U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

GENETIC CHARACTERIZATION OF THE INNATE IMMUNE SYSTEM OF LAGOMORPHS (ILS, CCLS)

FABIANA MARISA VIEIRA DAS NEVES

TESE DE DOUTORAMENTO APRESENTADA AO INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR DA UNIVERSIDADE DO PORTO EM PATOLOGIA E GENÉTICA MOLECULAR

2017

FABIANA MARISA VIEIRA DAS NEVES

GENETIC CHARACTERIZATION OF THE INNATE IMMUNE SYSTEM OF LAGOMORPHS (ILS, CCLS)

Tese de Candidatura ao grau de Doutor em Patologia e Genética Molecular submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

ORIENTADOR:
PROFESSOR DOUTOR PEDRO JOSÉ DE CASTRO ESTEVES
Professor Auxiliar Convidado
Departamento de Biologia
Faculdade de Ciências da Universidade do Porto
Investigador principal
Centro de Investigação em Biodiversidade e Recursos Genéticos

CO-ORIENTADOR:
PROFESSOR DOUTOR PAULO MANUEL DE CASTRO PINHO E COSTA
Professor Auxiliar Convidado
Departamento de Patologia e Imunologia Molecular
Instituto de Ciências Biomédicas Abel Salazar – Universidade do Porto
Investigador Proncipal
Departamento de Genética Humana
Instituto Nacional de Saúde Dr Ricardo Jorge

Ao meu pimpolhito…

FINANCIAL SUPPORT:

This study was supported by Fundação para a Ciência e Tecnologia (FCT) through a PhD grant (SFRH/BD/81916/2011) financed by Programa Operacional and União Europeia.

LISTA DE PUBLICAÇÕES

Ao abrigo do disposto do nº 2, alínea a) do artigo 31º do Decreto-Lei n.º115/2013 de 7 de Agosto fazem parte integrante desta tese de doutoramento os seguintes trabalhos já publicados ou submetidos para publicação:

Artigo I

Neves F, Abrantes J, Steinke JW, Esteves PJ. (2014) Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals. Innate Immunity, 20(2): 184–191. doi: 10.1177/1753425913486687

Artigo II

Neves F, Abrantes J, Pinheiro A, Almeida T, Costa PP, Esteves PJ. (2014) Convergent evolution of IL-6 in two leporids (*Oryctolagus* and *Pentalagus*) originated an extended protein. Immunogenetics, 66(9):589–595. doi: 10.1007/s00251-014-0787-0

Artigo III

Neves F, Abrantes J, Lissovsky AA, Esteves PJ. (2015) Pseudogenization of CCL14 in the Ochotonidae (pika) family. Innate Immunity 21(6):647-654. doi: 10.1177/1753425915577455

Artigo IV

Neves F, Abrantes J, Almeida T, de Matos AL, Costa PP, Esteves PJ. (2015) Genetic characterization of interleukins (IL-1α, IL-1β, IL-2, IL-4, IL-8, IL-10, IL-12A, IL-12B, IL-15 and IL-18) with relevant biological roles in lagomorphs. Innate Immunity 21(8):787-801. doi: 10.1177/1753425915606209

Artigo V

Neves F, Abrantes J, Almeida T, Costa PP, Esteves PJ. (2015) Evolutionary Insights into IL17A in Lagomorphs. Mediators of Inflammation, vol. 2015, Article ID 367670, 7 pages, 2015. doi: 10.1155/2015/367670

Artigo VI

Neves F, Abrantes J, Esteves PJ. (2016) Evolution of CCL11: genetic characterization in lagomorphs and evidence of positive and purifying selection in mammals. Innate immunity 22(5):336-343. doi: 10.1177/1753425916647471

Artigo VII

Neves F, Abrantes J, Lopes AM, Magalhães MJ, Esteves PJ. Evolution of CCL16 in Glires (Rodentia and Lagomorphs) shows an unusual random pseudogenization pattern. (*In preparation*)

AGRADECIMENTOS

Não podia finalizar este trabalho sem agradecer às pessoas que para ele contribuíram. Esta tese retrata não só o meu trabalho, mas também o contributo individual daqueles que estiveram ao meu lado, pessoal e profissionalmente. A todos os que me ajudaram a chegar ao final desta etapa, o meu mais sincero obrigado.

Ao meu orientador, Professor José Pedro Esteves pela oportunidade que me deu, por ter sido crucial no desenho e construção deste projecto. Agradeço o apoio, a inspiração, a paciência. e o fato de acreditar em mim e nas minhas capacidades para levar a cabo este projeto. Obrigada por todas as ideias pertinentes que foi dando ao longo deste trabalho, pelo seu rigor e espírito crítico constante, por ser um poço de saber inesgotável e por me abrir o caminho para o crescimento pessoal e profissional.

Ao meu co-orientador, Professor Paulo Pinho e Costa por me ter aceite como estudante e ajudar a percorrer este caminho.

À minha chefinha, Doutora Joana Abrantes, agradeço todo o apoio ao longo destes últimos anos. Obrigada pelo apoio, disponibilidade, pelo trabalho, dedicação e por todos os bons momentos científicos e pessoais vividos.

À Fundação para a Ciência e a Tecnologia (FCT), agradeço a concessão de uma Bolsa de Doutoramento (SFRH/BD/ 81916/2011).

Ao Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO) e ao Centro de Testagem Molecular (CTM), agradeço a ajuda técnica proporcionada e pelas condições laboratoriais proporcionadas. Ao Bruno Maia, à Diana, à Filipina, ao Javi, à Jolyta, à Patrícia, à Sara João, à Sandrolina à Sofia Mourão, à Sofia Silva, à Susana obrigada por todos os bons momentos, tanto de descontracção, e boa disposição como em todo o apoio técnico e laboratorial que precisei. À Sara Ferreira à Sandra Rodrigues e à Márcia, agradeço a prontidão com que sempre me ajudaram na resolução de diversas situações.

