e das respostas por parte destes receptores é particularmente importante para a compreensão do processo de infecção viral. No Coelho-bravo o estudo destes genes é ainda mais relevante pois as populações têm sido fortemente depauperadas por duas doenças virais, a Doença Hemorrágica Viral (RHD) e a Mixomatose.

4. *Oryctolagus cuniculus*

4.1. História evolutiva

O *Oryctolagus cuniculus* (*O. c.*), vulgarmente designado de Coelho-bravo, pertence à Ordem *Lagomorpha*, que está dividida em duas famílias, a *Ochotonidae* e a *Leporidae*, que se terão separado há cerca de 35 a 40 milhões de anos (Matthee 2009). A família *Ochotonidae* é composta pelo género *Ochotona* (também designadas por Pikas). A família *Leporidae* compreende 11 géneros que incluem os Coelhos e as Lebres. Este grupo terá divergido há cerca de 12 milhões de anos (Matthee 2009). O género *Oryctolagus* é monoespecífico e de acordo com registos fósseis ter-se-á originado na Península Ibérica há cerca de 6,5 milhões de anos (Lopez-Martinez 1989). Está dividido em duas subespécies, *Oryctolagus cuniculus algirus* e *Oryctolagus cuniculus cuniculus*, morfológica e geneticamente distintas. A subespécie *Oryctolagus cuniculus algirus* ocorre no sudoeste da Península Ibérica, Açores, Madeira e Canárias (Biju-Duval et al. 1991, Monnerot et al. 1994) e a *Oryctolagus cuniculus cuniculus* está distribuída pelo nordeste de Espanha, Sul de França, Europa Ocidental, América do Sul, África, Austrália e em mais de 800 ilhas na sua maioria introduzidos pelo Homem (Flux 1983, 1992). Dados baseados nos níveis de divergência mitocondrial, loci ligados ao cromossoma X, variabilidade genética de aloenzimas e microssatélites sugerem que as populações de *Oryctolagus cuniculus* ficaram restritas a dois isolados populacionais (Branco et al. 2000, Branco et al. 2002, Geraldès et al. 2005, Queney et al. 2001, van der Loo et al. 1991, van der Loo et al. 1999), um no nordeste de Espanha que originou a subespécie *O. c. cuniculus* e outro no Sudoeste da Península Ibérica que originou a subespécie *O. c. algirus* antes de se formar uma zona de contacto bem delimitada (Figura 8). Estes grupos populacionais evoluíram separadamente durante um grande período de tempo, estimado em cerca de 2 milhões de anos (Biju-Duval et al. 1991, Geraldès et al. 2006). Vários estudos sugeriram uma expansão pós-glacial do Sudoeste para Norte e do Nordeste para Sul, com possíveis contactos secundários entre os dois grupos populacionais na zona central da Península Ibérica (Branco et al. 2000, Branco et al. 2002, Geraldès et al. 2005, Queney et al. 2001, van der Loo et al. 1991, van der Loo et al. 1999). Estudos de diversidade apoiados nos genes da hemoglobina, DNA mitocondrial, marcadores ligados ao cromossoma X, aloenzimas, proteínas plasmáticas e marcadores imunológicos relatam uma maior diversidade genética nas populações Ibéricas relativamente às populações de *O. c. cuniculus* do resto da Europa e de outras partes do mundo (Branco et al. 2002, Campos et al. 2008, Esteves et al. 2004, Geraldès et al. 2006). As populações a Norte dos Pirenéus perderam variabilidade relativamente às da Península

Ibérica provavelmente devido ao *bottleneck* associado à expansão pós-glacial das populações francesas de Coelho-bravo (Queney et al. 2001).

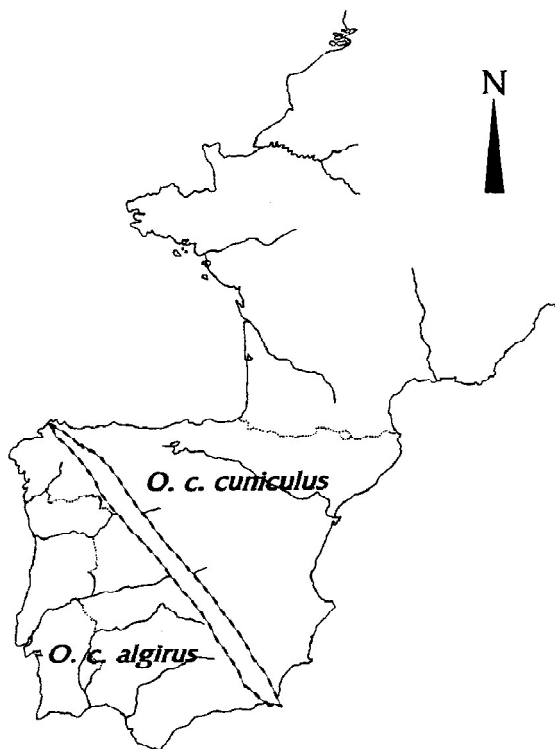


Figura 8. Separação das populações de *Oryctolagus cuniculus*. A Nordeste da Península Ibérica, a subespécie de *O. c. cuniculus*, e a Sudoeste, a subespécie de *O. c. algirus* (Esteves et al. 2006)

4.2. Características genómicas

No genoma dos eucariotas podem ser encontradas múltiplas cópias de diferentes sequências repetitivas intercaladas ao longo do genoma, que contribuem significativamente para o tamanho e possivelmente para a função desse genoma (Davidson and Britten 1973). Nos mamíferos estes elementos repetitivos têm sido designados de *short interspersed retrotransposable elements* (SINEs) quando o seu tamanho é inferior a 500pb, ou *long interspersed retrotransposable elements* (LINEs) com fragmentos de tamanho entre 6 e 8kb. Cada ordem de mamíferos possui uma ou mais famílias características de SINEs que estão presentes em grande número de cópias (Singer 1982). Nos Lagomorfos, particularmente no Coelho-bravo, essas famílias de SINEs são designadas por *C-repeats* que correspondem a cerca de 13% do seu genoma e parecem ter derivado de genes tRNA (Lawrence et al. 1985). Apresentam uma grande tendência para se inserirem dentro ou na proximidade de outros *C-repeats* e são

compostos por diferentes motivos, incluindo motivos semelhantes a RNA, um *core* conservado, repetições de CT e segmentos ricos em Adeninas (Krane et al. 1991).

Estes segmentos *C-repeats* podem ser agrupados em famílias distintas de acordo com as suas características. Estimativas acerca do tempo de divergência entre as várias famílias de *C-repeats* no *Oryctolagus cuniculus* apontam que o grupo mais antigo divergiu há cerca de 43 a 131 milhões de anos (média de 77 milhões de anos), o segundo grupo entre 22-59 milhões de anos (média de 37 milhões de anos) e por último, o grupo mais recente entre 8 e 42 milhões de anos (média 26 milhões de anos) (Krane et al. 1991). Os *C-repeats* que se inserem dentro de outros *C-repeats* pertencem a famílias que provavelmente se propagaram mais recentemente, logo, a divergência na família de *C-repeats* mais antiga é maior do que na mais recente (Krane et al. 1991).

4.3. Principais doenças virais

Durante as últimas décadas, as populações de Coelho-bravo têm sofrido um declínio acentuado devido à actuação de diversos factores, como por exemplo, a fragmentação e depauperamento do habitat, pressão cinegética excessiva (Le Gall et al. 1998, Stanford et al. 2007) e incidência de duas doenças virais: Doença Hemorrágica (RHD) – 1984 (Liu et al. 1984) e a Mixomatose – 1896 (Sanarelli 1898).

Doença Hemorrágica

A RHD era desconhecida até 1984 quando um grupo de Coelhos domésticos da estirpe Angorá foi importado da Alemanha para a província de Jiangsu na China (Liu. et al. 1984). A partir da China o vírus da doença hemorrágica (RHDV), disseminou-se por todo o mundo em Coelhos infectados ou produtos derivados (Kerr et al. 2009). Na Europa, a transmissão viral ocorreu dos Coelhos domésticos para os selvagens e está hoje em dia estabelecida nas populações selvagens por toda a Europa continental, Irlanda, Reino Unido e Austrália, com graves consequências ecológicas, particularmente na Península Ibérica (Anon 1989, Cooke 2002, Villafuerte et al. 1995, Villares et al. 1988). Aí, os surtos de RHD foram descritos pela primeira vez em 1988 e a doença é endémica desde então (Anon 1989, Villares et al. 1988). O número de indivíduos nas populações Ibéricas diminuiu cerca de 30% nessa altura, tendo levantado preocupações acerca da conservação da espécie o que conduziu a um impacto económico negativo na caça (Alves and Ferreira 2004, Moreno et al. 2007, Villafuerte et al. 1995). A recuperação dos efectivos populacionais em Portugal tem sido lenta levando à recente classificação do Coelho-bravo, de vulnerável no Livro Vermelho dos Vertebrados (ICN).

No que diz respeito ao vírus responsável por esta doença, parece ter surgido de mutações de um vírus não virulento já existente nas populações de Coelhos (Forrester et al. 2006, Nowotny et al. 1997). É um vírus de ssRNA com aproximadamente 7,4 kb, sentido positivo, pertencente à família *Caliciviridae*, género *Lagovirus* (Ohlinger et al. 1990, Parra and Prieto 1990). A partícula víral possui forma icosaédrica, sem envelope, capsíde com diâmetro de 35-40nm, composto principalmente por um grande polipeptídeo de 60 kDa. (Gorski et al. 1994, Smid et al. 1991).

RHD é uma condição aguda, altamente infecciosa e usualmente fatal que afecta Coelhos-bravos e domésticos. O período de incubação do vírus varia de 16 a 48h e os índices de morbilidade e mortalidade rondam os 90-100%. A morte ocorre na maioria dos casos 2 a 3 dias após a infecção, contudo, pode ocorrer mais tardiamente (Chasey 1997, Ohlinger et al. 1993, Smid et al. 1991). Os Coelhos com menos de 8 semanas de vida ficam infectados pelo vírus mas não manifestam a doença, podendo no entanto, disseminar o vírus que é relativamente resistente e persiste no meio ambiente (Cooke and Fenner 2002).

Os sinais característicos da doença são o aumento da temperatura corporal para 41°C ou mais, respiração rápida, observação de cianose, anorexia e prostração. Os sinais nervosos podem estar presentes nos estadios mais avançados, com movimentos agitados dos membros, ataxia e comportamento frenético com guinchos antes da morte (Chasey 1997). É observada perda de sangue espumado pelas narinas e mais raramente pela vagina. Ocasionalmente pode ocorrer diarreia ou obstipação (Chasey 1997, Marcato et al. 1991).

Não foram encontradas diferenças significativas a nível epidemiológico, clínico ou patológico entre o nível de susceptibilidade das duas subespécies de *O. cuniculus*, sugerindo assim que, ambas têm o mesmo grau de susceptibilidade ao RHDV e que o vírus é igualmente virulento em ambas as subespécies (Muller et al. 2009).

Mixomatose

A Mixomatose é uma doença debilitante causada pelo vírus do Mixoma (MV) que é membro do género *Leporipoxvirus* da família *Poxviridae*, subfamília *Chordopoxvirinae*. O seu genoma é constituído por dsDNA de grande comprimento, 161,8 kb que engloba 171 genes (Cameron et al. 1999).

Foi descrito pela primeira vez em 1896 na América do Sul como agente causador de uma infecção letal em Coelhos-bravos importados (Sanarelli 1898). Este vírus foi caracterizado como um patógeno natural do Taipeti da América do Sul (*Sylvilagus brasiliensis*), e do Coelho escova norte-americano (*Sylvilagus bruneii*) com estirpes

endémicas MV dessas áreas a serem designadas de linhagens do Brasil e da Califórnia, respectivamente (Fenner 2000). É de realçar que no *Sylvilagus*, esse vírus causa apenas uma infecção benigna, caracterizada por um fibroma cutâneo restrito ao local da inoculação. No entanto, nos Coelhos-bravos o contacto com o vírus, nos anos 50 do século XX, gerou uma infecção sistémica com elevadas taxas de mortalidade (Fenner 1983). Passados alguns anos, observou-se uma queda na mortalidade o que poderá estar relacionada com a adaptação vírus-hospedeiro, consistindo na evolução de estirpes virais atenuadas e na selecção natural dos Coelhos-bravos resistentes à infecção (Campos et al. 2008, Fenner and Ross 1994).

A gravidade do quadro clínico da doença pode variar com a estirpe de MV envolvido, bem como a linhagem de Coelhos infectados. Os sinais clínicos e as altas taxas de mortalidade associadas à mixomatose são resultado da disfunção de múltiplos órgãos, bem como das infecções bacterianas secundárias descontroladas do tracto respiratório que resultam da progressiva insuficiência da resposta imune do hospedeiro. Os sinais clínicos precoces dos tipos típicos de mixomatose são inchaço do focinho e região anogenital, bem como conjuntivite o que dá ao Coelho-bravo uma aparência leonina (Fenner 2000, Fenner and Ratcliffe 1965, Fields et al. 2001). Até o momento, são conhecidas pelo menos duas formas de mixomatose em Coelhos-bravos, a forma nodular ou mixomatose típica e a mixomatose atípica. A principal diferença reside na quase total ausência de sinais clínicos cutâneos e desconforto respiratório intenso observado apenas na mixomatose atípica (Marlier et al. 2000a, Marlier et al. 2000b, Stanford et al. 2007).

Hoje em dia, a mixomatose é endémica na América do Norte e do Sul, Europa e Austrália (DiGiacomo and Mare 1994, Stanford et al. 2007). Em grande parte da América do Sul, como Brasil e Uruguai, o vírus é endémico nos Coelhos do género *Sylvilagus*, particularmente *S. brasiliensis*. No entanto, em países como o Chile, o reservatório principal é o Coelho-bravo. Por outro lado, na América do Norte a estirpe Califórnia de MV é endémica em Coelhos americanos selvagens (*Sylvilagus*), particularmente *S. bachamani*, que actua como uma espécie de reservatório de infecção para Coelhos domésticos. Na Austrália, o vírus tem sido endémico nas populações de Coelho-bravo em estado selvagem desde a sua introdução em 1950 (DiGiacomo and Mare 1994, Stanford et al. 2007). Na Europa, a mixomatose disseminou-se nas populações de Coelho-bravo de vários países, após a sua primeira introdução em 1952. No entanto, outros Lagomorfos em que não ocorre mortalidade, tais como a Lebre (*Lepus europaeus* e *Lepus timidus*) e diversas espécies de *Sylvilagus* podem também servir como hospedeiros (DiGiacomo and Mare 1994, Stanford et al. 2007).

Objectivos

Com este trabalho pretende-se ajudar a colmatar a falta de informação do gene TLR3 em populações de Coelho-bravo de ambas as subespécies, *Oryctolagus cuniculus cuniculus* e *Oryctolagus cuniculus algirus*, comparar a organização molecular deste gene entre os mamíferos e avaliar a existência de padrões de selecção natural. Para isso, no estudo do gene TLR3, foram definidos os seguintes objectivos:

- Caracterizar molecularmente este gene no Coelho-bravo;
- Determinar a variabilidade genética das populações de Coelho-bravo da Península Ibérica através de sequenciação;
- Realizar em mamíferos, um estudo comparativo dos domínios conservados;
- Conhecer a filogenia do gene nos mamíferos;
- Averiguar a ocorrência de padrões selectivos nos mamíferos através da comparação de vários métodos específicos para a detecção de selecção natural;
- Aplicar e relacionar os padrões de selecção natural que ocorrem nos mamíferos com as características genéticas do gene no Coelho-bravo.

Material e Métodos

Neste trabalho, analisaram-se 40 amostras de *Oryctolagus cuniculus* provenientes de quatro localizações diferentes na Península Ibérica. Duas das populações pertencem à subespécie *O. c. cuniculus* (Saragoça e Tarragona) e as restantes duas populações à subespécie *O. c. algirus* (Mértola e Pancas) (Figura 9).

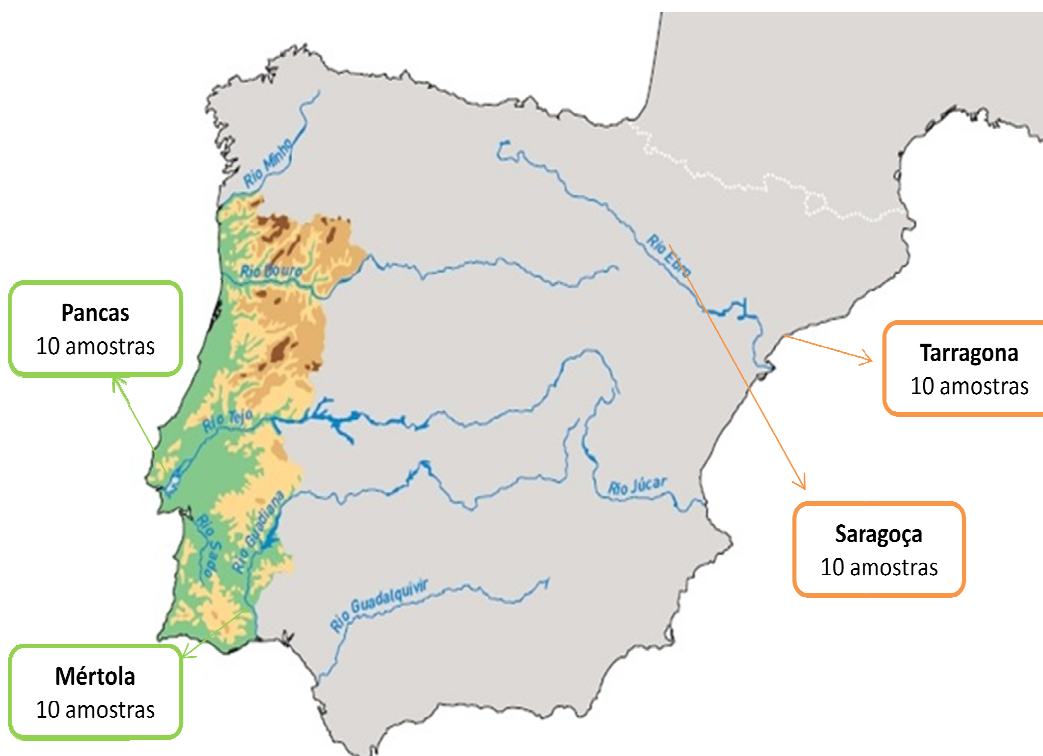


Figura 9. Localização geográfica das populações de Coelho-bravo da Península Ibérica analisadas neste trabalho.

Para estas amostras foi extraído DNA total utilizando o método fenol-clorofórmio (Chomczynski and Sacchi 2006). Para a amplificação do fragmento de DNA que continha o gene TLR3 foram desenhados *primers* a partir da sequência do GenBank® (Número de acesso: NM_001082219.1, <http://www.ncbi.nlm.nih.gov/gene>). Os *primers* usados estão listados na Tabela 1, onde para além da respectiva sequência, consta a sua localização no gene (Figura 10), temperatura de *annealing* (TA) e o tamanho de fragmento obtido após a amplificação. Procedeu-se à realização da reacção em cadeia da polimerase (PCR), utilizando a Taq DNA Polymerase (Taq PCR Master Mix Kit, Qiagen) num total de 30 ciclos de 95°C – 30 segundos (s), TA (Tabela 1) – 30s, 72°C – 45s, com uma desnaturação inicial de 15 minutos a 95°C e uma amplificação final de 20 minutos a 60°C.