Agradeço de um modo muito especial ao pessoal do CIBIO, e que cada um à sua maneira me fez sentir parte de um grupo e ajudou com que estes 4 anos fossem muito descontraídos e cheios de alegria, calor humano e bons momentos. Ao pessoal do Aquário: ao Gui, ao João Maia, à Licas, à Rita, à Té, ao Ubuntu e à Vânia pelos momentos descontraídos e excelentes miminhos culinários. Às

minhas companheiras de viagens e de pequeno-almoço, Soraia e Guida, um muitíssimo obrigada. Obrigada Guidinha pela valiosa ajuda nas figuras.

Aos meus “imídios” (Joana, Guida, Ana M, AnaP, PP; Tereza, Maria e Diogo) pela imprescindível ajuda e acompanhamento mas também companheirismo e a boa disposição.

À Teresinha “do bar” que tão bem me alimentou tanto física como psicologicamente um grande obrigada.

A todos os meus amigos que sempre foram estando presente ao longo destes 4 anos e que, de diferentes maneiras demonstraram a sua amizade, preocupação e apoio (Brunito, Sapinhos, Helena, Marlene, Stéphanie, Anabela, Tixa).

E, por fim, às pessoa mais importantes da minha vida:

À minha Mãe. A minha força. Aquela que acredita em mim incondicionalmente e a quem devo a pessoa em que me tornei. Sem ti nada na minha vida teria sido possível. O teu amor, o teu sacrifício, a tua dedicação e a tua motivação são e serão sempre o meu pilar. És sem dúvida o meu Modelo a seguir e espero conseguir continuar a deixar-te orgulhosa das minhas decisões.

À minha irmã, que me mostrou que nada é impossível, que vale a pena lutar pelos sonhos e seguir aquilo que acreditamos. Por toda o caminho que desbravou para mim desde sempre e por toda a amizade, amor e apoio.

Ao Filipe pela amizade e amor, pelo carinho e companheirismo, por não me deixares ir a baixo nos maus momentos e me fazeres crer que tudo iria sempre correr bem. Pela paciência para me aturares (e bem sei que não é nada fácil) e, principalmente, por tornares esta vida muito melhor.

Ao meu pimpolhito que todos os dias me ensina a ser e a tornar-me melhor pessoa. Sem dúvida o meu maior desafio e a minha maior e melhor prenda de sempre.

Index of contents

Index of figures

Chapter 1. General Introduction

Chapter 3. Innate immunity – Evolution of interleukins in lagomorphs

Chapter 4. Innate immunity – Genetic aspects of CC motif chemokines in lagomorphs

Index of tables

Chapter 1. General Introduction

Chapter 3. Innate immunity – Evolution of interleukins in lagomorphs

Chapter 4. Innate immunity – Genetic aspects of CC motif chemokines in lagomorphs

ABBREVIATIONS

2∆lnL - twice the difference in the natural logs of the likelihoods

3D - Three-dimensional

aa – amino acid

AIDS - Acquired immune deficiency syndrome

APOBEC - apolipoprotein B mRNA editing enzyme catalytic

B.C. – Before Christ's

BEB - Bayes Empirical Bayes

BP – Binding protein

bp – base pair

CCL – C-C chemokine ligand

CCR – C-C chemokine receptor

CD4 - cluster of differentiation 4

CDC - Center of Disease Control and Prevention

cDNA – complementary DNA

CDS – Coding sequence

Chr - Chromossome

Cks – Cytokines

CTLA - Cytotoxic T lymphocyte associated antigen

CXCL - C-X-C chemokine ligand

CXCR – C-X-C- chemokine receptor

DAMPs - Damage-associated molecular patterns

DC - dendritic cells

DiANNA - DiAminoacid Neural Network Application

DMSO - Dimethyl sulfoxide

dN – Nonsynonymous substitutions

DNA - Deoxyribonucleic acid

dS – Synonymous substitutions

Ebi3 - Epstein-Barr virus induced gene 3

eiF2 - Eukaryotic initiation factor 2

FEL - Fixed-Effect Likelihood

FHF - Fulminant hepatic failure

FUBAR – Fast Unconstrained Bayesian AppRoximation

gDNA – genomic DNA

GRO - Growth related oncogene

HCC - Hemofiltrate CC-Chemokine

HCV - Hepatitis C virus

HIV - Human Immunodeficiency Virus

HMGA2 - High Mobility Group AT-Hook 2

iFEL – Internal Branch Fel

IFN – Interferon

IP10 - IFNγ-inducible protein 10

Ig – Immunoglobulin

IgLλ – Immunoglobulin light chain lambda locus

ILR – Interleukine receptor

ILs – Interleukins

Kb – kilobases

LEC – Liver-expressed chemokine

LPS – Lipopolysaccharide

LRT - Likelihood ratio test

MC1R - Melanocortin 1 receptor

MCP - Monocyte chemoattractant protein

MEME - Mixed Effects Model of Evolution

MIP - Macrophage inflammatory protein

ML - Maximum Likelihood

mRNA - Messenger RNA

MSMD - Mendelian susceptibility to mycobacterial disease

MUSCLE - MUltiple Sequence Comparison by Log Expectation

Mya – Million years ago

MYXV – Myxoma virus

NK – Natural killer

NLRs - NOD-like receptors

PAMPs - Pathogen-associated molecular patterns

PCR - Polymerase chain reaction

PDB - Protein data bank

PGRPs . Peptidoglycan recognition proteins

PKR – Protein Kinase R

PRRs - Pattern recognition receptors

REL – Random Effect Likelihood

RELIK - Rabbit Endogenous Lentivirus type K

RHD – Rabbit hemorrhagic disease

RHDV – Rabbit hemorrhagic disease virus

RNA - Ribonucleic acid

SLAC – Single Likelihood Ancestor Counting

Th – T helper

TLRs - Toll-like receptors

TNF - Tumor necrosis factor

Treg – T regulatory

VDHC – Vírus da doença hemmorágica do coelho

VM – Virus do mixoma

RESUMO

As citocinas são proteínas do sistema imunitário inato que estão envolvidas em vários processos da resposta imunológica. De forma a diversificar e adaptar-se rapidamente a mudanças nas condições normais, as citocinas podem sofrer alterações. Até à data, a descrição de interleucinas e de ligandos das quimiocinas em lagomorfos está quase unicamente restringida ao coelho-bravo. De forma a complementar a informação disponível para os lagomorfos, caracterizou-se a organização genómica, a diversidade genética e a evolução das interleucinas (ILs) e dos ligandos das quimiocinas com importância biológica nos lagomorfos. Estes genes foram analisados para diversos géneros de lagomorfos pertencentes às famílias Ochotonidae e Leporidae.

De forma a compreender as alterações que têm acompanhado a evolução destes genes nos lagomorfos, as sequências obtidas neste trabalho foram comparadas com as sequências publicamente disponíveis para outros mamíferos. A deteção de assinaturas de seleção positiva com recurso a métodos de máxima verosimilhança em 46 ILs de diferentes mamíferos revelou que 60% destas proteínas têm codões sob seleção positiva. Porém, o número e a percentagem de codões sob seleção positiva variaram entre as diferentes ILs. Dada a importância destas proteínas na ativação da resposta imunitária, era esperado que a sua estrutura e função fossem mantidas ao longo de diferentes linhagens evolutivas e, consequentemente, seriam observados padrões de seleção negativa. No entanto, os resultados obtidos neste trabalho contradizem esta previsão, o que pode ser explicado pela multiplicidade de processos biológicos em que as interleucinas estão envolvidas.

Das ILs estudadas neste trabalho, destaca-se a IL6, com um padrão evolutivo único. Estudos anteriores demonstraram que a IL6 do coelho-bravo difere de outros mamíferos, nomeadamente de outros dois géneros de leporídeos (*Sylvilagus* spp. e *Lepus* spp.) cuja divergência ocorreu há aproximadamente 12 milhões de anos. Esta diferença consiste na extensão desta proteína em mais 27 aminoácidos. Neste trabalho, estudámos esta proteína em outros quatro géneros de leporídeos (coelho-pigmeu, coelho-bosquímano, coelho-de-Amami e coelho-zacatuche) e na pika-americana. No coelho-de-Amami, detetou-se uma deleção do codão stop que provoca uma extensão da IL6 em 17 aminoácidos. Estes

resultados indicam que a extensão da IL6 ocorreu de forma independente no coelho-bravo e no coelho-de-Amami: uma aconteceu entre 2 a 8 milhões de anos atrás no ancestral do coelho-bravo, e a outra ocorreu no ancestral do coelho-de-Amami há um máximo de 9 milhões de anos. A ausência desta extensão na IL6 no coelho-bosquímano, género mais próximo do coelho-bravo, mostra que este evento aconteceu por convergência, sugerindo alguma relevância funcional. Esta hipótese é reforçada pela presença de quatro cisteínas extra na extensão do coelho-europeu que, de acordo com as nossas previsões de modelação, alteram significativamente a estrutura da IL6.

As populações naturais de coelho-bravo têm sido afetadas por duas doenças virais, a doença hemorrágica do coelho, causada pelo vírus da doença hemorrágica do coelho (VDHC), e a mixomatose causada pelo vírus do mixoma (VM). Estudos anteriores com coelhos infetados com VDHC detetaram alterações na expressão das interleucinas IL1, IL2, IL6, IL8 e IL10. Alterações na expressão da IL4, IL12, IL15 e IL18 foram relacionadas com a infeção pelo vírus do mixoma. Neste contexto, caracterizamos estas ILs em seis espécies de lagomorfos: coelho-bravo, coelho-pigmeu, coelho-do-chaparral, coelho-da-Flórida, lebre-europeia e pika-americana. Em geral, estas ILs são conservadas entre os lagomorfos. No entanto, para a pika-americana verificámos algumas diferenças, nomeadamente na localização do codão stop na IL1α e na IL2, na existência de um transcrito diferente na IL8 e num elevado número de cisteínas na IL1β. Adicionalmente, foram também detetadas alterações em locais de N-glicosilação na IL1, IL10, IL12B e IL15.

O coelho-bravo e a lebre-europeia são afetados por uma doença bacteriana, a tularemia, cujo agente etiológico é a bactéria *Francisella tularensis*. Anteriormente, encontrou-se uma associação entre a IL17A e a defesa do hospedeiro contra esta bactéria. O estudo da IL17A em cinco géneros de lagomorfos mostrou uma similaridade de 97-99% entre leporídeos e ~88% entre os leporídeos e a pika-americana. Em geral, a estrutura da IL17A é muito conservada nos lagomorfos. No entanto, no codão 88, um dos locais de interação entre a IL17A e o seu recetor IL17RA, existe uma mutação Arg>Pro que só ocorre em coelhos e lebres-europeias, as duas espécies infetadas pela *F. tularensis*. Esta mutação pode induzir alterações críticas na estrutura e conformação da IL17A e consequentemente modificar a sua função.