Tabela 1. Características dos primers utilizados na amplificação do gene TLR3 no Coelho-bravo: identificação dos primers, sequência, posição (pb), temperatura de *annealing* (TA) (°C) e tamanho do fragmento amplificado (pb).

Exão	Identificação	Sequência do primer	Posição	TA (°C)	Tamanho do fragmento
1	OcTLR3_Fw1	5' ATGAGCCAGAGTGTGCCTTATC 3'	1pb	60°C	510pb
	OcTLR3_Rv1	5' CTTGAAGATTTAACATTGGATC 3'	510pb		
2	OcTLR3_Fw2	5' AGTGAAGTGACGATGTGCTG 3'	5949pb	58°C	453pb
	OcTLR3_Rv2	5' ACCAATACCCCAAGGCAGTA 3'	6402pb		
3	OcTLR3_Fw3	5' GGCCTCTTTCTGAACAATGC 3'	9057pb	58°C	643pb
	OcTLR3_Rv3	5' CCAATGTCAAGCACCTCTAG 3'	9700pb		
	OcTLR3_Fw4	5' CCTTACTCATACTCAACCTG 3'	9607pb	58°C	601pb
	OcTLR3_Rv4	5' GTCATCAAAGACAGATTGTGG 3'	10208pb		
	OcTLR3_Fw5	5' CTTGGATGAGATCCCAGAAG 3'	10103pb	56°C	765pb
	OcTLR3_Rv5	5' TTTGCATAATGGGTCTTTAG 3'	10868pb		
4	OcTLR3_Fw6	5' ATTGAGGAAAATACATAACCTGA 3'	11300pb	58°C	466pb
	OcTLR3_Rv6	5' TTATGATAGAATCCCGTCACC 3'	11766pb		

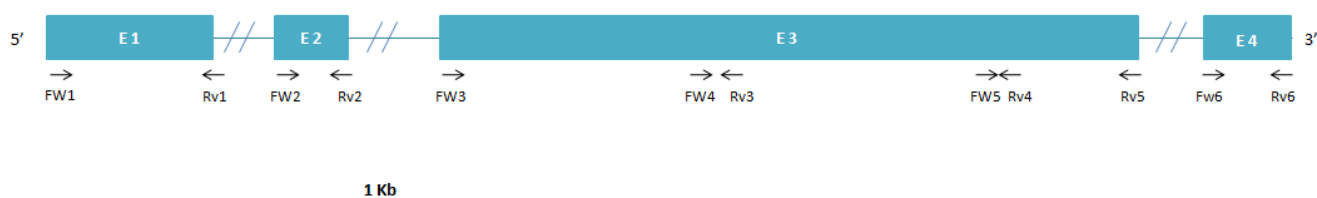


Figura 10. Representação esquemática da localização dos primers no gene TLR3 com representação dos exões (E1, E2, E3 e E4), e dos locais de homologia dos primers assinalados pelas setas.

Os produtos de PCR foram sujeitos a uma corrida electroforética num gel de agarose e os fragmentos obtidos foram visualizados por coloração com Gel Red™ (Biotium, E.U.A.). Para confirmar se o tamanho de pares de bases amplificados estava de acordo com o pretendido, comparou-se com o marcador de peso molecular, Marker 5 (Eurogentec, Bélgica). Em todas as amostras com presença da banda com o tamanho de pb pretendido no gel de agarose, foi efectuada a purificação dos produtos de PCR com a

adição de ExoSAP (Fermentas, Alemanha) e incubação a 37°C durante 15 minutos. De modo a assegurar a inactivação da enzima a reacção prolongou-se 15 minutos a 80°C.

A reacção de sequenciação foi efectuada com o kit BigDye® Terminator v3.1 Cycle Sequencing (Applied Biosystems, Foster City), com as seguintes condições de tempo e temperatura: 94°C – 3 minutos; 23 ciclos a 96°C-10s, 55°C – 5s, 60°C – 4 minutos. A purificação das sequências foi efectuada com passagem por coluna com Sephadex (Amersham, Inglaterra) e posteriormente procedeu-se à desidratação térmica e rehidratação com formamida. A análise dos electroferogramas foi efectuada no 3130xl Genetic Analyzer (Applied Biosystems, Foster City), e as sequências obtidas foram editadas, alinhadas e analisadas com o programa CLUSTAL W (Thompson et al. 1994) implantado no software BioEdit (Hall 1999).

Foi efectuada a análise dos domínios conservados no gene TLR3 de Coelho-bravo e de outros mamíferos listados na Tabela 2 utilizando o programa *Simple Modular Architecture Research Tool* (SMART) (Letunic et al. 2009, Schultz et al. 1998) (<http://smart.embl-heidelberg.de/>), isoladamente e em conjunto com a opção de procura de domínios Pfam (Finn et al. 2010) e através do programa LRRfinder (Offord et al. 2010) (www.lrrfinder.com/).

Foi pesquisada a existência de *C-repeats* nas regiões intrónicas do gene TLR3 de Coelho-bravo com o programa Censor (Kohany et al. 2006) (<http://www.girinst.org/censor/>). Os *C-repeats* identificados foram alinhados com a sequência consenso proposta por Krane et al. (1991) com outros 5 *C-repeats* analisados por este autor e classificados nos três grupos (I, II, e III), através do programa CLUSTAL W (Thompson et al. 1994) implantado no software BioEdit (Hall 1999). A identificação das sequências utilizadas encontra-se na Tabela 3. Para classificar os *C-repeats* encontrados na sequência do gene TLR3, foi construída uma árvore filogenética usando nucleotídeos e aminoácidos. Utilizaram-se os métodos de *Maximum likelihood* e *Neighbor-Joining* (Saitou and Nei 1987) disponíveis no software MEGA 4 (Tamura et al. 2007). O *Neighbor-Joining* foi utilizado com as opções *complete deletion* e modelo *p-distance*. Para analisar a consistência e robustez da árvore, utilizaram-se *bootstraps* obtidos com 1000 réplicas.

Para pesquisa dos locais de N-glicosilação no TLR3 de Coelho-bravo, bem como em genes ortólogos noutros mamíferos (Tabela 2), foi utilizado o programa Prosite Scan (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_prosite.html).

Paralelamente, realizaram-se estudos evolutivos, tais como, a filogenia do gene TLR3 nos mamíferos e a detecção de padrões de selecção natural. Para tal, foram retiradas as sequências do gene TLR3 de vários mamíferos das bases de dados GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) e Ensembl (<http://www.ensembl.org/index.html>) (ver Tabela 2), que foram compiladas e alinhadas

com o programa CLUSTAL W (Thompson et al. 1994) implantado no software BioEdit (Hall 1999). Estas sequências permitiram a construção de árvores filogenéticas usando o software MEGA 4 (Tamura et al. 2007) através do método de *Neighbor-Joining* (Saitou and Nei 1987) com a opção *complete deletion* e modelo *p-distance* e adicionalmente foi também usado o método de *Maximum likelihood*. Como *outgroup* foi utilizada a sequência de Opossum (*Monodelphis domestica*). Para avaliar a consistência e robustez da árvore, utilizaram-se *bootstraps* obtidos com 1000 réplicas.

Tabela 2. Identificação das sequências de mamíferos utilizadas na análise filogenética e na detecção de padrões de selecção natural.

Organismo	Número de acesso	Base de dados
Boi <i>Bos taurus</i>	NM_001008664.1	GenBank
Cão <i>Canis lupus familiaris</i>	XM_540020.2	GenBank
Cavalo <i>Equus caballus</i>	NM_001081798.1	GenBank
Chimpanzé <i>Pan troglodytes</i>	NM_001130470.1	GenBank
Coelho-bravo <i>Oryctolagus cuniculus</i>	NM_001082219.1	GenBank
Gato <i>Felis catus</i>	NM_001079829.1	GenBank
Golfinho <i>Tursiops truncatus</i>	ENSTTRG00000015619	Ensembl
Homem <i>Homo sapiens</i>	NM_003265.2	GenBank
Javali <i>Sus scrofa</i>	NM_001097444.1	GenBank
Macaco-rhesus <i>Macaca mulatta</i>	NM_001036685.1	GenBank
Musaranho <i>Sorex araneus</i>	ENSSARG00000011830	Ensembl
Opossum <i>Monodelphis domestica</i>	ENSMODG00000004861	Ensembl
Orangotango <i>Pongo abelii</i>	XM_002815350.1	GenBank
Ouriço-cacheiro <i>Erinaceus europaeus</i>	ENSEEUG00000008588	Ensembl
Ovelha <i>Ovis Áries</i>	NM_001135928.1	GenBank
Porquinho-da-índia <i>Cavia porcellus</i>	NM_001173029.1	GenBank
Raposa voadora <i>Pteropus vampyrus</i>	ENSPVAG00000005843	Ensembl
Ratazana <i>Rattus norvegicus</i>	NM_198791.1	GenBank
Rato <i>Mus musculus</i>	NM_126166.4	GenBank
Társio-das-Filipinas <i>Tarsius syrichta</i>	ENSTSYG00000009003	Ensembl

Tabela 3. Identificação das sequências de C-repeats utilizadas neste trabalho anteriormente classificadas por Krane et al. (2001).

C-repeat	Número de acesso	Início	Fim	Orientação	Localização	Base de dados
c3	X07786	8894	9206	3'→5'	5' γ globina	GenBank
c6	X07786	10866	11186	3'→5'	5' γ globina	GenBank
c8	X07786	18604	18936	5'→3'	5' γ globina	GenBank
c18	M35026	994	1996	5'→3'	5' ξ globina	GenBank
c19	M35026	1311	1655	3'→5'	5' ξ globina	GenBank

Para averiguação da ocorrência de padrões de selecção natural no gene TLR3 dos mamíferos (Tabela 2), foi efectuado um estudo comparativo de dois programas, de acordo com o estudo efectuado por Wlasiuk and Nachman (2010) em TLRs de primatas. Utilizou-se o programa CODEML implantado no pacote de programas PAML 4 (Yang 1997), e o servidor *online* DATA MONKEY (<http://www.datamonkey.org/>) (Delport et al. 2010, Pond and Frost 2005).

O programa PAML, opção CODEML, foi usado para estimar a razão de substituições nucleotídicas que provocam alteração de aminoácidos (Não-sinonímias) (d_N) com as substituições nucleotídicas que não provocam alterações de aminoácidos (Sinonímias) (d_S) ($\omega = d_N / d_S$). Assim, como a selecção natural opera maioritariamente ao nível proteico, as mutações d_N e d_S estão sob pressões selectivas diferentes e têm diferentes taxas fixas de mutação (Yang 2007). A comparação entre as taxas de substituição de d_N e d_S pode revelar a direcção e a força dos padrões de selecção natural a actuar nas proteínas (Kimura 1977, Miyata and Yasunaga 1980). Foi efectuada ainda a detecção de evolução molecular adaptada, segundo métodos de substituição de codões. A grande diferença deste modelo relativamente aos modelos de substituição de nucleotídeos ou aminoácidos, consiste no facto de o tripleto ser considerado a unidade evolutiva (Goldman and Yang 1994). O rácio ω mede a direcção e a magnitude da selecção nas mudanças de aminoácidos, com valores de $\omega < 1$, $= 1$ e > 1 , indicando selecção negativa, evolução neutra, ou selecção positiva, respectivamente. A selecção positiva é definida como a presença de codões onde $\omega > 1$ (Yang 2007). Comparou-se o ajuste de dois modelos específicos de local aos dados em análise – um modelo neutro que não admite selecção positiva ($\omega \leq 1$) e o modelo alternativo que admite selecção positiva ($\omega > 1$). Foram comparados os modelos M0 ao M3, M1a ao M2a, M7 ao M8 e M8a ao M8. Na primeira comparação o modelo M0 assume um rácio ω constante ao longo do gene, enquanto que M3 assume classes discretas com diferentes proporções nos locais com diferentes rácios ω . Na segunda comparação, o modelo neutro M1a, foi reformulado para permitir apenas duas classes locais para o ω : conservado $0 < \omega < 1$ e neutro $\omega = 1$. O modelo alternativo, M2a (modelo de selecção positiva) permite a terceira categoria de locais que tem um $\omega > 1$. Nas duas últimas comparações, o modelo neutro M7 assume que o valor de ω segue uma distribuição beta com 10 categorias, cada uma correspondendo a um distinto valor de ω que é sempre inferior a 1, enquanto o outro modelo neutro M8a (semelhante ao M7), permite uma classe extra de codões com $\omega = 1$. O modelo alternativo, M8 tem uma categoria extra com $\omega > 1$ (Yang et al. 2005, Yang et al. 2000). Para verificar a probabilidade de cada codão estar sob selecção positiva, os valores de logaritmo de *likelihood* para cada par de modelos foram comparados utilizando

o teste de *Maximum likelihood* (LRT). Os níveis de significância para estes LRTs foram calculados utilizando 4 graus de liberdade para a comparação entre os modelos M0-M3, 2 graus de liberdade nas comparações entre M1a-M2a e M7-M8, e um grau de liberdade para a comparação entre M8a-M8 (Tabela 4). Finalmente, o método Bayesiano Empírico foi aplicado para determinar quais os códons com maior probabilidade de estarem sob selecção positiva. Locais com probabilidade superior a 0,95 foram considerados sob evolução adaptativa (Nielsen and Yang 1998, Wong et al. 2004).

Tabela 4. Especificações usadas na análise com o software PAML

Comparação	2Δ1	Graus de liberdade	Nível de Significância
M0 vs. M3	452	4	<0,01
M1a vs. M2a	4	2	>0,05
M7 vs. M8	18	2	<0,001
M8a vs. M8	4	1	<0,05

Por outro lado, através do servidor *online* DATA MONKEY foram utilizados vários métodos de *Maximum Likelihood* designados de *Single Likelihood Ancestor Sequences* (SLAC), *Fixed-effect Likelihood* (FEL) e *Randon Effect Likelihood* (REL). O SLAC é baseado na reconstrução de sequências ancestrais e na contagem das alterações sinonímias e não sinonímias em cada codão da filogenia. O FEL estima a razão de substituições d_N/d_S por cada codão sem assumir à priori uma distribuição de razões d_N/d_S . Por último, o REL começa por atribuir valores de razões d_N/d_S e posteriormente infere sobre a razão de substituições para cada codão individualmente. Foram considerados como códons candidatos a evoluir sobre selecção os que apresentarem valores $p < 0,05$ para o modelo SLAC e FEL, e Factor de Bayes >95 para o modelo REL.

Foram considerados sob selecção positiva os códons identificados em pelo menos dois dos quatro métodos analisados neste trabalho (PAML, SLAC, FEL e REL).

Resultados

1. Caracterização da estrutura molecular do gene TLR3 do Coelho-bravo

Inicialmente, e de modo a completar a informação já conhecida para a estrutura e organização molecular do gene TLR3 em Coelho-bravo, recorreu-se ao programa SMART (Letunic et al. 2009, Schultz et al. 1998), (<http://smart.embl-heidelberg.de/>) para determinar os domínios conservados presentes neste gene. Para isso foi utilizada a sequência de nucleotídeos publicada no GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) com o número de acesso NM_001082219.1 e posteriormente traduzida para aminoácidos. O esquema obtido encontra-se representado na Figura 11-A, e a caracterização de cada um dos domínios está descrita na Tabela 5. A utilização deste programa permitiu a identificação dos domínios conservados e a sua localização no gene. Foram inicialmente identificados 16 domínios LRRs na região extracelular da molécula, um domínio LRR-CT, o domínio Transmembranar e o domínio TIR (Tabela 5). Esta pesquisa foi complementada com a ferramenta disponibilizada de pesquisa dos domínios Pfam (Finn et al. 2010) o que permitiu identificar mais 8 domínios LRR (Figura 11-B). Salienta-se a ausência do domínio LRR-NT em ambas as metodologias utilizadas. Em paralelo com a utilização do programa SMART, foi efectuada a pesquisa com o LRRfinder (Offord et al. 2010) (www.lrrfinder.com/), uma recente ferramenta de identificação de domínios LRRs em sequências proteicas com particular aplicação aos TLRs. Através deste programa, foram identificados 24 domínios LRRs, domínio LRR-NT, LRR-CT, Transmembranar e o domínio TIR (Figura 11-C), que estão indicados na Tabela 6.

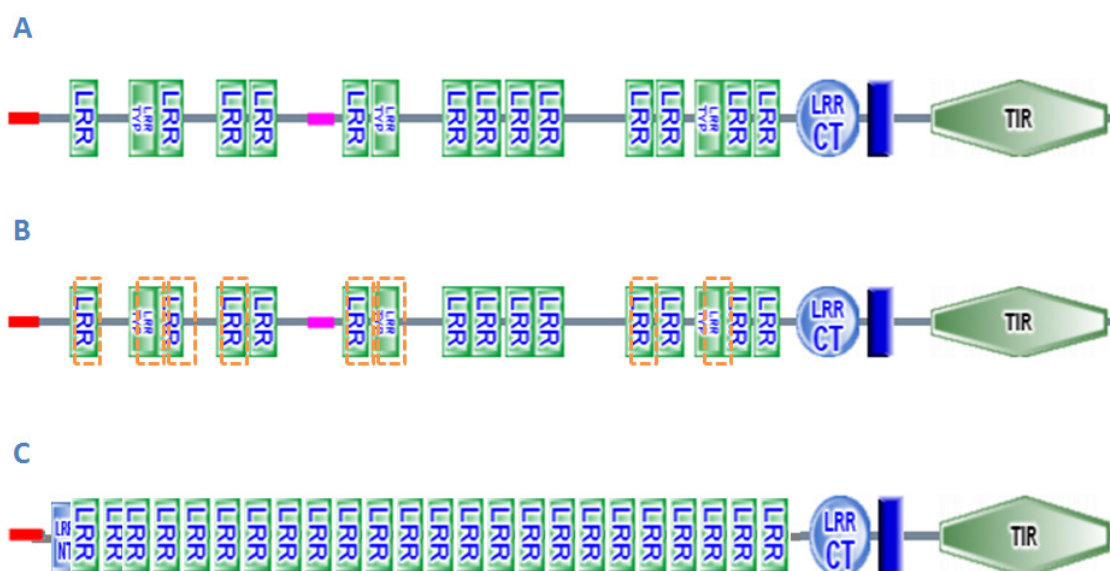


Figura 11. Representação dos domínios conservados no gene TLR3 de Coelho-bravo através do programa SMART (A), SMART + Pfam (B) e LRRfinder (C). No esquema B, os retângulos a tracejado laranja representam os 8 domínios encontrados através da opção de pesquisa de domínios Pfam existente no programa SMART.

Tabela 5. Caracterização dos domínios encontrados através do programa SMART em conjunto com a opção de pesquisa de domínios Pfam apresentados na Figura 11-A,B.






























Domínio	Início	Final	Representação
Sinal	1	26	
LRR	51	74	
LRR (Pfam)	53	75	
LRR_TYP (Típico)	99	122	
LRR (Pfam)	101	123	
LRR	123	144	
LRR (Pfam)	125	147	
LRR	171	196	
LRR (Pfam)	173	195	
LRR	197	220	
Baixa complexidade	246	267	
LRR	274	295	
LRR (Pfam)	276	298	
LRR_TYP (Típico)	298	321	
LRR (Pfam)	300	322	
LRR	355	378	
LRR	379	404	
LRR	407	430	
LRR	431	455	
LRR	506	529	
LRR (Pfam)	508	530	
LRR	530	553	
LRR_TYP (Típico)	562	585	
LRR (Pfam)	564	586	
LRR	586	609	
LRR	611	633	
LRR-CT	646	698	
Transmembranar	704	726	
TIR	756	901	

Tabela 6. Caracterização dos domínios encontrados através do programa LRRfinder e identificação das seqüências de aminoácidos detectadas em cada domínio.