A conversão génica entre os receptores das quimiocinas CCR5 e CCR2 descrita para o coelho-bravo, o coelho-de-Amami e o coelho-bosquímano, está ausente em *Sylvilagus* spp. e *Lepus* spp., o que sugere que este evento aconteceu há aproximadamente 9 milhões de anos no ancestral do coelho-bravo, coelho-de-Amami e coelho-bosquímano. Posteriormente, estudos realizados nos ligandos do CCR5 demonstraram que as quimiocinas CCL3, CCL4 e CCL5 são funcionais em leporídeos e evoluíram sob seleção negativa. Curiosamente, a evolução do CCL8 tem um padrão consistente com a conversão génica observada para o CCR5. O CCL8 é um pseudogene nas espécies em que a conversão génica CCR5-CCR2 ocorreu, sendo funcional em espécies em que este evento não se verificou. Esta relação de causa-efeito levou ao estudo dos restantes ligandos do CCR5, nomeadamente o CCL11, o CCL14 e o CCL16. Todos os lagomorfos têm um CCL11 potencialmente funcional. Contudo, o coelho-pigmeu tem uma mutação no codão stop codificando uma proteína mais longa. Através de abordagens de máxima verosimilhança, foram detetadas assinaturas de seleção negativa e de seleção positiva. A seleção negativa foi detetada em locais importantes para a ligação e ativação dos recetores e pode resultar das restrições funcionais da proteína. A seleção positiva, ao resultar num aumento da frequência de mutações vantajosas para a proteína, poderá estar associada com uma melhoria da resposta do hospedeiro contra vários agentes.

Os leporídeos têm um CCL14 funcional, enquanto em algumas espécies de pika este gene é um pseudogene. De facto, para a pika-americana, pika-alpina, pika-de-pallas e pika-de-turuchan, o CCL14 é um pseudogene devido a uma mutação Met>Thr no codão de iniciação. A pika-do-Norte apresenta dois alelos (um funcional e um pseudogene) e na pika-da-estepe o CCL14 é um pseudogene devido a uma inserção de 7 pares de bases. Adicionalmente, a análise das sequências de CCL14 de vários mamíferos revelou seis codões sob seleção positiva em regiões cruciais para a ativação do CCL14 e a ligação a recetores. Isto sugere que o CCL14 tem um importante papel biológico nos mamíferos que se perdeu na família Ochotonidae, nomeadamente nos subgéneros Pika e Lagotona.

O gene CCL16 foi identificado como pseudogene no coelho-bravo e em alguns roedores como rato, ratinho e porquinho-da-Índia, enquanto no esquilo parece ser funcional. Para clarificar a evolução deste gene na superordem Glires,

composta por roedores e lagomorfos, foram sequenciados seis géneros de leporídeos e nove espécies de pikas. Adicionalmente, recuperamos as sequências de CCL16 de roedores disponíveis em bases de dados públicas. Os resultados da análises dessas sequências sugerem que na superordem Glires o CCL16 sofreu vários eventos independentes de pseudogenização. Nos lagomorfos, em todos os leporídeos, com a exceção dos *Sylvilagus* spp., o CCL16 é um pseudogene devido a uma mutação não-sinónima no codão 45, o que conduz a um codão stop prematuro. Isto sugere que o gene sofreu a pseudogenização no ancestral dos leporídeos há cerca de 14 milhões de anos. Curiosamente, no coelho-mexicano, no coelho-brasileiro e no coelho-da-Flórida, o CCL16 não possui esta mutação. Nestas espécies, a cisteína 45 também sofreu uma mutação, codificando uma lisina. Apesar de codificar uma proteína putativamente funcional, o CCL16 não é expresso no coelho-da-Flórida. A alteração de uma das cisteínas características do CCL16 para uma lisina pode ter levado à perda de função da proteína. Em todas as espécies de pika, com a exceção de um alelo da pika-de-pallas, o CCL16 está intacto, sendo expresso na pika-americana. Nos roedores, em alguns membros das famílias Muridae (rato e ratinho), Heteromyidae (rato-canguru) e Caviidae (porquinho-da-Índia) o CCL16 é um pseudogene, enquanto em membros das famílias Sciuridae (esquilo de treze linhas e a marmota), Cricetidae (hamster chinês e dourado), Dipodidae (gerbo), Bathyergidae (rato-toupeira-nu e o rato-toupeira-de-damaraland), Chinchilloidea (chinchila) e Octodontidae (degu), o CCL16 está intacto. Isso pode indicar que, embora o CCL16 esteja presente e seja funcional no ancestral dos Glires, este foi posteriormente inativado em algumas espécies. Isto pode ter ocorrido estocasticamente ou em linhagens específicas em momentos diferentes durante a evolução do CCL16.

Estes trabalhos são os primeiros estudos centrados na caracterização genómica e na diversidade genética de genes do sistema imunitário inato (interleucinas e ligandos de quimiocinas) na ordem Lagomorfa. Além disso, os resultados obtidos levantam novas questões que poderão conduzir a novos estudos para a compreensão das funções destas proteínas. Assim, os trabalhos futuros deverão ser direcionados para o estudo da função e determinação da estrutura cristalográfica, de forma a compreender o impacto das mutações detetadas nestas proteínas.

Palavras-chave

Lagomorfos

Coelho-bravo

Interação patogénio-hospedeiro

Imunidade inata

Quimiocinas

Interleucinas

Seleção positiva

SUMMARY

Cytokines are innate immune modulators involved in several processes of the immune response that diversify and quickly adapt to alterations in normal conditions. Knowledge on interleukins and chemokine ligands in lagomorphs was mostly restricted to the European rabbit. To fill in this gap we characterize the genomic organization, genetic diversity and evolution of interleukins and chemokine ligands with biological relevance in lagomorphs. Indeed, we analyzed these genes for several Lagomorphs belonging to the Ochotonidae and Leporidae families.

In order to understand the changes associated with the evolutionary patterns of these genes in lagomorphs, the sequences obtained in this work were compared with sequences available for other mammals. The search of signatures of positive selection by using maximum-likelihood (ML) methods in 46 mammalian ILs revealed that 60% of these makers have codons under positive selection. The number and percentage of codons positively selected varied between the ILs. Given the relevance of these proteins in the early activation of the immune response, we expected that their structure and function would be maintained along evolutionary lineages and thus, signatures of purifying selection would be detected. However, our results contradict this prediction, which may be explained by the multitude of biological processes in which ILs are enrolled.

From the ILs studied in this work, IL6 showed a unique evolutionary pattern. The IL6 from the European rabbit had been previously shown to differ from other mammals, including two other leporids (*Sylvilagus* spp. and *Lepus* spp.) that diverged ~12 million years ago, by extending for 27 additional amino acids. We further studied this protein for more four leporids (pygmy rabbit, riverine rabbit, Amami rabbit and volcano rabbit) and also in the American pika. Remarkably, in the Amami rabbit we detected a deletion of the stop codon causing an extension of IL6 for 17 extra amino acids. Our results indicate that this extension occurred independently in these species. Indeed, one occurred between 2 and 8 million years ago in the ancestor of the European rabbit, and the other occurred in the ancestor of the Amami rabbit less than 9 million years ago. The absence of this extension in *Bunolagus*, a sister genus of *Oryctolagus*, indicates evolution by convergence, suggesting some functional relevance. This is reinforced by the

presence of four extra cysteines in the European rabbit extension that, according to our modelling predictions, significantly alter the IL6 structure.

The European rabbit populations have been affected by two viral diseases, the rabbit hemorrhagic disease caused by the rabbit hemorrhagic disease virus (RHDV), and myxomatosis caused by the myxoma virus (MYXV). Previous studies on RHDV-infected rabbits identified alterations in the expression of IL1, IL2, IL6, IL8 and IL10. IL4, IL12, IL15 and IL18 expression was also altered following myxoma virus infection. We characterized these ILs in six Lagomorpha species (European rabbit, pygmy rabbit, two cottontail rabbit species, European brown hare and American pika). Overall, these ILs are conserved between lagomorphs. However, for the American pika, an altered location of the stop codon in IL1α and IL2, a different transcript in IL8 and a high number of cysteine residues in IL1β were observed. Additionally, changes at N-glycosylation motifs were also detected in IL1, IL10, IL12B and IL15.

The European rabbit and *Lepus* spp. are affected by a tularemia, a disease whose etiological agent is the bacteria *Francisella tularensis*. IL17A had been implicated in the host defense against this pathogen. The study of IL17A in five lagomorphs showed a similarity between 97–99% in leporids and ~88% between leporids and American pika. Overall, the IL17A structure is very well conserved. However, at codon 88, one of the interaction sites between IL17A and its receptor IL17RA, there is an Arg>Pro mutation that only occurs in the European rabbit and in the European brown hare. Interestingly, these two species are infected by *F. tularensis*. This mutation may induce critical alterations in the IL17A structure and conformation and consequently modify its function.

A CCR5-CCR2 gene conversion was observed in the European rabbit, Amami rabbit and riverine rabbit, but it was not detected in cottontails or *Lepus* spp., suggesting that this event occurred at approximately 9 million years ago in the ancestor of the European, Amami and riverine rabbits. Further studies on CCR5 ligands demonstrated that CCL3, CCL4 and CCL5 are functional in leporids and evolved under purifying selection. Interestingly, the evolution of CCL8 in lagomorphs is consistent with the gene conversion event observed. Indeed, CCL8 is pseudogenized in species with the CCR5-CCR2 gene conversion while it is functional in species without this gene conversion. This prompted us to study the remaining CCR5 ligands, CCL11, CCL14 and CCL16. All lagomorphs have a

potentially functional CCL11. However, the pygmy rabbit has a mutation in the stop codon leading to a longer protein. By using maximum likelihood approaches, we detected both signatures of purifying and positive selection. While purifying selection is detected in sites important for receptor binding and activation and may result from the proteins' functional constraints, positive selection might result from an increase in frequency of advantageous mutations that probably enhance the host response against several agents.

CCL14 seems to be functional in leporids while in some *Ochotona* species it is a pseudogene. Indeed, the CCL14 of American, alpine, palla's and turuchan pikas is a pseudogene due to a Met>Thr mutation at the start codon, while Northern pika presents one functional and one non-functional allelle and steppe pika is pseudogenized due to an insertion of 7 base pairs. Analysis of the CCL14 sequences of several mammals reveals six codons under positive selection in regions crucial for CCL14 activation and binding to receptors. This suggests that CCL14 has an important biological role in mammals that was lost in the Ochotonidae family, namely in *Pika* and *Lagotona* subgenera.

CCL16 gene was previously identified as a pseudogene in the European rabbit and in some rodents, such as mouse, rat and guinea pig, while in the squirrel it seems to be functional. To elucidate the evolution of this gene in the superorder Glires (Rodents and Lagomorphs), we sequenced six leporids and nine species of *Ochotona*. Additionally we retrieved all the CCL16 sequences of rodents available in public databases. Our results suggest that in this superorder CCL16 suffered several independent pseudogenization events. In the order Lagomorpha, all Leporid species, with the exception of cottontails (*Sylvilagus* spp.), showed that pseudogenization is due to a non-synonymous mutation at codon 45 that leads to a premature stop codon. This suggests that the gene became pseudogenized in the Leporids ancestor at around 14 million years ago. Interestingly, the cottontail rabbits (Mexican, forest and Eastern cottontail) do not present this mutation. In cottontails, the Cys45 from the typical Cys-Cys motif also suffered a mutation and encodes a Lysine. Despite encoding a putatively functional protein, CCL16 is not expressed in the Eastern cottontail. The exchange of one of the CCL16 characteristic Cysteines for a Lysine may have caused the loss of the protein function. In all species of Pika, except for one allele of Hoffmann's pika, CCL16 is intact, and is expressed in the American Pika. In

Rodentia, some members of the Muridae (mouse and rat), Heteromyidae (kangoroo rat) and Cavioidea (guinea pig) have a pseudogenized CCL16 gene while in members of the Sciuroidea (thirteen lined ground squirrel and alpine marmot), Cricetidae (Chinese and golden hamsters), Dipodidae (lesser Egyptian jerboa), Bathyergidae (naked mole-rat and damaraland mole-rat), Chinchilloidea (long tailed chinchilla), and Octodontoidea (degu), CCL16 is intact. This may indicate that although CCL16 is present and functional in the ancestor of the Glires clade, it was later inactivated in some species. This may have occurred stochastically at different moments in the CCL16 evolution.