A cinzento está assinalada a parte conservada de cada domínio LRR.

Domínio	Início	Final	Seqüência
Sinal	1	26	MSQSVPYHISLWGLLSLWMLCTSSG
LRR-NT	27	53	NKCTVRQEVAADCShlKLTQIPDDLPTN
LRR	54	77	ITVLNLTHNQIKRLPPANFTRYSR
LRR	78	101	LTVLDGGFNTISKLEPELCQSLPS
LRR	102	125	LKVLNLQHNELSQLSDKTFVFCMN
LRR	126	149	LTELHLMSNSIQKIQNNPFKNQKN
LRR	150	173	LIKLDLSHNGLKSTKLGTQLQLQN
LRR	174	199	LQELLSKNKIQALKSEELAFGNSS
LRR	200	223	LKKLELSSNLKEFSPGCFHAIGQ
LRR	224	250	LFGLFLNNAQLGPSLIEKLCLELSNTS
LRR	251	276	IQNLSLSNSQLNRISNRTLLGLKHTN
LRR	277	300	LTMLDLSHNNLNMIDDDSFVWLPH
LRR	301	324	LKYLFLYNNIVHLSNHFSYGLFN
LRR	325	357	VRHLNLKRSFPKQSVSLTSLPKIDDFSFQWLRC
LRR	358	381	LEYLNMEDNDIPSIKRNMFGLIN
LRR	382	409	LRYSLSNSFTNLRTLKNETFSSLAHSP
LRR	410	433	LLILNLTKNKISKIESDAFSWLGS
LRR	434	458	LEVLDIGINEIGQELTGQEWRLLEN
LRR	459	482	IFEIYLSYNKYLQLTSNSFALVPS
LRR	483	508	LQRLMLRRVALKNVDISPSPFHPLHD
LRR	509	532	LTILDLSNNLANINEEMLEGLEK
LRR	533	564	LEVLDLQHNLLARLWKQANPGGPVHFLKGLSH
LRR	565	588	LHTLNLESNGLDEIPEGVFKNLFQ
LRR	589	612	LKSINLALNNLNVLPQSVFDDQVS
LRR	613	637	LKSLNLQKNLITSVEKKVFGPVFKN
LRR	638	661	LSNLDMSFNPFDCESIAWVFNW
LRR-CT	646	702	NPFDCESIAWVFNWINKTRTNISELSSHLCNT PPQYHGFSVRLFDTSCKDSAP
Transmembranar	703	725	FELLMINTSFLLLFIFVLLIH
TIR	726	905	FEGWRISFYWNVSVHRILGFKEIDGQPEQFEYTAY IIHAYKDRDWWEHFSSMEEEDPSIKFCLEERDFQ AGVLGLEAIVNSIKRSRKIFVITQHLLKDPLCKR FKVHHAVQQAIEQNLDIILIFLENIPDYKLNHAL CLRRGMFKSRCIFYWPVHKERIEAFHHKLQVALGS RNSLH

Adicionalmente, o programa LRRfinder permitiu caracterizar outros parâmetros relevantes para o conhecimento da molécula de TLR3 de Coelho-bravo, como por exemplo, a determinação da sua massa molecular, 120,308Da, e o seu ponto isoeléctrico, 8,33. A análise de frequência de ocorrência dos aminoácidos ao longo dos 905 resíduos de comprimento da molécula evidenciou uma prevalência de leucinas, confirmando a alta frequência de LRRs observados nesta proteína. É de realçar a elevada frequência de aminoácidos hidrofóbicos (51,1%).

A análise dos domínios conservados também foi efectuada para outros mamíferos utilizando o programa SMART e o LRRfinder, o que possibilitou a comparação dos domínios existentes em cada uma das espécies (Anexo 1 e 2). Relativamente à análise efectuada com o programa SMART (Anexo 1), a ausência do domínio NT anteriormente referida na sequência de Coelho-bravo também foi observada nas sequências do gene TLR3 de Musaranho (*Sorex araneus*), Golfinho (*Tursiops truncatus*) e Ouriço-cacheiro (*Erinaceus europaeus*). Estes dois últimos mamíferos partilham ainda outras semelhanças com o *Oryctolagus cuniculus*, por exemplo a presença de 16 LRRs dos quais 3 pertencem à família dos LRRs típicos. Para além da ausência deste domínio na sequência de TLR3 de Coelho-bravo, todos os outros domínios são comuns aos mamíferos analisados, com variações em relação ao número de LRRs de espécie para espécie. O número de LRRs presentes no domínio extracelular do TLR3 deste grupo de mamíferos varia entre 15 e 18, com o número de LRRs típicos a oscilar entre 1 e 4. Todos os genes observados têm presentes os domínios LRRs, LRR-CT, Transmembranar e TIR.

O programa LRRfinder apresenta uma constrição para a sua utilização, pois apenas identifica os motivos LRR em sequências que não estejam previamente inseridas e identificadas na sua base de dados. Por isso, não foi possível identificar os domínios LRR-NT, LRR-CT, Transmembranar e TIR em sete espécies de mamíferos: Cão, Golfinho, Raposa voadora, Musaranho, Orangotango, Ouriço-cacheiro e Társio-das-Filipinas. Deste modo, relativamente ao domínio LRR-NT, foi possível detectar a sua presença no Coelho-bravo (sequência presente na base de dados do programa), mas não foi possível confirmar a ausência/presença para Golfinho, Musaranho e Ouriço-cacheiro. Na análise da sequência de Ratazana foi observado uma incongruência entre os dois programas, apenas no programa SMART foi encontrando um domínio Transmembranar anterior ao LRR-NT. O número de LRRs encontrados no programa LRRfinder foi superior ao detectado no SMART, variando entre 21 e 24, sendo que a maior parte das sequências analisadas (14/19) apresentou 24 destes domínios. Também no Coelho-bravo, se detectaram 24 domínios, em claro contraste com os apenas 16 domínios LRRs observados na análise efectuada com o programa SMART. No entanto,

para esta espécie, a comparação entre os domínios identificados por ambos os programas mostrou que não existem diferenças significativas na localização da região sinal e na localização dos domínios LRR-CT, Transmembranar e TIR.

O domínio TIR da proteína TLR3 é um dos mais bem caracterizados, apresentando três regiões importantes para a sinalização celular que são altamente conservadas entre os mamíferos. A comparação deste domínio entre o Homem e o Coelho-bravo mostrou a ocorrência de três substituições aminoacídicas, uma Alanina por uma Treonina na região 1, um Ácido Glutâmico por uma Lisina na região 2 e na região 3 uma Arginina por uma Histidina (Figura 13).

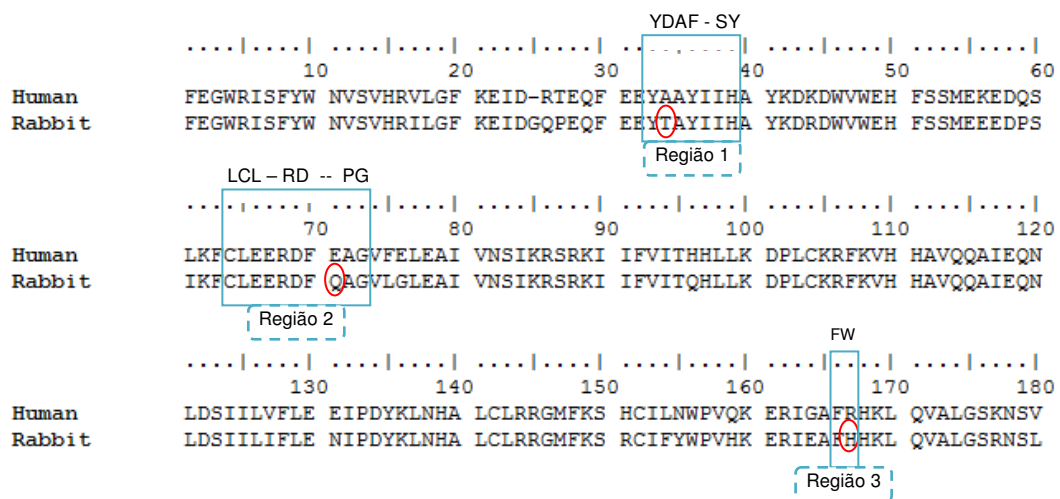


Figura 12. Alinhamento da sequência do domínio TIR do Homem e Coelho-bravo. As regiões conservadas estão assinaladas pelas caixas azuis. A vermelho estão assinalados os aminoácidos da sequência de Coelho-bravo que diferem da sequência de Homem (dentro de cada região). A sequência consenso para cada região encontra-se acima de cada caixa azul.

É conhecida a inserção de sequências repetitivas no genoma do Coelho-bravo, denominadas *C-repeats*. Neste trabalho foi utilizado o software Censor (Kohany et al. 2006) (<http://www.girinst.org/censor/index.php>) para verificar a sua existência no gene TLR3. O programa identificou 5 *C-repeats* com tamanhos de 287pb, 347pb, 301pb, 294pb, 332pb, distribuídos pelos 3 intrões. O CR1 e CR2 encontram-se entre o primeiro e o segundo exão, o CR3 e CR4 por sua vez estão entre o exão dois e três, e por último o CR5 encontra-se entre o exão três e quatro. A sua localização no gene está esquematizada na Figura 14. A sua inserção ocorreu em ambos os sentidos. Dois dos *C-repeats* foram inseridos no sentido 5'→3' (CR3 e CR5), e os restantes foram inseridos na direcção contrária, 3'→5' (CR1, CR2 e CR4).

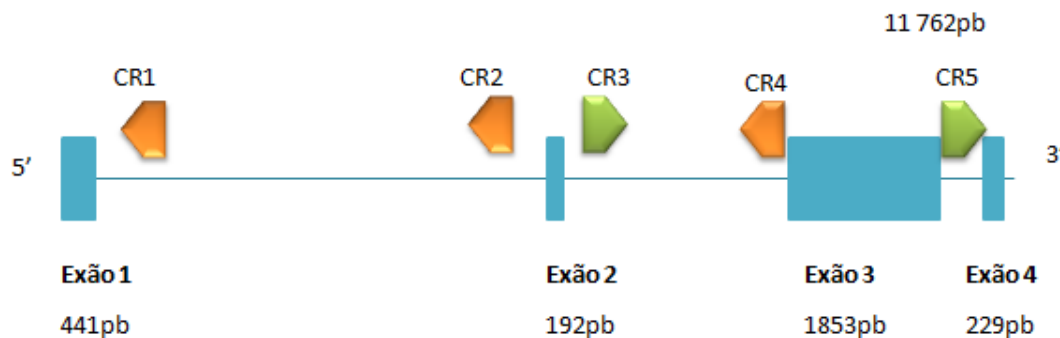


Figura 13. Localização dos C-repeats na sequência genómica do TLR3.

A verde estão assinalados os exões que se inseriram no sentido 5'-3' e a cor de laranja os que se inseriram no sentido 3'-5'.

Foi efectuado o alinhamento destas 5 sequências com a sequência consenso proposta por Krane et al. (2001) e com outros C-repeats já classificadas em cada um dos três grandes grupos (Krane et al. 1991) que se terão inserido no genoma em diferentes alturas do processo evolutivo dos Lagomorfos. O alinhamento pode ser observado na figura seguinte (Figura 15). Estas sequências permitiram a construção de uma árvore filogenética que se encontra na Figura 16 através dos métodos *Neighbour-joining* e *Maximum likelihood*. Como se obtiveram padrões filogenéticos semelhantes apenas se apresenta a árvore filogenética obtida com o método *Neighbour-joining*.



Figura 14. Alinhamento dos *C-repeats* encontrados na sequência de TLR3 de Coelho-bravo (CR1 a CR5) com a sequência consenso (ConC) e os *C-repeats* já classificados (c3, c6, c8, c18 e c19) (Krane et al. 1991). No alinhamento, os pontos significam que a base é a mesma da sequência consenso e os traços são espaços introduzidos nas sequências para se conseguir o alinhamento. A- Adenina, C – Citosina, G- Guanina, T- Timina, S=C+G, R=A+G, Y=C+T, M=A+C, K=T+G. A cinzento estão assinalados os locais diagnósticos identificados por Krane et al. (2001) para a classificação dos *C-repeats*.

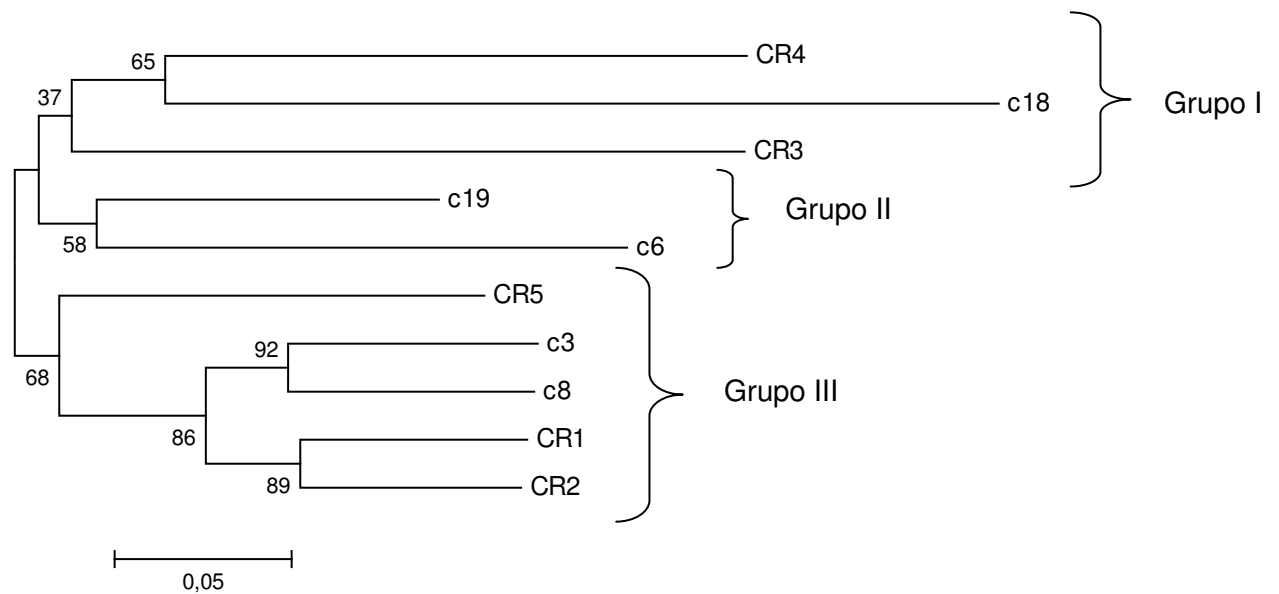


Figura 15. Árvore filogenética de classificação dos *C-repeats* com base nas sequências de nucleotídeos. CR1 a CR5 correspondem aos *C-repeats* identificados na sequência do TLR3 de Coelho-bravo. De acordo com Krane et al. (1991), a sequência c18 pertence ao grupo I, a c6 e a c19 ao grupo II e a c3 e c8 pertencem ao grupo III.

Nesta árvore filogenética pode-se observar o agrupamento das sequências de acordo com a sua similaridade genética o que permite a sua classificação em diferentes grupos. Deste modo, as sequências CR3 e CR4, identificadas no gene TLR3 de Coelho-bravo, estão no mesmo grupo que a c18 sendo por isso classificada no grupo I (Krane et al. 1991). As restantes três sequências, CR1, CR2 e CR5, agrupam com as sequências c3 e c8 que se inseriram mais recentemente no genoma dos Lagomorfos e que pertencem ao grupo III.

2. Caracterização genética do gene TLR3 nas populações de Coelho-bravo.

As sequências obtidas nas 40 amostras de Coelho-bravo estudadas neste trabalho foram alinhadas com a sequência publicada no GenBank (número de acesso NM_001082219.1). A sua comparação permitiu a identificação de 41 alterações nucleotídicas (Tabela 7). Essas substituições, estavam associadas a 15 alterações de aminoácidos. A identificação da região do TLR3 em que foram identificadas estas alterações de aminoácidos está presente na Figura 16. Para além disso, foi caracterizada a sua polaridade, carga e possibilidade de estar envolvido num sítio de N-glicosilação (Anexo 3).

Os resultados sugerem uma maior diversidade genética nas populações *O. c. algirus*, pois, das 41 posições polimórficas observadas, 36 delas variaram nas

populações de *O. c. algerus*, enquanto nas populações de *O. c. cuniculus*, apenas 23 locais de variação foram observados.

Embora não tenha sido encontrada nenhuma posição diagnóstica, foram encontradas 18 posições polimórficas exclusivas das populações de *O. c. algerus* que dão origem a 5 alterações de aminoácidos e 5 posições polimórficas exclusivas da população de *O. c. cuniculus* onde todas originam alteração de aminoácido. Dos 41 nucleotídeos com variação, 23 são raros (frequência inferior a 12,5% das amostras). Em todas as posições polimórficas foram observados indivíduos heterozigóticos, excepto na posição 136 (A/G), localizada no domínio-NT, que origina uma alteração aminoacídica I46V, no entanto, o nucleotídeo G ocorreu apenas numa amostra de *O. c. algerus*.

Relativamente às amostras de *O. c. algerus* destacam-se as posições nucleotídicas onde a variação ocorre exclusivamente nesta população: G243A, C246Y, C408A, A618M, A632R, C781Y, C822Y, A899M, G1059S, G1074A, G1281R, C1299G, T1439Y, T1665Y, A1953R, G2124R, C2409M e T2505C. É interessante realçar que em 13 das posições anteriormente referidas (246, 618, 632, 781, 822, 899, 1059, 1281, 1439, 1665, 1953, 2124 e 2409) não foram encontrados indivíduos homozigóticos para um dos nucleotídeos. As 5 alterações de aminoácidos que acontecem nesta população são originadas pelas alterações observadas nas posições nucleotídicas 632, 899, 1059, 1439, 2124 que originam as substituições amonoacídicas K211R, H300P, Q353H, V480A, M708I, respectivamente. Nestas, apenas na posição 300 ocorre a substituição de aminoácidos com propriedades diferentes, hidrofílico de carga positiva (Histidina), por um hidrofóbico (Prolina). Todas as alterações de aminoácidos nesta população ocorrem em domínios LRRs, excepto no aminoácido 708 que se encontra no domínio Transmembranar.

Por outro lado, na população de *O. c. cuniculus*, as posições polimórficas exclusivas são: A136G, T395C, C892Y, T943Y e A1678R das quais, nas últimas três não foi observada homozigotia para um dos nucleotídeos. Estas alterações nucleotídicas reflectem-se respectivamente nas alterações aminoacídicas I46V, M132T, L298F, S315P, K560E. Em duas das posições verificou-se alteração da polaridade, mas não da carga dos aminoácidos envolvidos. Assim, na posição 132 houve substituição de um resíduo hidrofóbico (Metionina) por um hidrofílico (Treonina) e na 315, registou-se a alteração de um aminoácido hidrofílico (Serina) por um hidrofóbico (Prolina). Todas as alterações ocorrem nos domínios LRRs, excepto a alteração já referida no domínio-NT, I46V.