This work is the first study on the genomic characterization and genetic diversity of innate immune system genes (Interleukins and Chemokine ligands) in the order Lagomorpha. The results obtained raise new questions and may lead to a set of future studies that could help to fully understand these proteins' functions. Thus, future works should rely on functional and crystallographic studies to fully understand the impact of the mutations detected in these proteins.

Keywords:
Lagomorphs
European rabbit
Host-pathogen interaction
Innate immunity
Chemokines
Interleukins
Positive selection

CHAPTER 1

General Introduction

1. THE LAGOMORPHA ORDER

1.1. Phylogeny

Lagomorphs were first classified as rodents (Duplicidentata) (Vaughan et al., 2011). The order Lagomorpha was only proposed in 1912 by Gidley (Gidley, 1912) in order to distinguish lagomorphs from rodents. Nowadays, and despite some controversy, lagomorphs and rodents are clustered in the Glires clade (Douzery and Huchon, 2004). The order Lagomorpha includes rabbits, hares and pikas, which are adapted to a wide range of environments, from deserts to the artic (Chapman and Flux, 2008). Comprising a total of 91 living species, the order Lagomorpha has two families, Ochotonidae and Leporidae that diverged approximately 30 million years ago (Fontanesi et al., 2016; Lopez-Martinez, 2008; Matthee et al., 2004). With a common origin in Asia and the first fossils dating from the early Eocene, these families became differentiated in their routes and sizes, probably due to climatic changes (Ge et al., 2013; Lopez-Martinez, 2008).

1.1.1. The Ochotonidae family

The Ochotonidae family includes a single genus (Ochotona) divided in four subgenera (*Pika*, *Ochotona*, *Conothoa* and *Lagotona*) (Figure 1.1) (Lissovsky, 2014). The radiation of these subgenera occurred at approximately 7 million years ago (Melo-Ferreira et al., 2015). The first records of pikas date back to the Eocene, in Asia, spreading to Europe and North America during the Oligocene (Lopez-Martinez, 2008). Pikas are small-sized (70-300 grams) with short round ears and concealed tail and tend to vocalize (Chapman et al., 1990). Currently, there are 30 living species, well-adapted to extreme climates, being found mostly in alpine or meadow environments, such as in high mountains of North America and Asia (Chapman and Flux, 2008; Fontanesi et al., 2016; Fostowicz-Frelik et al., 2010; Lissovsky, 2014; Melo-Ferreira et al., 2015; Smith, 2008; Vaughan et al., 2011). In the past few years, many pika subspecies have been considered threatened and some have already became extinct (Fontanesi et al., 2016; Smith, 2008).

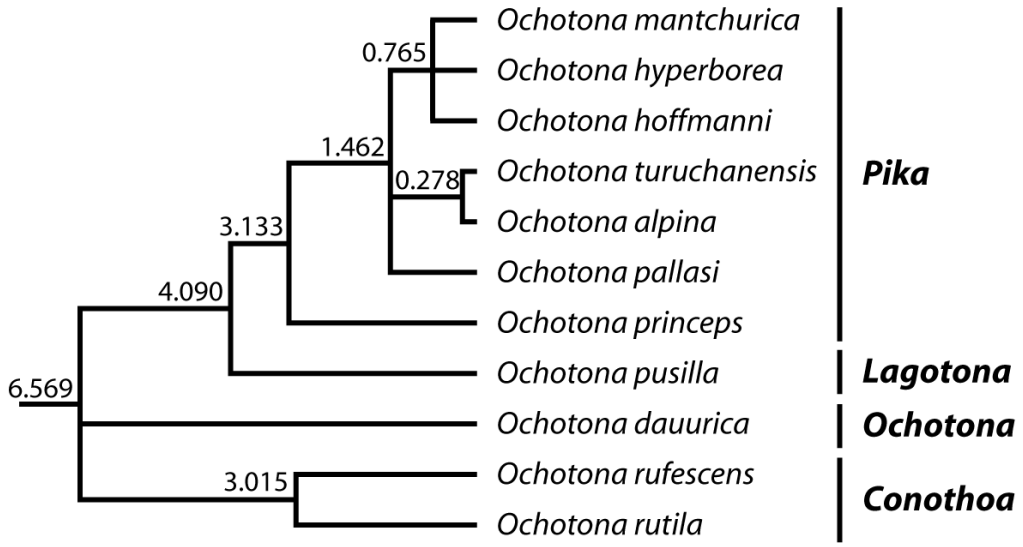


Figure 1. 1. Evolutionary relationships within the Ochotonidae family. The divergence time (in million years) is given for each node (phylogeny based on data from Melo-Ferreira et al. 2015).

1.1.2. The Leporidae family

The Leporidae family includes rabbits and hares in a total of 11 genera (*Poelagus*, *Pronolagus*, *Nesolagus*, *Oryctolagus*, *Bunolagus*, *Caprolagus*, *Pentalagus*, *Brachylagus*, *Sylvilagus*, *Lepus* and *Romerolagus*) with 61 extant species (29 species of rabbits and 32 species of hares) widely distributed (Figure 1.2). From these, seven genera are monotypic: *Poelagus* (Bunyoro rabbit), *Oryctolagus* (European rabbit), *Bunolagus* (riverine rabbit), *Caprolagus* (hispid hare), *Pentalagus* (Amami rabbit), *Brachylagus* (pygmy rabbit) and *Romerolagus* (volcano rabbit) (Alves and Hacklander, 2008; Fontanesi et al., 2016). Despite some species being referred as hares, e.g. the hispid hare, only one genus represents the true hares (*Lepus*), with the remaining genera corresponding to true rabbits. Morphologically, leporids present larger body size (1-5 kilograms), with long, large and tubular ears, long tails and long hind legs (Chapman and Flux, 2008; Fontanesi et al., 2016; Vaughan et al., 2011).

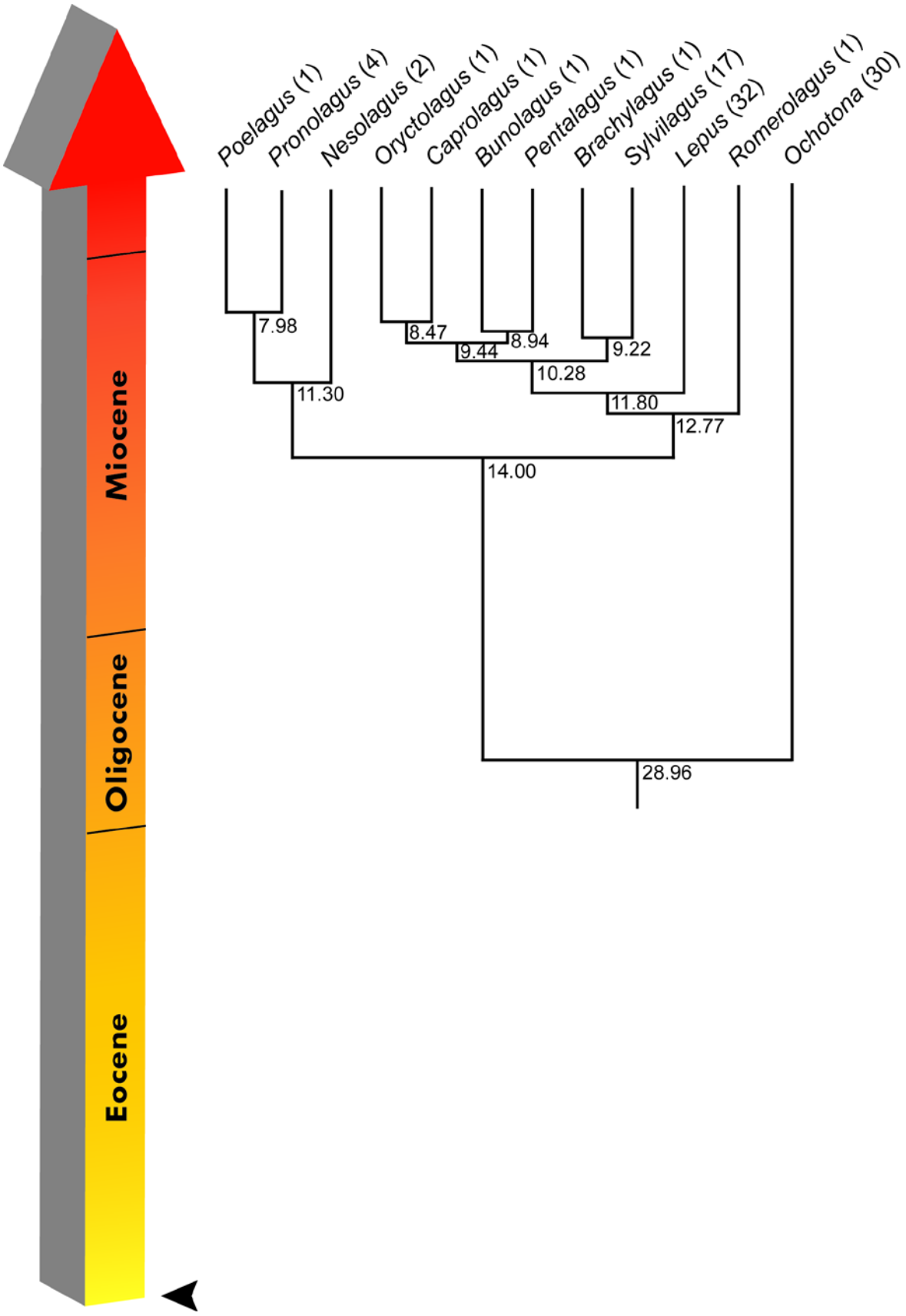


Figure 1. 2. Evolutionary relationships in the Leporidae family. The divergence time (in million years) is given for each node (phylogeny based on data from Matthee et al. 2004). The number of species per genus is indicated in brackets.

2. IMMUNE SYSTEM

The immune system uses a broad array of mechanisms in order to recognize specific features of foreign organisms that distinguish them from the host and to consequently develop an immune response towards their elimination (Chaplin, 2010). Innate immunity is the first line of defense and is present in all multicellular organisms, while adaptive immunity is characteristic of jawed vertebrates. Lagomorphs are affected by several diseases that have distinct repercussions according to the species/genera. This is probably due to specific differences in genes of the immune system, both innate and adaptive.