Relativamente a posições polimórficas comuns às populações de *O. c. cuniculus* e *O. c. algerus*, verificaram-se 17 posições nucleotídicas das quais 5 dão lugar às seguintes alterações de aminoácidos: L101S, E204D, L575F, L716F e A864S. Três destas

alterações estão localizadas em domínios LRRs (L101S, E204D, L575F), uma delas está no domínio Transmembranar (L716F) e outra no domínio TIR (A864S).

É de realçar que nas posições nucleotídicas 1143 e 1144, onde grande parte das amostras apresenta um padrão heterozigótico (C/T) em ambas as posições, nos indivíduos *O. c. algirus*, dos quatro haplótipos possíveis, detectaram-se em homozigótia os haplótipos CT, TC e CC, e nos indivíduos *O. c. cuniculus*, detectaram-se apenas dois desses haplótipos CT e TC. O resultado obtido nos *O. c. cuniculus* parece indicar que existem duas linhagens distintas em desequilíbrio gamético, os haplótipos CT e TC terão sido os parentais e não terão sofrido recombinação entre si, pois não são observados os haplótipos recombinantes, CC e TT. No entanto, para confirmar esta observação terão que ser clonados todos os indivíduos heterozigóticos de modo a determinar os seus haplótipos. Nenhuma destas alterações origina substituição de aminoácidos.

A posição nucleotídica 1725 apresentou uma grande variabilidade, tendo sido a única onde se observaram as quatro bases, que dão origem aos codões TTA, TTC, TTG, TTT. No entanto, apenas dois aminoácidos são observados, Fenilalanina codificado por TTC ou TTT, e Leucina codificado por TTA ou TTG. Realça-se ainda a posição nucleotídica 1911, onde na população de *O. c. algirus* existe uma base fixa, Citosina, enquanto na população de *O. c. cuniculus* se observam tanto Timinas como Citosinas, no entanto não existe alteração a nível aminoacídico.

As alterações de aminoácidos detectadas neste trabalho foram analisadas com o programa SMART e LRRfinder e não se verificaram alterações nas delimitações dos domínios anteriormente identificados.

Tabela 7. Caracterização das alterações de nucleotídeos nas 40 amostras analisadas.

Sub-espécie	Exão	1							2				3										4																						
	Posição	90	136	243	246	302	303	395	408	585	612	618	632	781	810	822	892	899	943	1059	1074	1143	1144	1281	1299	1439	1606	1626	1656	1665	1678	1725	1911	1953	2124	2148	2208	2229	2409	2505	2523	2590			
Algirus	Amostra/NF	G	A	G	C	T	G	T	C	C	G	A	A	C	T	C	C	A	T	G	G	C	T	G	C	T	T	T	C	T	A	G	T	A	G	G	T	C	C	T	C	G			
	1	T	Y	C	C	.	.	.	C	.	.	T	C	T	.	.	.			
	2	C	.	.	.	T	.	M	.	.	C	Y	Y	.	S	.	.	C	Y	Y	.	.	A	C	R	R	K	Y	Y	.	.	Y	K		
	3	C	A	T	C	C	C	.	.	.	A	C	
	4	C	C	.	.	.	K	C	.	.	T	C	T	
	5	C	.	.	.	T	.	M	R	.	C	Y	Y	.	S	.	.	Y	Y	.	.	.	K	C	.	.	K	Y	T		
	6	Y	M	.	.	C	C	.	.	.	C	C	.	Y	.	K	C	.	.	.	Y	T		
	7	T	.	M	R	.	C	Y	C	C	.	.	T	C	.	T	C	T	
	8	C	Y	C	.	.	M	Y	.	.	.	Y	.	Y	C	C	.	.	C
	9	C	A	.	A	Y	C	C	.	.	.	K	C	.	.	T	C	T	
	10	C	A	.	.	T	.	.	R	.	C	Y	.	S	.	C	C	.	.	.	K	C	.	.	K	Y	T	.	Y	.	.		
	11	.	.	R	.	C	A	.	M	T	.	M	.	.	C	C	Y	Y	.	Y	C	.	.	K	.	.	M	.	.	K			
	12	.	.	A	.	C	A	.	A	T	.	M	.	.	C	R	Y	Y	C	C	.	.	.	K	C	.	.	K	Y	Y		
	13	T	C	Y	C	.	G	.	.	Y	Y	Y	.	.	K	C	Y	.	Y	.	.		
	14	C	A	.	A	C	Y	.	.	S	.	.	Y	Y	Y	Y	.	.	K	C	T	.	C	.	.		
	15	S	Y	S	Y	Y	Y	.	.	C	.	R	Y	.	.		
	16	C	A	.	.	T	.	M	.	.	Y	.	.	.	R	Y	C	.	S	.	.	Y	Y	.	.	.	K	C				
	17	C	T	C	.	.	.	C	Y	Y	C	R	S	.	C	C	.	.	.	K	C	T	.	Y	.	.			
	18	C	A	M	R	C	.	S	.	Y	Y	Y	.	.	K	C	M	.	.	.				
	19	.	.	A	.	C	A	.	A	T	.	M	R	.	C	Y	Y	.	S	.	Y	Y	Y	.	.	K	C	.	.	K	Y	Y			
20	.	.	.	Y	C	A	.	.	Y	C	.	.	.	S	.	.	Y	Y	.	S	.	.	.	T	.	.	K	C	Y	.				
Cuniculus	21	Y	Y	Y	K		
	22	Y	Y	Y	
	23	C	G	T	Y	S		
	24	T	Y	T	C	.	.	Y	Y	.	.	C	Y	.	.	K	Y	Y			
	25	T	C	T	C	.	.	C	C	.	.	S	C	.	.	T	C	T			
	26	Y	T	C	
	27	Y	Y	Y	Y	.	.	K	C	.	.	K	.	Y	.	.	T	K			
	28	Y	.	.	T	Y	Y	S	
	29	C	Y	Y	Y	Y	C	Y	.	.	.	Y	.	.	K	Y	Y			
	30	Y	Y	Y	Y	.
	31	Y	.	.	.	Y	Y	.
	32	Y	Y	Y	Y	Y	
	33	Y	R	.	T	C	.	.	.	C	Y	Y	C	Y	.	.	.	K	Y	.	.	K	.	Y	
	34	Y	Y	Y	
	35	Y	R	.	Y	S	.	.	.	Y	Y	Y	Y	.	Y	.	R	T	C	.	.	K	.	Y	.	.	Y	.			
	36	Y	Y	.	Y	.	.	K	Y	T	.
	37	Y	Y	Y	Y	.	.	.	T	Y	.	.	K	Y	Y	Y	.
	38	C	K
	39	Y	R	.	Y	S	.	.	.	Y	Y	Y	C	Y	.	.	T	Y	.	.	K	.	Y	Y	.
	40	C	.	Y	.	.	T	Y	Y	.

A=Adenina, C=Citosina, G=Guanina, T= Timina, S=C+G, R=A+G, Y=C+T, M=A+C, K=T+G. A **negrito** estão assinaladas as alterações de nucleotídeo que se reflectem em alteração de aminoácido. Os pontos correspondem a conservação do nucleotídeo.

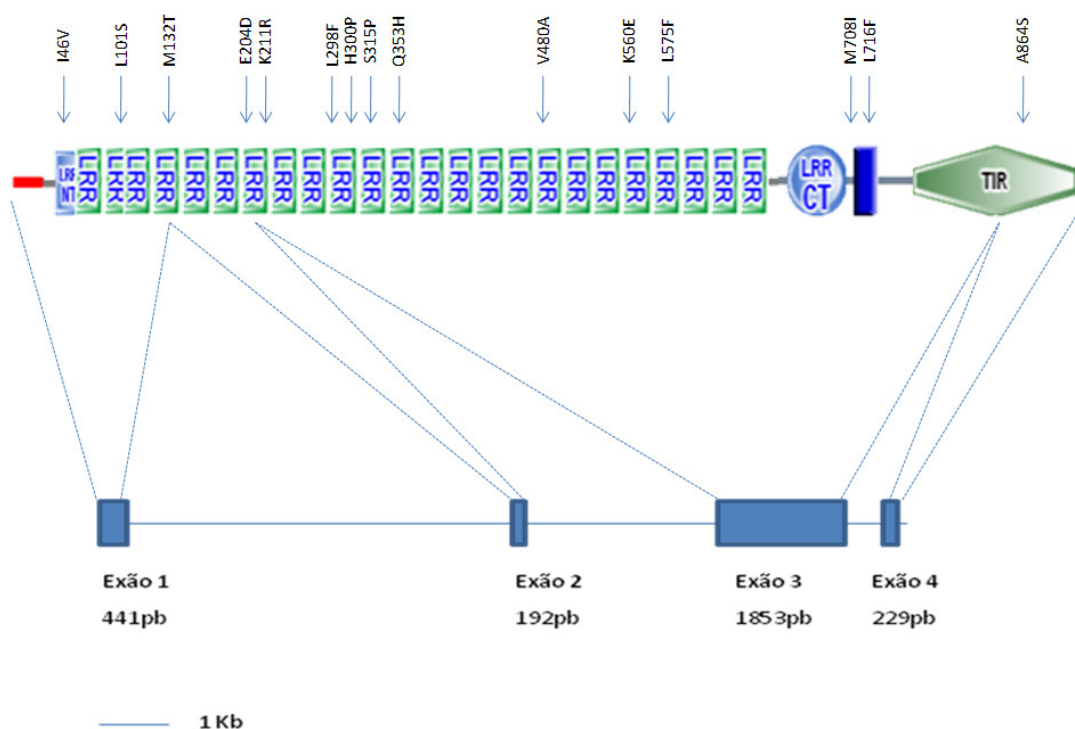


Figura 16. Localização das alterações de aminoácidos nos domínios conservados identificados pelo programa LRRfinder na sequência de TLR3 de Coelho-bravo. Correspondência entre os domínios e as regiões codificantes do gene.

3. Caracterização dos locais de N-glicosilação do TLR3

Para detectar possíveis locais de N-glicosilação recorreu-se ao programa Prosite Scan (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_prosite.html) Os resultados obtidos para o Coelho-bravo encontram-se representados na tabela seguinte (Tabela 8).

Fazendo uma análise dos possíveis locais de N-glicosilação presentes na molécula de TLR3 de Coelho-bravo, podem-se verificar 17 locais, dos quais 15 se localizam do domínio extracelular. A análise da sequência de TLR3 de Homem, com o mesmo *software* (Anexo 4), permitiu detectar também 17 locais de N-glicosilação, no entanto, dois locais diferem entre a sequência de Coelho-bravo e Homem. O Coelho-bravo perdeu o local correspondente às posições aminoacídicas 288-291, e ganhou um local entre as posições aminoacídicas 312-315, que aliás, é de entre os mamíferos analisados, exclusivo para esta espécie. Perdeu ainda o sítio de glicosilação 504-507 e ganhou o 664-667, que é comum à maioria dos mamíferos analisados.

Tabela 8. Caracterização dos possíveis locais de N-glicosilação da sequência do gene TLR3 de Coelho-bravo. A negrito estão realçados os locais da sequência de Coelho-bravo que diferem do Homem.

Início	Final	Sequência	Domínio Provável (LRRfinder)
51	54	NITV	LRR-NT
56	59	NLTH	LRR
69	72	NFTR	LRR
123	126	NLTE	LRR
193	196	NSSL	LRR
244	247	NTSI	LRR
249	252	NLSL	LRR
262	265	NRTL	LRR
272	275	NLTM	LRR
312	315	NHSF	LRR
395	398	NETF	LRR
410	413	NLTK	LRR
633	636	NLSN	LRR
659	662	NKTR	LRR
664	667	NISE	LRR-CT
706	709	NTSF	Transmembranar
732	735	NVSV	TIR

4. Padrões de selecção natural no gene TLR3 em mamíferos

Foi encontrada evidência de selecção positiva no gene TLR3 utilizando o método de *Maximum Likelihood* instalado no programa informático PAML 4 (Yang 1997). Foram comparados os modelos que incorporam ou proíbem a selecção positiva (Tabela 9). Observou-se que os modelos que permitem selecção positiva adequam-se melhor aos dados em estudo. Todos os modelos que permitem selecção positiva têm uma classe de codões com valores de $\omega > 1,0$. Assim, foram avaliados os aminoácidos mais prováveis de estarem sobre selecção positiva, com uma probabilidade posterior de 95%, apresentados na tabela seguinte. Adicionalmente, usando o servidor DATA MONKEY através dos modelos SLAC, FEL e REL, determinou-se também os codões que estariam sob selecção positiva.

Tabela 9. Sumário dos parâmetros estimados e valores de *likelihood* dos diferentes modelos de evolução para o TLR3 de Coelho-bravo.

Métodos	PAML							Data Monkey		
	M0: 1 ^a	M3:5 ^a	M1a:1 ^a	M2a:3 ^a	M7:2 ^a	M8a 3 ^a	M8: 4 ^a	SLAC	FEL	REL
Log likelihood	-18499,9	-17997	-18023,98	-1809.67	-18013	-1799.81	-17995.08	-18354.7	---	---
Parâmetros estimados	$\omega=0,09$	$p_0=0,54;$ $p_1=0,33;$ $(p_2=0,12);$ $(\omega_0=0,04);$ $(\omega_1=0,39);$ $\omega_2=1,26$	$p_0 = 0,75;$ $(p_1 =0,25);$ $(\omega_0=0,09);$ $(\omega_1 = 1)$	$p_0 = 0,75;$ $p_1 = 0,24;$ $(p_2=0,01);$ $(\omega_0=0,10);$ $(\omega_1 = 1);$ $\omega_2=2,45$	$p =0,10;$ $q=0,9$	$p_0= 0,80;$ $(p_1 =0,20);$ $p=0,06;$ $q = 1;$ $(\omega =1)$	$p_0= 0,91;$ $(p_1=0,09);$ $p=0,04;$ $q = 1,5;$ $\omega=1,59$	$p=0,02$	$p=0,02$	---
Codões sob seleção positiva	NP	NP	NP	<u>3, 4, 12,</u> <u>16, 79, 86,</u> <u>97, 120,</u> <u>185, 258,</u> <u>261, 266,</u> <u>269, 273,</u> <u>285, 295,</u> <u>473, 506,</u> <u>663, 749</u>	NP	NP	<u>3, 4, 12, 16,</u> <u>21, 79, 86,</u> <u>97, 120,</u> <u>184, 185,</u> <u>258, 261,</u> <u>262, 266,</u> <u>269, 273,</u> <u>285, 295,</u> <u>411, 434,</u> <u>456, 473,</u> <u>506, 586,</u> <u>663, 712,</u> <u>749, 780</u>	<u>4</u>	<u>4, 12,</u> <u>69, 79,</u> <u>258,</u> <u>285,</u> <u>456,</u> <u>712,</u> <u>715,</u> <u>780</u>	<u>12, 79,</u> <u>258,</u> <u>285,</u> <u>712</u>

Os números sublinhados são evidência de selecção positiva, em pelo menos dois dos modelos analisados;
 a Número de graus de liberdade na distribuição ω ;
 b Valores entre parênteses não são parâmetros livres;
 c Parâmetros p e q são parâmetros de forma da distribuição beta subjacentes ao modelo M7 e M8;
 d Codões sobre seleção positiva a um nível de confiança de 95%;
 NP Não Permitido

Utilizando a mesma abordagem de Wlasiuk and Nachman (2010), foi considerado que um codão estaria sob selecção positiva se fosse detectado em pelo menos dois dos quatro métodos considerados neste trabalho (PAML, SLAC, FEL e REL). Assim, foram detectados, com 95% de confiança, 24 codões (3,4,12,16, 79, 86, 97, 120, 185, 258, 261, 266, 269, 273, 285, 295, 456, 473, 506, 586, 663, 712,749, 780). Através do programa PAML, foram detectados 29 codões sob selecção positiva, no entanto, em conjunto com o programa DATA MONKEY, este número reduziu para 24.

Como existem diferenças entre as delimitações dos domínios da molécula TLR3 (Ectodomínio, domínio Transmembranar e domínio TIR), entre as espécies de mamíferos, utilizou-se a estrutura do Coelho-bravo para a localização dos codões sob selecção positiva. Assim, na sequência de Coelho-bravo, no domínio Transmembranar apenas 1 codão se encontrava nesta situação, no domínio TIR foram identificados dois codões e os restantes 21 estão localizados no Ectodomínio.

A caracterização das posições aminoacídicas onde foi detectada selecção positiva nas sequências de mamíferos do gene TLR3 está presente no Anexo 5.

5. Análise filogenética do gene TLR3

As vinte sequências do gene TLR3 obtidas em mamíferos e presentes nas bases de dados disponíveis no Genbank (<http://www.ncbi.nlm.nih.gov/genbank/>) e no Ensembl (<http://www.ensembl.org/index.html>), foram alinhadas e analisadas com o programa MEGA 4 (Tamura et al. 2007). Essa análise permitiu a construção de uma árvore filogenética usando nucleotídeos e aminoácidos. Foram utilizados os métodos *Neighbour-joining* e *Maximum likelihood*. Como se obtiveram padrões filogenéticos semelhantes apenas se mostra a árvore filogenética obtida com o método *Neighbour-joining* (Figura 18). Como *outgroup* foi usada a sequência de um mamífero marsupial, *Opossum*. A filogenia obtida não diferiu da comumente aceite como representativa da filogenia dos mamíferos, com o agrupamento das sequências das espécies de acordo com os grandes grupos, Artiodáctilos, Carnívoros e Primatas (Prasad et al. 2008, Sims et al. 2009). A posição do Coelho-bravo com um roedor é expectável uma vez que os Lagomorfos juntamente com os roedores formam a superordem Glires (Asher et al. 2005). O aparecimento de outras duas espécies roedoras, *Rattus norvegicus* e *Mus musculus*, como basais dos mamíferos placentários e separadas do Coelho-bravo pode ser explicado por uma já documentada aceleração da taxa de mutação nessas espécies (Nei and Glazko 2002).

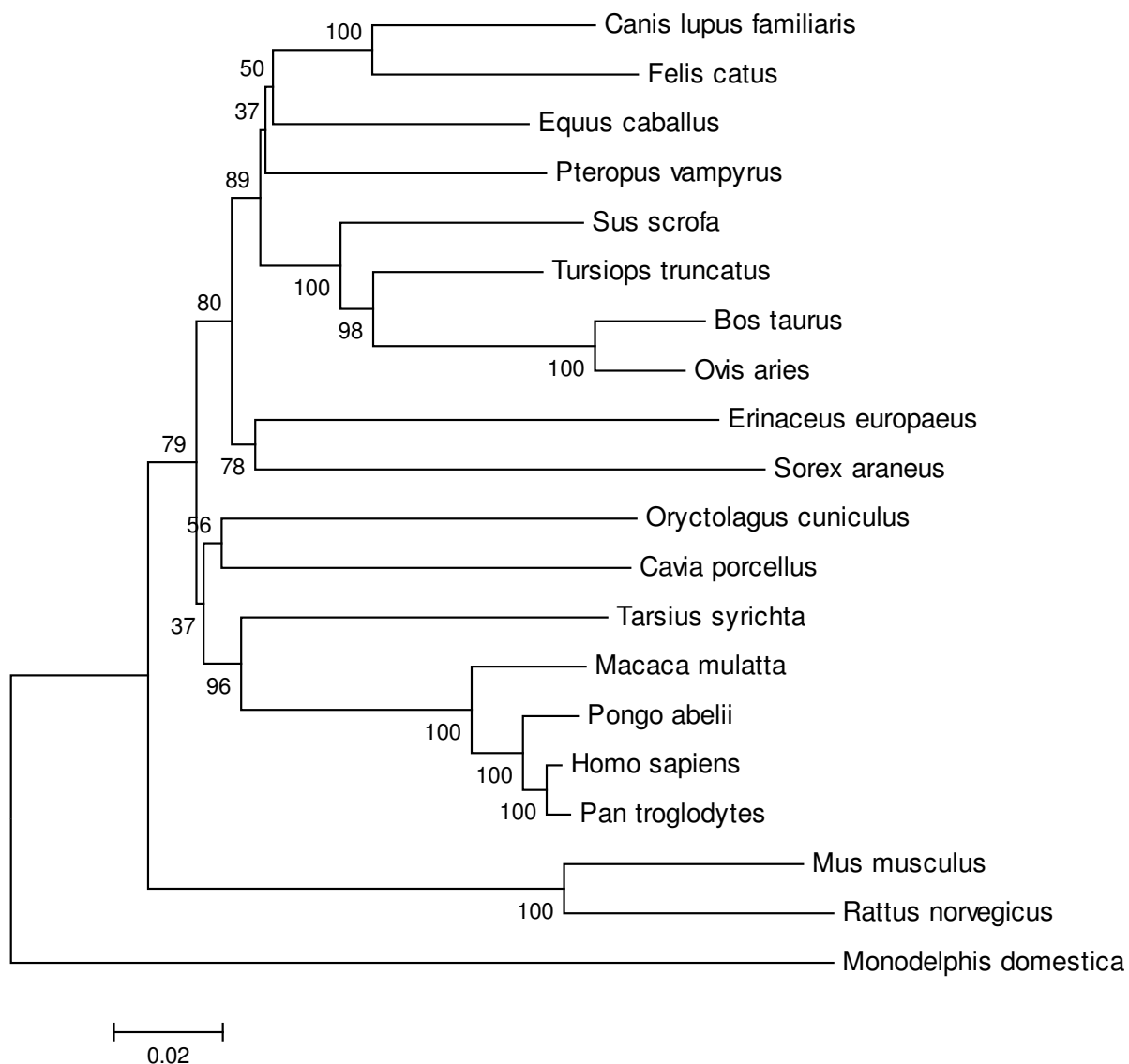


Figura 17. Árvore filogenética para o gene TLR3 com base na sequência de nucleotídeos do grupo de mamíferos analisado, utilizando o método *Neighbour Joining* e as opções *p-distance* e *complete deletion*. Os valores de *Bootstrap* obtidos com 1000 réplicas estão presentes nos nodos da árvore filogenética.

Discussão

O conhecimento do gene TLR3 nos processos de identificação de agentes patogénicos e em mecanismos fisiológicos, como a sinalização celular, tem crescido de uma forma exponencial nos últimos três anos. No entanto, questões acerca da sua estrutura e dos seus níveis de diversidade genética em populações selvagens permanecem por responder. Por isso, procurou-se integrar a informação obtida neste trabalho com a disponível acerca destas duas importantes temáticas.

1. Caracterização dos domínios do gene TLR3

A identificação dos domínios dos TLRs não está completamente esclarecida. Apesar de os domínios encontrados na sequência de aminoácidos do gene TLR3 de Coelho-bravo através do programa SMART serem semelhantes à informação disponível no GenBank, apresentando apenas pequenas oscilações nas posições de aminoácidos que delimitam os domínios, em ambas as fontes, não foi identificado o domínio LRR-NT, que apenas foi detectado com o programa LRRfinder. Fazendo uma comparação entre os programas utilizados para a detecção de domínios conservados, em particular de domínios LRRs, o LRRfinder pareceu mais efectivo na identificação destes últimos domínios (LRRs), no entanto não foi possível utilizar o programa para algumas espécies no que diz respeito à identificação de outros domínios (LRR-NT, LRR-CT, Transmembranar e TIR), devido à falta dessa informação na base de dados do programa. Ainda assim, para as espécies passíveis de serem analisadas em ambos os programas, com a respectiva identificação de todos os domínios, encontraram-se algumas diferenças. O LRRfinder identificou o domínio LRR-NT no Coelho-bravo e não identificou um segundo domínio Transmembranar que apenas ocorre na ratazana, tendo assim resultados contrários ao SMART. Relativamente ao domínio LRR-NT de Coelho-bravo, este parece estar presente na sequência pois foi efectuada a comparação com o domínio NT de Homem e a sequência encontra-se conservada (Anexo 6). Relativamente à existência de um segundo domínio Transmembranar na ratazana, serão necessários estudos para confirmar qual o método mais fiável, no entanto, parece improvável que exista este segundo domínio, dado ser o único mamífero analisado em que foi detectado. Estes resultados parecem sugerir que a delimitação dos diferentes domínios terá que ser revista e provavelmente sofrerá alterações com o desenvolvimento de novas metodologias e com a realização de estudos cristalográficos.

1.1. Caracterização do domínio LRR-NT

O domínio LRR-NT, de acordo com o programa LRRfinder, está localizado entre os aminoácidos 27 e 53. Tem sido demonstrada a sua importância na ligação ao dsRNA, pois um dos locais de ligação está aí localizado, LRR-NT-LRR3, formado pelos resíduos His39, His60, Arg64, Phe84, Ser86, His108 e Glu110 (Liu et al. 2008). No LRR-NT o resíduo conservado, His39, é essencial para a ligação (Liu et al. 2008) e encontra-se presente em todas as sequências das amostras analisadas. Outras mutações descritas neste domínio do TLR3 para o Homem, como por exemplo Gln62, Arg64 e His108 mantêm a função quando mutadas para Alanina, o que demonstra que não são essenciais à ligação (Pirher et al. 2008). Na sequência de TLR3 de Coelho-bravo o Gln62 está presente, assim como a His108, no entanto a Arg64 deu lugar a Lisina, desconhecendo-se no entanto as consequências na estrutura ou função da molécula. Nas amostras de populações de Coelho-bravo analisadas neste trabalho não foram detectadas mutações nos locais de ligação ao dsRNA (LRR-NT-LRR3), tendo sido apenas detectada uma alteração de aminoácido nesta região, I46V. Na comparação de sequências de TLR3 de mamíferos, em nenhum dos códons desta região foi encontrada evidência de selecção positiva.

1.2. Caracterização do domínio LRR

No domínio externo da molécula de TLR3, podem ainda ser encontrados um elevado número de domínios ricos em repetições de leucina, LRRs. O número de LRRs detectados para o Coelho-bravo através do programa SMART foi inferior (16 LRRs) ao detectado pelo programa LRRfinder (24 LRRs), contudo o número de LRRs foi semelhante quando foi utilizada a pesquisa de domínios Pfam em conjunto com o SMART (24 LRRs). No entanto, dentro de cada metodologia utilizada os valores foram semelhantes a outros mamíferos (ver Anexo 1 e 2).

Na análise de domínios LRRs nos mamíferos, foram observadas variações em relação ao número de LRRs de espécie para espécie, esta observação está de acordo com a previamente descrita por Chang et al. (2009). Como estes domínios são os principais pontos de interacção com o ligando, e como as diferentes espécies interagem com dsRNA de origens diversas, cada uma terá necessidade de se adaptar e evoluir de uma forma independente, o que poderá resultar em diferente número de LRRs.

Analisando todos os mamíferos, verificou-se que em cada espécie, o número de LRRs obtido foi diferente de acordo com os dois programas utilizados (SMART e

LRRfinder). Esta variação poderá estar associada às características próprias de cada *software*. Por exemplo, em relação ao gene TLR3 no Homem, foram descritos 23 LRRs por Bell et al. (2003) e foram detectados 17 com o programa SMART e 24 com o LRRfinder.

Das 15 alterações de aminoácidos detectadas nas populações de Coelho-bravo estudadas neste trabalho, 11 ocorrem em LRRs, sendo que 4 alterações são exclusivas das populações de *O. c. algirus* e outras 4 ocorrem exclusivamente nas populações de *O. c. cuniculus* analisadas. Este facto poderá estar associado a diferentes interações com agentes patogénicos ou com características do meio ambiente, que se reflectem em mutações características de cada subespécie. No entanto, estas alterações poderão estar associadas a mutações estocásticas que ocorreram após a separação destas duas subespécies, que evoluíram de uma forma independente desde a sua separação há cerca de 2 milhões de anos. Nenhuma das 11 alterações aminoacídicas provoca perda destes domínios, desconhecendo-se a possibilidade de originar mudanças de conformação da molécula ou mesmo da sua função. No entanto, a ocorrência de alterações conformacionais ou funcionais é pouco expectável, visto que as alterações de aminoácidos detectadas ocorrem na região variável do LRR, ou nos locais conservados da sequência que permitem a presença de qualquer aminoácido (X). Apenas num caso, L575F, a mutação provocou a substituição da leucina terminal da sequência consenso do LRR21, que pertence ao grupo de LRRs, LRR20-LRR22, que são muito importantes para a dimerização do receptor, bem como para a ligação e sinalização (Liu et al. 2008). No entanto, as leucinas podem ser substituídas por outros aminoácidos hidrofóbicos, tal como é o caso da fenilalanina, não perdendo as características dos LRRs (Bell et al. 2003, Matsushima et al. 2007). Foram ainda detectadas alterações de aminoácido no LRR20, K560E, que ocorre na região variável do domínio, o que provavelmente não terá consequências na molécula.

Estudos de caracterização genética do ectodomínio da molécula de TLR3 em mamíferos indicam que os resíduos H539 a N541, são essenciais para a ligação e sinalização (Bell et al. 2006). Estes encontram-se conservados em todas as amostras analisadas. Foram ainda descritas alterações de aminoácidos no domínio externo do TLR3 como consequência de polimorfismo de um único nucleótido (SNP), N284I, Y307D, L412F e P554S no Homem (Botos et al. 2009) e K144Q no Cavallo (Astakhova et al. 2009). A alteração N284I está localizada no core hidrofóbico do LRR e é responsável pela perda parcial ou total de actividade, o que presumivelmente desestabiliza a arquitectura do LRR (Botos et al. 2009). Por sua vez, a P554S está relacionada com a encefalite provocada por HSV, e portanto não deverá estar

implicada na ligação TLR3-dsRNA. Nenhuma destas alterações estão presentes nas amostras de Coelho-bravo analisadas.

Pequenas modificações na região de ligação do TLR3 ao dsRNA podem originar perda de função, pois é esse o ponto inicial da sinalização. As regiões entre Lys330 a Lys493 e Arg65 a Lys163 foram descritas como regiões com carga positiva importantes na ligação (Sang et al. 2008). Nas amostras analisadas, detectaram-se duas alterações de aminoácido em cada uma das regiões supracitadas. As alterações L101S, M132T e V480A ocorrem entre aminoácidos de carga neutra, no entanto a Q353H dá origem a um resíduo de carga positiva. As consequências de cada uma destas alterações em particular não são conhecidas no Coelho-bravo.

Relativamente à pesquisa de padrões de selecção natural, dos 24 codões sob selecção positiva, 16 estão localizados em LRRs, ou seja, o domínio extracelular está sob uma maior pressão selectiva, o que está de acordo com a literatura disponível para outros TLRs (Chang et al. 2009) e que poderá ser compreendido por uma maior necessidade de adaptação ao reconhecimento de diferentes agentes patogénios.

1.3. Caracterização do domínio LRR-CT

No domínio que sucede aos LRRs, o LRR-CT, não foi encontrada nenhuma alteração de aminoácido. Este facto é apoiado pela literatura que existe sobre este tema, pois até à data não foram descritas mutações neste domínio (Botos et al. 2009). Esta ausência de variação estará provavelmente associada à importância que este domínio apresenta na formação de dímeros assim como na estabilização e activação do complexo TLR3-dsRNA (Botos et al. 2009, Liu et al. 2008). Na comparação de sequências de TLR3 de mamíferos, foi encontrada evidência de selecção positiva em um dos codões desta região (663). Será interessante estudar mais aprofundadamente este resultado devido às características conservadas já enunciadas neste domínio.

1.4. Caracterização do domínio Transmembranar

Nas populações de Coelho-bravo, o domínio Transmembranar apresentou duas alterações de aminoácido, M708I e L716F, sendo que em ambos os casos estão envolvidos aminoácidos hidrofóbicos, o que não deverá provocar grandes alterações na molécula de TLR3. Nenhuma destas mutações está descrita na literatura, no entanto, para esta região foram já encontradas outras, como é o caso de T710A encontrada em raças de cavalos (Astakhova et al. 2009) da qual não se conhecem as possíveis consequências funcionais ou conformacionais. Os resultados obtidos na

pesquisa de selecção natural realizada neste trabalho, mostraram a ocorrência de selecção positiva num dos codões (712) deste domínio. De facto, para esta posição são observados 5 aminoácidos (ver Anexo 5).

1.5. Caracterização do domínio TIR

O domínio TIR é essencial para a sinalização intracelular e foi um importante sinal de reconhecimento desta classe de receptores, os receptores *Toll-like*. O primeiro receptor *Toll*, foi identificado na *Drosophila melanogaster* e possui um domínio TIR conservado, que é homólogo do receptor IL-1 e IL-18. A conservação genética da sequência do domínio TIR em várias espécies animais, plantas (Jebanathirajah et al. 2002) e microrganismos (Turner 2003), sublinha a importância deste domínio na sinalização. Estudos de comparação das sequências do domínio TIR do Macaco-rhesus com Humanos e Roedores revelaram um elevado nível de conservação da sequência quer na análise de nucleotídeos (entre 79,5% a 98,1%) quer na de aminoácidos (79,3% a 98,6%). Existem três regiões deste domínio (Região 1: FDAFISY; Região 2: GYKLC-RD-PG; Região 3: W rodeado de resíduos básicos), que são fundamentais na transdução de sinal e estão altamente conservadas entre Macaco-rhesus, Humanos e Murinos (Slack et al. 2000, Xu et al. 2000). As únicas diferenças encontradas nestas regiões entre a sequência de Macaco e de Homem é na região 3, R890H, no entanto esta alteração é conservativa e como tal não afectará significativamente a estrutura e função do domínio TIR do Macaco (Sanghavi et al. 2004). No presente estudo, foi efectuado o alinhamento das sequências do domínio TIR do gene TLR3 de Homem e Coelho-bravo de modo a identificar os resíduos presentes nas regiões conservadas e as alterações de aminoácidos. Constatou-se que existe alteração de um aminoácido em cada uma das regiões identificadas. Na região 1, registou-se a alteração A758T, na região 2, E795Q, e na região 3 a R890H. Esta última alteração tinha já sido descrita no Macaco (R890H) (Sanghavi et al. 2004).

Ainda relativamente à análise do domínio TIR nas amostras analisadas, foi detectada uma alteração de aminoácido, A864S. A alteração de um aminoácido hidrofóbico por um hidrofílico resulta, neste caso, num possível local de N-glicosilação. Adiante serão listadas algumas das possíveis consequências desta modificação. Esta mutação situa-se fora das 3 regiões conservadas, logo, à partida não se prevêem graves alterações na função do domínio, pois as alterações de aminoácidos não significativas no domínio TIR foram primeiramente detectadas fora das 3 regiões conservadas, contudo, estes resíduos podem também ser importantes na sinalização (Sanghavi et al. 2004), apesar de não estarem estabelecidas as

possíveis consequências. Existem estudos que descrevem outras alterações do domínio TIR com implicações conhecidas, tal como é o caso da eliminação da região 3 que conduz à perda da activação da IRAK e do NF- κ B (Slack et al. 2000). A não activação do NF- κ B também se pode dever à mutação A795H, que para além disso também não activa o promotor do IFN- β nas células HEK293 sob estimulação de Poly(I:C), por falha na ligação à TICAM-1 (Matsumoto and Seya 2008). Alguns resíduos estão descritos como essenciais para a ligação e indução de NF- κ B e IFN- β , como por exemplo o Phe732, Leu742 e Gly743. Existem ainda algumas Tirosinas no domínio TIR cuja fosforilação é essencial à sinalização, logo, a alteração dos resíduos Tyr759, Tyr858, Tyr733 e Tyr756 é comprometedora neste processo, sendo que os dois primeiros são essenciais. Funami et al. (2004), Sarkar et al. (2003) descreveram ainda dois resíduos do domínio TIR do TLR3, Arg740 e Val741, localizados na ligação à região citoplasmática, que são determinantes para a localização intracelular e expressão do TLR3 nas células dendríticas no Homem. Todas as alterações anteriormente citadas no domínio TIR, encontram-se conservadas nas amostras de Coelho-bravo analisadas, excepto o resíduo de localização intracelular, Val741, que se encontra mutado em todas as amostras de Coelho-bravo por uma Isoleucina. Esta observação, levanta a hipótese da localização intracelular do TLR3 de Coelho-bravo seja diferente da do Homem.

Os dois codões no domínio TIR que apresentaram evidências de selecção natural, 749 e 780 (ver Tabela 9 e Anexo 5), não se encontram dentro das regiões conservadas. Este resultado, é expectável uma vez que as regiões conservadas deverão estar sob uma forte selecção negativa. No entanto, não está descrita nenhuma função destes aminoácidos que possa explicar estas evidências de selecção positiva.

2. Diversidade genética das populações de Coelho-bravo

Por comparação da diversidade genética encontrada neste trabalho, 41 posições nucleotídicas polimórficas, das quais 15 originaram alterações de aminoácidos, com estudos efectuados noutros mamíferos verificou-se que esta foi consideravelmente superior. Como exemplo, um estudo efectuado sobre o gene TLR3 num grupo de 9 raças de bovinos encontrou 7 SNPs, em que 3 substituições foram não sinónimas (Cargill and Womack 2007). Em 84 porcos pertencentes a 11 raças foi detectado um baixo número de SNPs (Morozumi and Uenishi 2009). Em 16 cavalos pertencentes a 5 raças foram detectados 4 SNPs em que duas substituições são não

sinonímias (Astakhova et al. 2009). Em 37 raças de gado foram encontrados apenas 3 polimorfismos não sinonímios no gene TLR3. Em estudos efectuados na totalidade dos TLRs, em chimpanzés, o nível de polimorfismo varia entre 0,01 e 0,06% (Wlasiuk and Nachman 2010). Nos humanos, a combinação de regiões codificantes e não codificantes varia entre 0,03% e 0,23% na população africana e entre 0,03% e 0,12% na população europeia (Wlasiuk and Nachman 2010). A maior diversidade obtida neste trabalho poderá estar associada à origem das amostras testadas, pois foram utilizadas populações selvagens, que não apresentavam selecção mediada pelo homem e que estiveram em contacto com diversos agentes patogénicos.

3. Caracterização dos *C-repeats* do gene TLR3 em Coelho-bravo

Em paralelo com a pesquisa de domínios conservados na região codificante do gene TLR3, foram também pesquisadas sequências repetitivas nas regiões intrónicas, caracteristicamente inseridos no genoma dos Lagomorfos, *C-repeats*. Foram detectados 5 *C-repeats* dos quais três estão inseridos no sentido 3'→5'. As sequências de *C-repeats* CR1, CR2 e CR5 foram classificadas como pertencentes ao grupo III, pois agruparam com os *C-repeats* c3 e c8 já anteriormente classificados neste grupo (Krane et al. 1991). Este resultado indica que estes *C-repeats* se terão inserido neste gene entre 8 e 42 milhões de anos e que portanto estarão provavelmente também presentes na família *Ochotonidae* que divergiu da *Leporidae* há cerca de 35 a 40 milhões de anos (Matthee 2009). Os dados disponíveis não permitem tirar conclusões acerca da possível presença destes *C-repeats* nos restantes géneros da família *Leporidae* que divergiram há cerca de 12 milhões de anos (Matthee 2009). Por outro lado, as sequências de CR3 e CR4, identificadas nas sequências de TLR3 de Coelho-bravo, agruparam-se com *C-repeats* do grupo mais antigo (c18)

, grupo I que se terão inserido entre 43 e 131 Milhões de anos, estando provavelmente generalizadamente presentes em toda a ordem *Lagomorpha*.

Na árvore filogenética conseguida com o alinhamento de todas as sequências de *C-repeats* em estudo, observou-se uma maior divergência entre as sequências que pertencem ao grupo I e II. Este dado está de acordo com a literatura que refere que os *C-repeats* pertencentes ao grupo II apresentam mais diferenças do que as do grupo III, que são muito similares entre si. Por sua vez, as sequências do grupo I apresentam tantas diferenças entre si como com os outros grupos, justificando assim os *bootstraps* mais elevados nas ramificações das sequências classificadas no grupo III do que nos restantes dois.

4. Glicosilação

A maioria das proteínas não executa a sua função sem antes ser submetida a modificações pós-traducionais (PTM) que ocorrem após a tradução do mRNA para a sequência peptídica. A glicosilação é uma PTM muito comum que tem implicações na conformação proteica, transporte e vida-média da proteína. Está ainda envolvida em interacções entre células e em processos de antigenicidade (Hart 1992, Seitz 2000, Spiro 2002). É uma modificação que ocorre por processos enzimáticos (com excepção da glicação), envolvendo a adição de açúcares à proteína de modo a construir uma cadeia de glicano. Existem 4 tipos de glicosilação: ligada ao N, ligada ao O, C-manosilação e Glicofosfatidil-inositol (GPI). A N-glicosilação é dos tipos mais comuns, em conjunto com a O-glicosilação, e consiste na adição de uma cadeia de glicano a um resíduo Asn. A cadeia de glicano que é adicionada consiste em Xxx-Ser/Thr, sendo que o motivo será Asn-Xxx-Ser/Thr (Blom et al. 2004, Varki et al. 1999), ou em alguns casos raros Asn-Xxx-Cys, onde o X corresponde a qualquer aminoácido excepto Prolina (Blom et al. 2004).

Esta PTM é um mecanismo capaz de gerar diversidade de funções, incluindo por vezes, ausência de função ou pouca relevância no contexto da molécula. Na maior parte das vezes relevam-se modificações cruciais para o desenvolvimento, função e sobrevivência do organismo, podendo também assumir um papel modulador da interacção proteica, activar ou alterar funções. As alterações dos locais de glicosilação podem ter consequências variáveis e imprevisíveis, que podem ir deste indetectáveis até completas perdas de função ou mesmo perda da glicoproteína (Varki et al. 1999)

Nas amostras analisadas foram encontrados 17 locais de glicosilação na molécula de TLR3 de Coelho-bravo dos quais 15 locais se localizam no domínio externo do TLR3, tal como no Homem (Botos et al. 2009), Búfalo e Nilgó (Dhara et al. 2007). Segundo Botos et al. (2009) a alteração da maioria dos locais de glicosilação não produz alteração significativa na actividade ou expressão do TLR3, no entanto a alteração de Asn247 para Alanina ou Ácido Aspártico provoca perda total ou parcial de actividade sem no entanto perder a expressão. Por outro lado a mutação do resíduo Asn413, que é o único carboidrato que mantém contacto com o dsRNA, tem sido descrito com efeitos contraditórios que podem indicar perda parcial de resposta (Botos et al. 2009). Estes dois resíduos estão conservados na sequência de TLR3 de Coelho-bravo, nas posições Asn244 e Asn410 respectivamente.

Na análise conjunta com outros mamíferos, em comparação com o Homem, o Coelho-bravo perde um local de glicosilação entre as posições aminoacídicas 288-

291, assim como a ratazana, no entanto ganhou um entre as posições 312-315, que aliás, é exclusivo neste grupo de mamíferos analisados. Esta alteração ocorre num domínio LRR, logo, poderá estar relacionada com a interacção entre o TLR3 e o ligando. O Coelho-bravo perdeu ainda um local de N-glicosilação entre as posições 504-507, tal como aconteceu em outros 6 mamíferos, e ganhou um local de glicosilação compreendido entre os aminoácidos 664-667, que é comum à maioria dos organismos analisados. Esta perda de um sítio de glicosilação deu-se num domínio LRR (LRR18), e o ganho do local de N-glicosilação localiza-se no domínio LRR-CT. Será interessante esclarecer se esta alteração do local de N-glicosilação estará relacionada com a adaptação a reconhecimento de diferentes ligandos, e/ou sinalização intracelular. Por comparação dos possíveis locais de N-glicosilação com os codões sob selecção positiva, verificou-se que nenhum deles diz respeito a posições N ou S/T dos motivos de N-glicosilação no Coelho-bravo.

Considerações Finais

Os resultados descritos neste trabalho revelaram elevados níveis de polimorfismo para o gene TLR3 em populações de Coelho-bravo, quando comparados com os obtidos para outras espécies. Estes resultados já eram espectáveis, pois foram analisadas populações selvagens que não terão tido ao longo da sua evolução restrições populacionais muito severas e não terão sofrido selecção natural mediada pelo homem. Em outros trabalhos que estudaram este gene, apenas se obtiveram resultados usando raças domésticas fortemente seleccionadas. Para poder comparar os níveis de polimorfismo obtidos no Coelho-bravo com os obtidos em outras espécies, será necessário estudar raças domésticas de Coelho-bravo fortemente seleccionadas para manter determinadas características, ou populações insulares que se formaram a partir de um reduzido número de indivíduos. Infelizmente, para a maioria das espécies que sofreram processos de domesticação as populações selvagens já desapareceram ou encontram-se muito ameaçadas o que torna impossível a sua comparação com os dados obtidos neste trabalho.

A análise de padrões de selecção natural em sequências do gene TLR3 em mamíferos revelou a existência de vários codões sob forte selecção positiva, o que pode ser explicado pelo facto de o gene TLR3 estar associado a respostas virais. Por isso, nos sítios da proteína que estão mais expostos, é natural que apresentem respostas associadas a vírus específicos de cada espécie, originando uma elevada mutabilidade e conseqüentemente uma elevada taxa de variabilidade aminoacídica. No entanto, a ocorrência de selecção natural para este gene não está presente em todos os grupos animais. Por exemplo, nos primatas não foi detectada selecção positiva (Wlasiuk and Nachman 2010).

Até à data não é conhecido nenhum estudo especificamente conduzido para verificar se no Coelho-bravo o ligando do TLR3 é o ácido nucleico de origem viral, dsRNA, tal como se verifica em outros mamíferos. Contudo, foi assumido este dado com base em estudos efectuados em diferentes organismos (Matsuo et al. 2008). Por exemplo, a expressão de TLR3 tanto em células humanas como de peixe, apresenta resposta ao poly(I:C) e a dsRNA (Matsuo et al. 2008). No entanto, este reconhecimento nem sempre é linear pois também foram observadas diferenças específicas no reconhecimento de alguns ligandos. É o caso de ssRNA, DNA bacteriano e flagelina, descritos por exemplo entre o Homem e o rato (Oshiumi et al. 2008). Estas diferenças reflectem, presumivelmente, diferentes pressões selectivas a actuar em cada hospedeiro em virtude da adaptação a novos ambientes e a agentes patogénios (Chang et al. 2009). Para uma melhor compreensão do papel fisiológico do TLR3 no Coelho-bravo, será importante a realização de estudos de expressão tecidual.

As populações de Coelho-bravo têm sofrido decréscimos significativos provocados por uma incorrecta gestão cinegética e pelo aparecimento de duas doenças virais altamente letais. O processo infeccioso dessas doenças, Mixomatose e Doença Hemorrágica Viral, permanece por esclarecer. Está demonstrada a ligação de vários TLRs a respostas virais, nomeadamente os TLR7, 8 e 9. A análise do genoma de Coelho-bravo, mostrou a pseudogenização do TLR8 e ausência do TLR7 (Astakhova et al. 2009). A ocorrência de uma outra doença viral com características semelhantes à RHD, Doença da Lebre Parda, que é responsável por elevadas mortalidades em lebre e a ocorrência de uma forma benigna de mixomatose no género *Sylvilagus*, Coelho-americano, sugere que o estudo destes genes em outros Lagomorfos poderá ser útil na compreensão dos processos infecciosos no Coelho-bravo. Este estudo não só seria importante para esta espécie mas também serviria como modelo animal para compreender diferentes enfermidades virais que atacam outros mamíferos, nomeadamente o Homem.

A comparação da diversidade genética em populações naturais é uma excelente ferramenta para detectar vantagens selectivas de determinados alelos. Por isso, o estudo de populações com diferentes níveis de mortalidade será de grande importância e a metodologia seguida neste trabalho deverá ser replicada para os outros TLRs. Este tipo de abordagem poderá ser útil para responder a questões fundamentais de biologia da conservação: Será que a detecção de selecção positiva está associada a um aumento de resistência a doenças? Existem variações na distribuição geográfica de um variante genético que possam explicar a ocorrência de diferenças de susceptibilidade entre populações? Podem ser estes dados usados como instrumentos de melhor controle das populações naturais de Coelho-bravo?

Referências Bibliográficas

- Abrantes J, Posada D, Guillon P, Esteves PJ, Le Pendu J. Widespread gene conversion of alpha-2-fucosyltransferase genes in mammals. *J Mol Evol.* 2009. 69: 22-31.
- Akashi-Takamura S, Miyake K. TLR accessory molecules. *Curr Opin Immunol.* 2008. 20: 420-425.
- Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. *Cell.* 2006. 124: 783-801.
- Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by *Toll-like* receptor 3. *Nature.* 2001. 413: 732-738.
- Alves P, Ferreira C. 2004. Revisão do Livro Vermelho dos Vertebrados de Portugal. Determinação da abundância relativa das populações de coelho-bravo (*Oryctolagus cuniculus*) em Portugal Continental.: CIBIO, ICETA, Universidade do Porto. Report no.
- Androulidaki A, Iliopoulos D, Arranz A, Doxaki C, Schworer S, Zacharioudaki V, Margioris AN, Tsihchlis PN, Tsatsanis C. The kinase Akt1 controls macrophage response to lipopolysaccharide by regulating microRNAs. *Immunity.* 2009. 31: 220-231.
- Anon. Doença hemorrágica a vírus do Coelho em Portugal. *Revista Portuguesa de Ciências Veterinárias.* 1989. 57-58.
- Asher RJ, Meng J, Wible JR, McKenna MC, Rougier GW, Dashzeveg D, Novacek MJ. Stem Lagomorpha and the antiquity of Glires. *Science.* 2005. 307: 1091-1094.
- Astakhova NM, Perelygin AA, Zharkikh AA, Lear TL, Coleman SJ, MacLeod JN, Brinton MA. Characterization of equine and other vertebrate TLR3, TLR7, and TLR8 genes. *Immunogenetics.* 2009. 61: 529-539.
- Balkwill F, Coussens LM. Cancer: an inflammatory link. *Nature.* 2004. 431: 405-406.
- Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD. MicroRNAs: new regulators of immune cell development and function. *Nat Immunol.* 2008. 9: 839-845.
- Barton GM, Kagan JC. A cell biological view of *Toll-like* receptor function: regulation through compartmentalization. *Nat Rev Immunol.* 2009. 9: 535-542.
- Bazzoni F, Rossato M, Fabbri M, Gaudiosi D, Mirolo M, Mori L, Tamassia N, Mantovani A, Cassatella MA, Locati M. Induction and regulatory function of miR-9 in human monocytes and neutrophils exposed to proinflammatory signals. *Proc Natl Acad Sci U S A.* 2009. 106: 5282-5287.
- Bell JK, Askins J, Hall PR, Davies DR, Segal DM. The dsRNA binding site of human *Toll-like* receptor 3. *Proc Natl Acad Sci U S A.* 2006. 103: 8792-8797.
- Bell JK, Mullen GE, Leifer CA, Mazzoni A, Davies DR, Segal DM. Leucine-rich repeats and pathogen recognition in *Toll-like* receptors. *Trends Immunol.* 2003. 24: 528-533.

- Bell JK, Botos I, Hall PR, Askins J, Shiloach J, Segal DM, Davies DR. The molecular structure of the *Toll-like* receptor 3 ligand-binding domain. *Proc Natl Acad Sci U S A*. 2005. 102: 10976-10980.
- Belvin MP, Anderson KV. A conserved signaling pathway: the *Drosophila* toll-dorsal pathway. *Annu Rev Cell Dev Biol*. 1996. 12: 393-416.
- Beutler B, Jiang Z, Georgel P, Crozat K, Croker B, Rutschmann S, Du X, Hoebe K. Genetic analysis of host resistance: *Toll-like* receptor signaling and immunity at large. *Annu Rev Immunol*. 2006. 24: 353-389.
- Beutler BA. TLRs and innate immunity. *Blood*. 2009. 113: 1399-1407.
- Biju-Duval C, Ennafaa H, Dennebouy N, al e. Mitochondrial DNA evolution in Lagomorphs: origin of systematic heteroplasmy, organization of diversity in European rabbits. *J Mol Evol*. 1991. 92-102.
- Biragyn A, et al. *Toll-like* receptor 4-dependent activation of dendritic cells by beta-defensin 2. *Science*. 2002. 298: 1025-1029.
- Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics*. 2004. 4: 1633-1649.
- Boraschi D, Tagliabue A. The interleukin-1 receptor family. *Vitam Horm*. 2006. 74: 229-254.
- Botos I, Liu L, Wang Y, Segal DM, Davies DR. The *Toll-like* receptor 3:dsRNA signaling complex. *Biochim Biophys Acta*. 2009. 1789: 667-674.
- Branco M, Ferrand N, Monnerot M. Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome b gene. *Heredity*. 2000. 85 Pt 4: 307-317.
- Branco M, Monnerot M, Ferrand N, Templeton AR. Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution*. 2002. 56: 792-803.
- Cameron C, Hota-Mitchell S, Chen L, Barrett J, Cao JX, Macaulay C, Willer D, Evans D, McFadden G. The complete DNA sequence of myxoma virus. *Virology*. 1999. 264: 298-318.
- Cameron JS, et al. *Toll-like* receptor 3 is a potent negative regulator of axonal growth in mammals. *J Neurosci*. 2007. 27: 13033-13041.
- Campos R, Storz JF, Ferrand N. Evidence for contrasting modes of selection at interacting globin genes in the European rabbit (*Oryctolagus cuniculus*). *Heredity*. 2008. 100: 602-609.
- Cargill EJ, Womack JE. Detection of polymorphisms in bovine *Toll-like* receptors 3, 7, 8, and 9. *Genomics*. 2007. 89: 745-755.
- Carpenter S, O'Neill LA. How important are *Toll-like* receptors for antimicrobial responses? *Cell Microbiol*. 2007. 9: 1891-1901.
- Carty M, Bowie AG. Recent insights into the role of *Toll-like* receptors in viral infection. *Clin Exp Immunol*. 2010.

Chang JS, Russell GC, Jann O, Glass EJ, Werling D, Haig DM. Molecular cloning and characterization of *Toll-like* receptors 1-10 in sheep. *Vet Immunol Immunopathol.* 2009. 127: 94-105.

Chasey D. Rabbit haemorrhagic disease: the new scourge of *Oryctolagus cuniculus*. *Lab Anim.* 1997. 31: 33-44.

Chaturvedi APS. How location governs *Toll-like* receptor signaling. *Traffic.* 2009. 10: 621-628.

Choe J, Kelker MS, Wilson IA. Crystal structure of human *Toll-like* receptor 3 (TLR3) ectodomain. *Science.* 2005. 309: 581-585.

Chomczynski P, Sacchi N. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc.* 2006. 1: 581-585.

Conley ME. Immunodeficiency: UNC-93B gets a toll call. *Trends Immunol.* 2007. 28: 99-101.

Cooke BD. Rabbit haemorrhagic disease: field epidemiology and the management of wild rabbit populations. *Rev Sci Tech.* 2002. 21: 347-358.

Cooke BD, Fenner F. Rabbit haemorrhagic disease and the biological control of wild rabbits, *Oryctolagus cuniculus*, in Australia and New Zealand. *Wildl. Res.* 2002. 689-706.

Davidson EH, Britten RJ. Organization, transcription, and regulation in the animal genome. *Q Rev Biol.* 1973. 48: 565-613.

de Bouteiller O, Merck E, Hasan UA, Hubac S, Benguigui B, Trinchieri G, Bates EE, Caux C. Recognition of double-stranded RNA by human *Toll-like* receptor 3 and downstream receptor signaling requires multimerization and an acidic pH. *J Biol Chem.* 2005. 280: 38133-38145.

Delport W, Poon AF, Frost SD, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics.* 2010. 26: 2455-2457.

Dhara A, Saini M, Das DK, Swarup D, Sharma B, Kumar S, Gupta PK. Molecular characterization of coding sequences and analysis of *Toll-like* receptor 3 mRNA expression in water buffalo (*Bubalus bubalis*) and nilgai (*Boselaphus tragocamelus*). *Immunogenetics.* 2007. 59: 69-76.

DiGiacomo RF, Mare CJ. Viral diseases. Pages 171-183 in Manning P, DH R, Newcomer C, eds. *The biology of the laboratory rabbit.* San Diego: Academic Press. 1994.

Edelmann KH, Richardson-Burns S, Alexopoulou L, Tyler KL, Flavell RA, Oldstone MB. Does *Toll-like* receptor 3 play a biological role in virus infections? *Virology.* 2004. 322: 231-238.

Esteves, Lanning D, Ferrand N, Knight KL, Zhai SK, van der Loo W. Allelic variation at the VHa locus in natural populations of rabbit (*Oryctolagus cuniculus*, L.). *J Immunol.* 2004. 172: 1044-1053.

Esteves P, Alves P, Ferrand N. O uso de marcadores genéticos na gestão e conservação de populações de coelho-bravo (*Oryctolagus cuniculus*). Pages 18

in Caçadores FAd, ed. Gestão de Populações de Coelho Bravo, C Ferreira, P C Alves.2006.

Fenner F. The Florey lecture, 1983. Biological control, as exemplified by smallpox eradication and myxomatosis. Proc R Soc Lond B Biol Sci. 1983. 218: 259-285.

Fenner F. Adventures with poxviruses of vertebrates. FEMS Microbiol Lett. 2000. 123-133.

Fenner F, Ratcliffe F. Myxomatosis. Cambridge: University Press.1965.

Fenner F, Ross J. Myxomatosis. Pages 205-239 in Thompson HV, King CM, eds. The European Rabbit. The History and Biology of a Successful Colonizer. Oxford: Oxford University Press.1994.

Fields BN, Knipe DM, Howley PM, EGriffin D. Fields' virology
Philadelphia: Lippincott Williams & Wilkins.2001.

Finn RD, et al. The Pfam protein families database. Nucleic Acids Res. 2010. 38: D211-222.

Flux JEC FP. World distribution of the rabbit (*Oryctolagus cuniculus*). Acta Zoológica Fennica. 1983. 75-77.

Flux JEC FP. World distribution of the rabbit (*Oryctolagus cuniculus*) on islands. Mammal Review. 1992. 151-202.

Forrester NL, Trout RC, Turner SL, Kelly D, Boag B, Moss S, Gould EA. Unravelling the paradox of rabbit haemorrhagic disease virus emergence, using phylogenetic analysis; possible implications for rabbit conservation strategies. Biol. Conserv. 2006. 296-306.

Funami K, Matsumoto M, Oshiumi H, Akazawa T, Yamamoto A, Seya T. The cytoplasmic 'linker region' in *Toll-like* receptor 3 controls receptor localization and signaling. Int Immunol. 2004. 16: 1143-1154.

Gay NJ, Keith FJ. Drosophila Toll and IL-1 receptor. Nature. 1991. 351: 355-356.

Geraldes A, Rogel-Gaillard C, Ferrand N. High levels of nucleotide diversity in the European rabbit (*Oryctolagus cuniculus*) SRY gene. Anim Genet. 2005. 36: 349-351.

Geraldes A, Ferrand N, Nachman MW. Contrasting patterns of introgression at X-linked loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus cuniculus*). Genetics. 2006. 173: 919-933.

Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol. 1994. 11: 725-736.

Gorski J, Mizak B, Chrobocinska M. Control of viral haemorrhagic disease of rabbits in Poland. Rev Sci Tech. 1994. 13: 881-891.

Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. 1999. 95-98.

Hart GW. Glycosylation. Curr Opin Cell Biol. 1992. 4: 1017-1023.

- Hashimoto C, Hudson KL, Anderson KV. The Toll gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell*. 1988. 52: 269-279.
- He W, Liu Q, Wang L, Chen W, Li N, Cao X. TLR4 signaling promotes immune escape of human lung cancer cells by inducing immunosuppressive cytokines and apoptosis resistance. *Mol Immunol*. 2007. 44: 2850-2859.
- Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C, Akira S, Lipford G, Wagner H, Bauer S. Species-specific recognition of single-stranded RNA via *Toll-like* receptor 7 and 8. *Science*. 2004. 303: 1526-1529.
- Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RA. Phylogenetic perspectives in innate immunity. *Science*. 1999. 284: 1313-1318.
- Hold GL, El-Omar EM. Genetic aspects of inflammation and cancer. *Biochem J*. 2008. 410: 225-235.
- Huang B, Zhao J, Unkeless JC, Feng ZH, Xiong H. TLR signaling by tumor and immune cells: a double-edged sword. *Oncogene*. 2008. 27: 218-224.
- Huang B, Zhao J, Li H, He KL, Chen Y, Chen SH, Mayer L, Unkeless JC, Xiong H. *Toll-like* receptors on tumor cells facilitate evasion of immune surveillance. *Cancer Res*. 2005. 65: 5009-5014.
- Janeway CA, Jr., Medzhitov R. Innate immune recognition. *Annu Rev Immunol*. 2002. 20: 197-216.
- Janssens S, Beyaert R. Role of *Toll-like* receptors in pathogen recognition. *Clin Microbiol Rev*. 2003. 16: 637-646.
- Jarrossay D, Napolitani G, Colonna M, Sallusto F, Lanzavecchia A. Specialization and complementarity in microbial molecule recognition by human myeloid and plasmacytoid dendritic cells. *Eur J Immunol*. 2001. 31: 3388-3393.
- Jebanathirajah JA, Peri S, Pandey A. Toll and interleukin-1 receptor (TIR) domain-containing proteins in plants: a genomic perspective. *Trends Plant Sci*. 2002. 7: 388-391.
- Jiang Z, et al. Details of *Toll-like* receptor:adapter interaction revealed by germline mutagenesis. *Proc Natl Acad Sci U S A*. 2006. 103: 10961-10966.
- Kadowaki N, Ho S, Antonenko S, Malefyt RW, Kastelein RA, Bazan F, Liu YJ. Subsets of human dendritic cell precursors express different *Toll-like* receptors and respond to different microbial antigens. *J Exp Med*. 2001. 194: 863-869.
- Kawai T, Akira S. The roles of TLRs, RLRs and NLRs in pathogen recognition. *Int Immunol*. 2009. 21: 317-337.
- Kerr PJ, Kitchen A, Holmes EC. Origin and phylodynamics of rabbit hemorrhagic disease virus. *J Virol*. 2009. 83: 12129-12138.
- Kim JI, Lee CJ, Jin MS, Lee CH, Paik SG, Lee H, Lee JO. Crystal structure of CD14 and its implications for lipopolysaccharide signaling. *J Biol Chem*. 2005. 280: 11347-11351.
- Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature*. 1977. 267: 275-276.

- Kohany O, Gentles AJ, Hankus L, Jurka J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics*. 2006. 7: 474.
- Krane DE, Clark AG, Cheng JF, Hardison RC. Subfamily relationships and clustering of rabbit C repeats. *Mol Biol Evol*. 1991. 8: 1-30.
- Kumagai Y, Kumar H, Koyama S, Kawai T, Takeuchi O, Akira S. Cutting Edge: TLR-Dependent viral recognition along with type I IFN positive feedback signaling masks the requirement of viral replication for IFN- α production in plasmacytoid dendritic cells. *J Immunol*. 2009. 182: 3960-3964.
- Kumar H, Kawai T, Akira S. *Toll-like* receptors and innate immunity. *Biochem Biophys Res Commun*. 2009. 388: 621-625.
- Kurt-Jones EA, et al. Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. *Nat Immunol*. 2000. 1: 398-401.
- Lawrence CB, McDonnell DP, Ramsey WJ. Analysis of repetitive sequence elements containing tRNA-like sequences. *Nucleic Acids Res*. 1985. 13: 4239-4252.
- Le Gall G, Arnauld C, Boilletot E, Morisse JP, Rasschaert D. Molecular epidemiology of rabbit haemorrhagic disease virus outbreaks in France during 1988 to 1995. *J Gen Virol*. 1998. 79 (Pt 1): 11-16.
- Lee HK, Dunzendorfer S, Soldau K, Tobias PS. Double-stranded RNA-mediated TLR3 activation is enhanced by CD14. *Immunity*. 2006. 24: 153-163.
- Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell*. 1996. 86: 973-983.
- Leonard JN, Ghirlando R, Askins J, Bell JK, Margulies DH, Davies DR, Segal DM. The TLR3 signaling complex forms by cooperative receptor dimerization. *Proc Natl Acad Sci U S A*. 2008. 105: 258-263.
- Leonard JP, et al. Phase I trial of *Toll-like* receptor 9 agonist PF-3512676 with and following rituximab in patients with recurrent indolent and aggressive non Hodgkin's lymphoma. *Clin Cancer Res*. 2007. 13: 6168-6174.
- Letunic I, Doerks T, Bork P. SMART 6: recent updates and new developments. *Nucleic Acids Res*. 2009. 37: D229-232.
- Liu L, Botos I, Wang Y, Leonard JN, Shiloach J, Segal DM, Davies DR. Structural basis of *Toll-like* receptor 3 signaling with double-stranded RNA. *Science*. 2008. 320: 379-381.
- Liu S, Xue HP, Pu BQ, Qian N. A new viral disease in rabbits. *Animal Husbandry and Veterinary Medicine*. 1984. 253±255.
- Liu YJ, Kanzler H, Soumelis V, Gilliet M. Dendritic cell lineage, plasticity and cross-regulation. *Nat Immunol*. 2001. 2: 585-589.
- Lopez-Martinez N. *Revision sistemática y biostratigráfica de los lagomorphos (Mammalia) del neogeno y cuaternario de España*. 1989.
- Lu LF, Liston A. MicroRNA in the immune system, microRNA as an immune system. *Immunology*. 2009. 127: 291-298.

- Marcato PS, Benazzi C, Vecchi G, Galeotti M, Della Salda L, Sarli G, Lucidi P. Clinical and pathological features of viral haemorrhagic disease of rabbits and the European brown hare syndrome. *Rev Sci Tech*. 1991. 10: 371-392.
- Marlier D, Mainil J, Linde A, Vindevogel H. Infectious agents associated with rabbit pneumonia: isolation of amyxomatous myxoma virus strains. *Vet J*. 2000a. 159: 171-178.
- Marlier D, Mainil J, Sulon J, Beckers JF, Linden A, Vindevogel H. Study of the virulence of five strains of amyxomatous myxoma virus in crossbred New Zealand White/Californian conventional rabbits, with evidence of long-term testicular infection in recovered animals. *J Comp Pathol*. 2000b. 122: 101-113.
- Matsuguchi T, Musikachoen T, Ogawa T, Yoshikai Y. Gene expressions of *Toll-like* receptor 2, but not *Toll-like* receptor 4, is induced by LPS and inflammatory cytokines in mouse macrophages. *J Immunol*. 2000. 165: 5767-5772.
- Matsumoto M, Seya T. TLR3: interferon induction by double-stranded RNA including poly(I:C). *Adv Drug Deliv Rev*. 2008. 60: 805-812.
- Matsumoto M, Kikkawa S, Kohase M, Miyake K, Seya T. Establishment of a monoclonal antibody against human *Toll-like* receptor 3 that blocks double-stranded RNA-mediated signaling. *Biochem Biophys Res Commun*. 2002. 293: 1364-1369.
- Matsuo A, Oshiumi H, Tsujita T, Mitani H, Kasai H, Yoshimizu M, Matsumoto M, Seya T. Teleost TLR22 recognizes RNA duplex to induce IFN and protect cells from birnaviruses. *J Immunol*. 2008. 181: 3474-3485.
- Matsushima N, Tanaka T, Enkhbayar P, Mikami T, Taga M, Yamada K, Kuroki Y. Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate *Toll-like* receptors. *BMC Genomics*. 2007. 8: 124.
- Matthee. Pikas, Hares and Rabbits (Lagomorpha). Pages 487-489 in Press OU, ed. *The Timetree of Life*, Hedges Kumar. 2009.
- Matzinger P. The danger model: a renewed sense of self. *Science*. 2002. 296: 301-305.
- McGettrick AF, O'Neill LA. The expanding family of MyD88-like adaptors in *Toll-like* receptor signal transduction. *Mol Immunol*. 2004. 41: 577-582.
- Medzhitov R. *Toll-like* receptors and innate immunity. *Nat Rev Immunol*. 2001. 1: 135-145.
- Medzhitov R. Recognition of microorganisms and activation of the immune response. *Nature*. 2007. 449: 819-826.
- Medzhitov R, Janeway CA, Jr. Innate immunity: the virtues of a nonclonal system of recognition. *Cell*. 1997. 91: 295-298.
- Medzhitov R, Preston-Hurlburt P, Janeway CA, Jr. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature*. 1997. 388: 394-397.

- Mellman I, Steinman RM. Dendritic cells: specialized and regulated antigen processing machines. *Cell*. 2001. 106: 255-258.
- Miettinen M, Sareneva T, Julkunen I, Matikainen S. IFNs activate *Toll-like* receptor gene expression in viral infections. *Genes Immun*. 2001. 2: 349-355.
- Miyata T, Yasunaga T. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *J Mol Evol*. 1980. 16: 23-36.
- Monnerot M, Vigne JD, Biju-Duval C, Casane D, Callou C, Hardy C, Mougel F, Soriguer R, Dennebouy N, Mounolou JC. Rabbit and man: genetic and historic approach. *Genet Sel Evol*. 1994. 26: 167-182.
- Moreno S, et al. Long-term decline of the European wild rabbit (*Oryctolagus cuniculus*) in south-western Spain. *Wildl. Res*. 2007. 652-658.
- Morozumi T, Uenishi H. Polymorphism distribution and structural conservation in RNA-sensing *Toll-like* receptors 3, 7, and 8 in pigs. *Biochim Biophys Acta*. 2009.
- Muller A, et al. Evolution of rabbit haemorrhagic disease virus (RHDV) in the European rabbit (*Oryctolagus cuniculus*) from the Iberian Peninsula. *Vet Microbiol*. 2009. 135: 368-373.
- Muzio M, et al. Differential expression and regulation of *Toll-like* receptors (TLR) in human leukocytes: selective expression of TLR3 in dendritic cells. *J Immunol*. 2000. 164: 5998-6004.
- Nei M, Glazko GV. The Wilhelmine E. Key 2001 Invitational Lecture. Estimation of divergence times for a few mammalian and several primate species. *J Hered*. 2002. 93: 157-164.
- Nielsen R, Yang Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*. 1998. 148: 929-936.
- Nowotny N, Bascuñana CR, Ballagi-Pordany A, Gavier-Widen D, Uhlen M, Belak S. Phylogenetic analysis of rabbit haemorrhagic disease and European brown hare syndrome viruses by comparison of sequences from the capsid protein gene. *Arch. Virol*. 1997. 657-673.
- O'Brien AD, Rosenstreich DL, Scher I, Campbell GH, MacDermott RP, Formal SB. Genetic control of susceptibility to *Salmonella typhimurium* in mice: role of the LPS gene. *J Immunol*. 1980. 124: 20-24.
- O'Neill LA, Bowie AG. The family of five: TIR-domain-containing adaptors in *Toll-like* receptor signalling. *Nat Rev Immunol*. 2007. 7: 353-364.
- O'Neill LA, Bryant CE, Doyle SL. Therapeutic targeting of *Toll-like* receptors for infectious and inflammatory diseases and cancer. *Pharmacol Rev*. 2009. 61: 177-197.
- Offord V, Coffey TJ, Werling D. LRRfinder: a web application for the identification of leucine-rich repeats and an integrative *Toll-like* receptor database. *Dev Comp Immunol*. 2010. 34: 1035-1041.

- Ohlinger VF, Haas B, Thiel HJ. Rabbit hemorrhagic disease (RHD): characterization of the causative calicivirus. *Vet Res.* 1993. 24: 103-116.
- Ohlinger VF, Haas B, Meyers G, Weiland F, Thiel HJ. Identification and characterization of the virus causing rabbit hemorrhagic disease. *J Virol.* 1990. 64: 3331-3336.
- Okahira S, Nishikawa F, Nishikawa S, Akazawa T, Seya T, Matsumoto M. Interferon-beta induction through *Toll-like* receptor 3 depends on double-stranded RNA structure. *DNA Cell Biol.* 2005. 24: 614-623.
- Oshiumi H, Matsuo A, Matsumoto M, Seya T. Pan-vertebrate *Toll-like* receptors during evolution. *Curr Genomics.* 2008. 9: 488-493.
- Oshiumi H, Matsumoto M, Funami K, Akazawa T, Seya T. TICAM-1, an adaptor molecule that participates in *Toll-like* receptor 3-mediated interferon-beta induction. *Nat Immunol.* 2003. 4: 161-167.
- Parodi U. *The Rabbit in Contemporary Immunological Research.* Longman Scientific & Technical.1907.
- Parra F, Prieto M. Purification and characterization of a calicivirus as the causative agent of a lethal hemorrhagic disease in rabbits. *J Virol.* 1990. 64: 4013-4015.
- Pasteur L CC, Roux PPE, Thuillier L. *The Rabbit in Contemporary Immunological Research.:* Longman Scientific & Technical.1881.
- Pirher N, Ivicak K, Pohar J, Bencina M, Jerala R. A second binding site for double-stranded RNA in TLR3 and consequences for interferon activation. *Nat Struct Mol Biol.* 2008. 15: 761-763.
- Pond SL, Frost SD. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics.* 2005. 21: 2531-2533.
- Prasad AB, Allard MW, Green ED. Confirming the phylogeny of mammals by use of large comparative sequence data sets. *Mol Biol Evol.* 2008. 25: 1795-1808.
- Queney G, Ferrand N, Weiss S, Mougél F, Monnerot M. Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus cuniculus*). *Mol Biol Evol.* 2001. 18: 2169-2178.
- Rakoff-Nahoum S, Medzhitov R. *Toll-like* receptors and cancer. *Nat Rev Cancer.* 2009. 9: 57-63.
- Ranjith-Kumar CT, et al. Biochemical and functional analyses of the human *Toll-like* receptor 3 ectodomain. *J Biol Chem.* 2007. 282: 7668-7678.
- Rassa JC, Ross SR. Viruses and *Toll-like* receptors. *Microbes Infect.* 2003. 5: 961-968.
- Roelofs MF, Boelens WC, Joosten LA, Abdollahi-Roodsaz S, Geurts J, Wunderink LU, Schreurs BW, van den Berg WB, Radstake TR. Identification of small heat shock protein B8 (HSP22) as a novel TLR4 ligand and potential involvement in the pathogenesis of rheumatoid arthritis. *J Immunol.* 2006. 176: 7021-7027.

- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987. 4: 406-425.
- Salaun B, Coste I, Risoan MC, Lebecque SJ, Renno T. TLR3 can directly trigger apoptosis in human cancer cells. *J Immunol.* 2006. 176: 4894-4901.
- Salaun B, Lebecque S, Matikainen S, Rimoldi D, Romero P. *Toll-like* receptor 3 expressed by melanoma cells as a target for therapy? *Clin Cancer Res.* 2007. 13: 4565-4574.
- Sanarelli G. Das myxomatogene Virus. Beitrag zum Stadium der Krankheitserreger ausserhalb des Sichtbaren. *Zbl. Bakt.* 1898. 865-873.
- Sang Y, Yang J, Ross CR, Rowland RR, Blecha F. Molecular identification and functional expression of porcine *Toll-like* receptor (TLR) 3 and TLR7. *Vet Immunol Immunopathol.* 2008. 125: 162-167.
- Sanghavi SK, Shankarappa R, Reinhart TA. Genetic analysis of Toll/Interleukin-1 Receptor (TIR) domain sequences from rhesus macaque *Toll-like* receptors (TLRs) 1-10 reveals high homology to human TLR/TIR sequences. *Immunogenetics.* 2004. 56: 667-674.
- Schon M, Bong AB, Drewniok C, Herz J, Geilen CC, Reifenberger J, Benninghoff B, Slade HB, Gollnick H, Schon MP. Tumor-selective induction of apoptosis and the small-molecule immune response modifier imiquimod. *J Natl Cancer Inst.* 2003. 95: 1138-1149.
- Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A.* 1998. 95: 5857-5864.
- Seitz O. Glycopeptide synthesis and the effects of glycosylation on protein structure and activity. *Chembiochem.* 2000. 1: 214-246.
- Sims GE, Jun SR, Wu GA, Kim SH. Whole-genome phylogeny of mammals: evolutionary information in genic and nongenic regions. *Proc Natl Acad Sci U S A.* 2009. 106: 17077-17082.
- Singer MF. SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell.* 1982. 28: 433-434.
- Slack JL, Schooley K, Bonnert TP, Mitcham JL, Qwarnstrom EE, Sims JE, Dower SK. Identification of two major sites in the type I interleukin-1 receptor cytoplasmic region responsible for coupling to pro-inflammatory signaling pathways. *J Biol Chem.* 2000. 275: 4670-4678.
- Smid B, Valicek L, Rodak L, Stepanek J, Jurak E. Rabbit haemorrhagic disease: an investigation of some properties of the virus and evaluation of an inactivated vaccine. *Vet Microbiol.* 1991. 26: 77-85.
- Smits EL, Ponsaerts P, Berneman ZN, Van Tendeloo VF. The use of TLR7 and TLR8 ligands for the enhancement of cancer immunotherapy. *Oncologist.* 2008. 13: 859-875.
- Spaner DE, Masellis A. *Toll-like* receptor agonists in the treatment of chronic lymphocytic leukemia. *Leukemia.* 2007. 21: 53-60.

- Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*. 2002. 12: 43R-56R.
- Stanford MM, Werden SJ, McFadden G. Myxoma virus in the European rabbit: interactions between the virus and its susceptible host. *Vet Res*. 2007. 38: 299-318.
- Supajatura V, Ushio H, Nakao A, Okumura K, Ra C, Ogawa H. Protective roles of mast cells against enterobacterial infection are mediated by *Toll-like* receptor 4. *J Immunol*. 2001. 167: 2250-2256.
- Szajnik M, Szczepanski MJ, Czystowska M, Elishaev E, Mandapathil M, Nowak-Markwitz E, Spaczynski M, Whiteside TL. TLR4 signaling induced by lipopolysaccharide or paclitaxel regulates tumor survival and chemoresistance in ovarian cancer. *Oncogene*. 2009. 28: 4353-4363.
- Szczepanski MJ, Czystowska M, Szajnik M, Harasymczuk M, Boyiadzis M, Kruk-Zagajewska A, Szyfter W, Zeromski J, Whiteside TL. Triggering of *Toll-like* receptor 4 expressed on human head and neck squamous cell carcinoma promotes tumor development and protects the tumor from immune attack. *Cancer Res*. 2009. 69: 3105-3113.
- Tabeta K, et al. *Toll-like* receptors 9 and 3 as essential components of innate immune defense against mouse cytomegalovirus infection. *Proc Natl Acad Sci U S A*. 2004. 101: 3516-3521.
- Taganov KD, Boldin MP, Baltimore D. MicroRNAs and immunity: tiny players in a big field. *Immunity*. 2007. 26: 133-137.
- Takeda K, Kaisho T, Akira S. *Toll-like* receptors. *Annu Rev Immunol*. 2003. 21: 335-376.
- Takeuchi O, Akira S. Innate immunity to virus infection. *Immunol Rev*. 2009. 227: 75-86.
- Takeuchi O, Hoshino K, Akira S. Cutting edge: TLR2-deficient and MyD88-deficient mice are highly susceptible to *Staphylococcus aureus* infection. *J Immunol*. 2000. 165: 5392-5396.
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*. 2007. 24: 1596-1599.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994. 22: 4673-4680.
- Turner JD. A bioinformatic approach to the identification of bacterial proteins interacting with Toll-interleukin 1 receptor-resistance (TIR) homology domains. *FEMS Immunol Med Microbiol*. 2003. 37: 13-21.
- van der Loo W, Ferrand N, Soriguer RC. Estimation of gene diversity at the b locus of the constant region of the immunoglobulin light chain in natural populations of European rabbit (*Oryctolagus cuniculus*) in Portugal, Andalusia and on the Azorean Islands. *Genetics*. 1991. 127: 789-799.

- van der Loo W, Mougél F, Bouton C, Sanchez MS, Monnerot M. The allotypic patchwork pattern of the rabbit IGKC1 allele b5wf: genic exchange or common ancestry? *Immunogenetics*. 1999. 49: 7-14.
- Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J. 1999. *Essentials of Glycobiology*. New York: Cold Spring Harbor.
- Villafuerte R, Calvete C, Blanco JC, Lucientes J. Incidence of viral haemorrhagic disease in wild rabbit populations in Spain. *Mammalia*. 1995. 651-659.
- Villares JLA, Pellitero AL, Garcia LMPO. Enfermedad vírica hemorrágica del conejo en España. *Medicina Veterinária*. 1988. 645-650.
- Wang T, Town T, Alexopoulou L, Anderson JF, Fikrig E, Flavell RA. *Toll-like* receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. *Nat Med*. 2004. 10: 1366-1373.
- Wassermann A NA, Bruck C. Eine serodiagnostische reaktion bei Syphilis in S D, ed. *The Rabbit in Contemporary Immunological Research*, Longman Scientific & Technical. 1906.
- West AP, Koblansky AA, Ghosh S. Recognition and signaling by *Toll-like* receptors. *Annu Rev Cell Dev Biol*. 2006. 22: 409-437.
- Wlasiuk G, Nachman MW. Adaptation and constraint at *Toll-like* receptors in primates. *Mol Biol Evol*. 2010. 27: 2172-2186.
- Wong WS, Yang Z, Goldman N, Nielsen R. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics*. 2004. 168: 1041-1051.
- Xu Y, Tao X, Shen B, Horng T, Medzhitov R, Manley JL, Tong L. Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature*. 2000. 408: 111-115.
- Yamamoto M, Sato S, Mori K, Hoshino K, Takeuchi O, Takeda K, Akira S. Cutting edge: a novel Toll/IL-1 receptor domain-containing adapter that preferentially activates the IFN-beta promoter in the *Toll-like* receptor signaling. *J Immunol*. 2002. 169: 6668-6672.
- Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci*. 1997. 13: 555-556.
- Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007. 24: 1586-1591.
- Yang Z, Wong WS, Nielsen R. Bayes empirical bayes inference of amino acid sites under positive selection. *Mol Biol Evol*. 2005. 22: 1107-1118.
- Yang Z, Nielsen R, Goldman N, Pedersen AM. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*. 2000. 155: 431-449.
- Yoneyama M, Fujita T. Recognition of viral nucleic acids in innate immunity. *Rev Med Virol*. 2010. 20: 4-22.
- Zhang, Feng, Li CL, Wang XY, Chang ZL. NF-kappa B regulates the LPS-induced expression of interleukin 12 p40 in murine peritoneal macrophages:

roles of PKC, PKA, ERK, p38 MAPK, and proteasome. *Cell Immunol.* 2000. 204: 38-45.

Zhang, et al. TLR3 deficiency in patients with herpes simplex encephalitis. *Science.* 2007. 317: 1522-1527.

Anexos

Anexo 1

Resumo dos domínios encontrados no gene TLR3 de mamíferos (SMART).

Legenda: Os números representados na tabela significam a quantidade do respectivo domínio detectada na sequência em causa. A negrito está assinalado o Coelho-bravo e os mamíferos com maior semelhança em termos de LRRs. As siglas utilizadas têm o seguinte significado: LRR – Repetições ricas em leucina; NT – N terminal; CT – C terminal; TYP – típico; (---), domínio não detectado.

Organismo	Domínios						
	Transmembranar	LRR-NT	LRRs total	LRRs TYP	LRR-CT	Transmembranar	TIR
Boi	---	1	17	3	1	1	1
Cão	---	1	18	2	1	1	1
Cavalo	---	1	17	2	1	1	1
Chimpanzé	---	1	17	2	1	1	1
Coelho-bravo	---	---	16	3	1	1	1
Gato	---	1	18	2	1	1	1
Golfinho	---	---	16	3	1	1	1
Homem	---	1	17	2	1	1	1
Javali	---	1	16	1	1	1	1
Macaco-rhesus	---	1	17	2	1	1	1
Raposa voadora	---	1	17	2	1	1	1
Musaranho	---	---	15	2	1	1	1
Orangotango	---	1	17	2	1	1	1
Ouriço-cacheiro	---	---	16	3	1	1	1
Ovelha	---	1	18	3	1	1	1
Porquinho-da-índia	---	1	18	2	1	1	1
Ratazana	1	1	16	3	1	1	1
Rato	---	1	17	4	1	1	1
Társio-das-Filipinas	---	1	15	1	1	1	1

Anexo 2

Resumo dos domínios encontrados no gene TLR3 de mamíferos (LRRfinder).

	LRR-NT	LRRs total	LRR-CT	Transmembranar	TIR
Boi	1	24	1	1	1
Cão	*	22	*	*	*
Cavalo	1	24	1	1	1
Chimpanzé	1	24	1	1	1
Coelho-bravo	1	24	1	1	1
Gato	1	24	1	1	1
Golfinho	*	21	*	*	*
Homem	1	24	1	1	1
Javali	1	24	1	1	1
Macaco-rhesus	1	24	1	1	1
Raposa voadora	*	24	*	*	*
Musaranho	*	21	*	*	*
Orangotango	*	24	*	*	*
Ouriço-cacheiro	*	21	*	*	*
Ovelha	1	24	1	1	1
Porquinho-da-índia	1	24	1	1	1
Ratazana	1	24	1	1	1
Rato	1	24	1	1	1
Társio-das-Filipinas	*	23	*	*	*

*Não foi possível identificar os domínios por falta desta sequência na base de dados do programa

Anexo 3

Alterações de aminoácidos nas populações de *Oryctolagus cuniculus* analisadas. As posições provavelmente localizadas no ectodomínio e assinaladas com * localizam-se em LRRs.

Posição	Região Provável ^a	Possíveis aminoácidos	Polaridade	Carga	Local potencial de N-glicosilação
46	LRR-NT	I	Hidrofóbico	Neutra	---
		V	Hidrofóbico	Neutra	---
101*	Ectodomínio	L	Hidrofóbico	Neutra	---
		S	Hidrofílico	Neutra	---
132*	Ectodomínio	M	Hidrofóbico	Neutra	---
		T	Hidrofílico	Neutra	---
204*	Ectodomínio	E	Hidrofílico	Negativa	---
		D	Hidrofílico	Negativa	---
211*	Ectodomínio	K	Hidrofílico	Positiva	---
		R	Hidrofílico	Positiva	---
298*	Ectodomínio	L	Hidrofóbico	Neutra	---
		F	Hidrofóbico	Neutra	---
300*	Ectodomínio	H	Hidrofílico	Positiva	---
		P	Hidrofóbico	Neutra	---
315*	Ectodomínio	S	Hidrofílico	Neutra	---
	Ectodomínio	P	Hidrofóbico	Neutra	---
353 *	Ectodomínio	Q	Hidrofílico	Neutra	---
	Ectodomínio	H	Hidrofílico	Positiva	---
480 *	Ectodomínio	V	Hidrofóbico	Neutra	---
		A	Hidrofóbico	Neutra	---
560*	Ectodomínio	K	Hidrofílico	Positiva	---
		E	Hidrofílico	Negativa	---
575*	Ectodomínio	L	Hidrofóbico	Neutra	---
		F	Hidrofóbico	Neutra	---
708	Domínio Transmembranar	M	Hidrofóbico	Neutra	---
		I	Hidrofóbico	Neutra	---
716	Domínio Transmembranar	L	Hidrofóbico	Neutra	---
		F	Hidrofóbico	Neutra	---
864	Domínio TIR	A	Hidrofóbico	Neutra	---
		S	Hidrofílico	Neutra	Sim

^a De acordo com a delimitação dos domínios obtida através do programa LRRfinder.

Anexo 4

Caracterização dos locais de glicosilação no grupo de mamíferos analisados.

	51-54	56-59	69-72	123-126	193-196	233-236	244-247	249-252	262-265	269-272	272-275	277-280	288-291	312-315	345-348	382-385	395-398	410-413	421-424	491-494	504-507	557-560	633-636	659-662	664-667	706-709	732-735
Human	x	x	x	x	x		x	x	x		x		x				x	x			x		x	x		x	x
Boi	x	x	x	x	x		x	x			x		x				x						x	x	x		x
Cão	x	x	x	x	x		x	x	x		x		x				x	x					x	x	x	x	x
Cavalo	x	x	x	x	x		x	x	x		x		x				x	x			x		x	x	x	x	x
Chimpanzé	x	x	x	x	x		x	x	x		x		x				x	x			x		x	x		x	x
Coelho-bravo	x	x	x	x	x		x	x	x		x			x			x	x					x	x	x	x	x
Gato	x	x	x		x	x	x	x	x		x		x				x	x					x	x	x	x	x
Golfinho	x	x	x	x	x		x	x	x		x		x				x	x			x		x	x	x	x	x
Javali	x	x	x		x		x	x	x		x		x				x	x			x		x	x			x
Macaco-rhesus	x	x	x	x	x		x	x	x		x		x				x	x			x		x			x	x
Raposa voadora	x		x		x			x	x	x	x	x	x		x		x	x			x		x	x	x	x	x
Musaranho	x	x	x		x		x	x	x		x		x				x	x					x	x	x		
Orangotango	x	x	x	x	x		x	x	x		x		x				x	x			x		x	x		x	x
Ouriço-cacheiro	x	x	x		x		x	x			x		x				x	x			x		x	x	x	x	x
Ovelha	x	x	x	x	x		x	x			x		x			x	x	x					x	x	x		x
Porquinho-da-índia	x	x	x	x	x		x	x	x		x		x				x	x					x	x			
Ratazana	x	x	x	x	x		x	x	x		x						x	x			x			x			x
Rato	x	x	x	x	x		x	x			x		x				x	x	x			x		x	x		x
Társio-das-Filipinas	x	x	x		x		x		x		x		x		x		x	x			x	x	x	x	x	x	

Os espaços preenchidos de azul dizem respeito a ausência desse local de glicosilação. O x, indica presença de local de glicosilação. Os espaços demilitados a linha mais escura correspondem a locais de glicosilação que não existem no homem.

Anexo 5

Características dos aminoácidos possíveis nos codões sobre selecção positiva com probabilidades >95%

Posição	Região provável ^a	Possíveis aminoácidos	Polaridade	Carga
3	Sinal	Q	Hidrofílico	Neutra
		G	Hidrofílico	Neutra
		R	Hidrofílico	Positiva
		K	Hidrofílico	Positiva
		W	Hidrofóbico	Neutra
		L	Hidrofóbico	Neutra
4	Sinal	T	Hidrofílico	Neutra
		C	Hidrofílico	Neutra
		R	Hidrofílico	Positiva
		S	Hidrofílico	Neutra
		P	Hidrofílico	Neutra
		Y	Hidrofóbico	Neutra
		K	Hidrofílico	Positiva
		D	Hidrofílico	Negativa
12	Sinal	N	Hidrofílico	Neutra
		G	Hidrofílico	Neutra
		L	Hidrofóbico	Neutra
		S	Hidrofílico	Neutra
16	Sinal	W	Hidrofóbico	Neutra
		P	Hidrofílico	Neutra
		S	Hidrofílico	Neutra
		T	Hidrofílico	Neutra
		L	Hidrofóbico	Neutra
		F	Hidrofóbico	Neutra
79 *	Ectodomínio	S	Hidrofílico	Neutra
		I	Hidrofóbico	Neutra
		L	Hidrofóbico	Neutra
		T	Hidrofílico	Neutra
		V	Hidrofóbico	Neutra
86 *	Ectodomínio	T	Hidrofílico	Neutra
		S	Hidrofílico	Neutra
		C	Hidrofílico	Neutra
97 *	Ectodomínio	K	Hidrofílico	Positiva
		I	Hidrofóbico	Neutra
		S	Hidrofílico	Neutra
		N	Hidrofílico	Neutra
		Q	Hidrofílico	Neutra
		V	Hidrofóbico	Neutra
120*	Ectodomínio	A	Hidrofóbico	Neutra
		V	Hidrofóbico	Neutra
		I	Hidrofóbico	Neutra
		M	Hidrofóbico	Neutra
		F	Hidrofóbico	Neutra
		T	Hidrofílico	Neutra
		A	Hidrofóbico	Neutra
185*	Ectodomínio	V	Hidrofóbico	Neutra
		S	Hidrofílico	Neutra
		T	Hidrofílico	Neutra
		T	Hidrofílico	Neutra

		L	Hidrofóbico	Neutra
258*	Ectodomínio	S	Hidrofílico	Neutra
		N	Hidrofílico	Neutra
		T	Hidrofílico	Neutra
		I	Hidrofóbico	Neutra
		V	Hidrofóbico	Neutra
261*	Ectodomínio	S	Hidrofílico	Neutra
		L	Hidrofóbico	Neutra
		H	Hidrofílico	Positiva
		D	Hidrofílico	Negativa
		Y	Hidrofóbico	Neutra
		N	Hidrofílico	Neutra
266*	Ectodomínio	T	Hidrofílico	Neutra
		S	Hidrofílico	Neutra
		M	Hidrofóbico	Neutra
		R	Hidrofílico	Positiva
		L	Hidrofóbico	Neutra
		V	Hidrofóbico	Neutra
269 *	Ectodomínio	L	Hidrofóbico	Neutra
		S	Hidrofílico	Neutra
		D	Hidrofílico	Negativa
		F	Hidrofóbico	Neutra
		V	Hidrofóbico	Neutra
		I	Hidrofóbico	Neutra
		A	Hidrofóbico	Neutra
		T	Hidrofílico	Neutra
273 *	Ectodomínio	W	Hidrofóbico	Neutra
		Q	Hidrofílico	Neutra
		H	Hidrofílico	Positiva
285*	Ectodomínio	N	Hidrofílico	Neutra
		S	Hidrofílico	Neutra
		K	Hidrofílico	Positiva
		G	Hidrofílico	Neutra
295*	Ectodomínio	A	Hidrofóbico	Neutra
		S	Hidrofílico	Neutra
		V	Hidrofóbico	Neutra
		R	Hidrofílico	Positiva
		T	Hidrofílico	Neutra
456*	Ectodomínio	E	Hidrofílico	Negativa
		R	Hidrofílico	Positiva
		G	Hidrofílico	Neutra
		D	Hidrofílico	Negativa
		K	Hidrofílico	Positiva
		A	Hidrofóbico	Neutra
473 *	Ectodomínio	R	Hidrofílico	Positiva
		T	Hidrofílico	Neutra
		S	Hidrofílico	Neutra
		K	Hidrofílico	Positiva
		P	Hidrofílico	Neutra
		E	Hidrofílico	Negativa
506*	Ectodomínio	R	Hidrofílico	Positiva
		Y	Hidrofóbico	Neutra
		P	Hidrofílico	Neutra
		F	Hidrofóbico	Neutra
		H	Hidrofílico	Positiva
		G	Hidrofílico	Neutra

586 *	Ectodomínio	C	Hidrofílico	Neutra
		F	Hidrofóbico	Neutra
		S	Hidrofílico	Neutra
		R	Hidrofílico	Positiva
		L	Hidrofóbico	Neutra
663	LRR-CT	E	Hidrofílico	Negativa
		Q	Hidrofílico	Neutra
		S	Hidrofílico	Neutra
		I	Hidrofóbico	Neutra
		G	Hidrofílico	Neutra
		K	Hidrofílico	Positiva
		N	Hidrofílico	Neutra
		D	Hidrofílico	Negativa
712	Domínio Transmembranar	I	Hidrofóbico	Neutra
		M	Hidrofóbico	Neutra
		T	Hidrofílico	Neutra
		V	Hidrofóbico	Neutra
		F	Hidrofóbico	Neutra
749	Domínio TIR	R	Hidrofílico	Positiva
		T	Hidrofílico	Neutra
		A	Hidrofóbico	Neutra
		K	Hidrofílico	Positiva
		S	Hidrofílico	Neutra
		G	Hidrofílico	Neutra
780	Domínio TIR	E	Hidrofílico	Negativa
		Q	Hidrofílico	Neutra
		K	Hidrofílico	Positiva

^a De acordo com a delimitação dos domínios obtida através do programa LRRfinder.

Anexo 6

Alinhamento da sequência do domínio LRR-NT de Homem e Coelho-bravo
 O quadrilátero está a delimitar a região correspondente ao domínio LRR-NT no Homem.
 A verde está assinalado um aminoácido essencial para a ligação do dsRNA e a vermelho os aminoácidos da sequência de Coelho-bravo que diferem da sequência do Homem.