2.1. Innate Immunity

This first line of defense prevents pathogen entry through several physical (skin and mucosal surfaces), chemical (pH and soluble factors) and microbiological (phagocytic cells) barriers (Figure 1.3). Encoded by genes of the host's germline, innate immunity is able to distinguish between "self" and "non-self" antigens (Chaplin, 2010; Thaiss et al., 2016; Turvey and Broide, 2010). When a pathogen invades the innate immune system acts immediately through pattern recognition receptors (PRRs) that detect conserved pathogen-associated molecular patterns (PAMPs) such as pathogen specific lipoproteins and carbohydrates. The host is also capable to detect damage-associated molecular patterns (DAMPs) released from broken cells, and quickly eliminate them. PRRs include the Toll-like receptors (TLRs), peptidoglycan recognition proteins (PGRPs) and NOD-like receptors (NLRs), among others (Deng et al., 2013; Thaiss et al., 2016). The outcome from the activation of these PRRs is an inflammatory response through the activation of innate immune cells such as macrophages and dendritic cells that produce several types of cytokines and chemokines, crucial to regulate the nature of the immune responses (Deng et al., 2013; Kasamatsu, 2013; Newton and Dixit, 2012; Thaiss et al., 2016). Despite being originally defined as unable to develop immunological memory, recent findings showed that innate immunity is able to exhibit "trained immunity", i.e., cells of the innate immune system like monocytes and natural killer (NK) cells are able to recognize pathogens and protect the host against re-infections. Indeed, this is observed in

lower organisms as plants and invertebrates (Netea et al., 2016). In vertebrates, this “trained immunity” acts in line with the adaptive immunity (Netea et al., 2016; Netea et al., 2011; Sun et al., 2014; van der Meer et al., 2015).

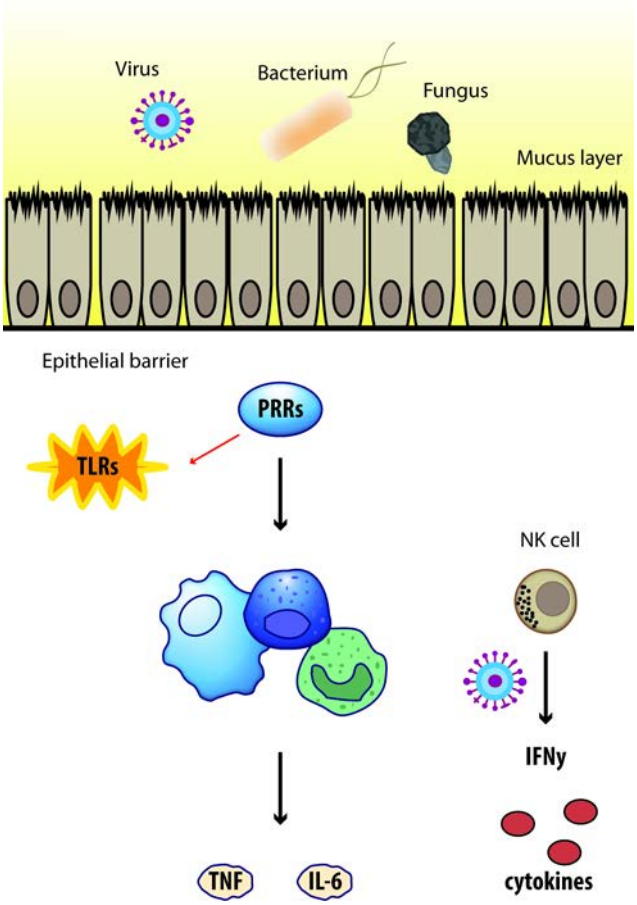


Figure 1. 3. Schematic representation of the innate imune response. Adapted from (Fontes et al., 2015).

2.1.1. Cytokines

Cytokines are small cell-signaling glycoproteins that include a group of chemotactic cytokines also known as chemokines. Cytokines and chemokines are secreted by a variety of cells in response to different stimuli and can act in other cells affecting their activities in order to define and control the nature of the immune responses. These proteins are also important to regulate immune cell trafficking and the cellular arrangement of the immune organs (Akdis et al., 2011;

Borish and Steinke, 2003; Commins et al., 2010). When these cytokines are secreted by leukocytes and act on other leukocytes they are named interleukins (ILs) (Beadling and Slifka, 2006; Kaiser et al., 2004; Schrader, 2003).

Interleukins are proteins of low molecular weight that participate in several different biological activities, such as immunity, inflammation, hematopoiesis, oncogenesis, neurogenesis, fertility, etc (Afzal et al., 2012; Heinrich et al., 2003; Ishihara and Hirano, 2002). Despite being molecules of the innate immune system, they may influence the result and nature of the adaptive immune response, being crucial for growth, differentiation and survival of immune cells (Brocker et al., 2010; Kaiser et al., 2004; O'Connell and McInerney, 2005; Zelus et al., 2000; Zhang and Nei, 2000). In order to exert their functions, interlukins bind to specific cell-surface receptors that are able to transmit an intracellular signal through different signaling pathways (Afzal et al., 2012; Fumagalli et al., 2009; Ishihara and Hirano, 2002; O'Connell and McInerney, 2005). Currently, there are 38 interleukins known in the human genome, which are classified according to their numeric order of discovery (IL1-IL38). In other mammals this number varies between species.

IL1 is mainly produced by mononuclear phagocytes and is a strong proinflammatory cytokine and an endogenous pyrogen. It includes two proteins (IL1α and IL1β) with similar biological activities. With an essential role in T helper 17 (Th17) cells differentiation, it is also a major mediator of inflammation, contributing to an increase of B cells proliferation and immunoglobulin (Ig) synthesis (Garlanda et al., 2013; Sims and Smith, 2010). IL1 may also be important for viral clearance through interferon alpha (IFNα) production (Zhu and Liu, 2003).

IL2, also known as T cell growth factor, is produced by CD4 T cells and has a central role in T cell dependent immune responses. This pleiotropic cytokine is a potent pyrogen with a key role in CD4 and CD8 differentiation and T regulatory (Treg) cells maintenance (Boyman and Sprent, 2012; Liao et al., 2013; Malek, 2008).

IL3 is a Th2 cytokine mainly produced by T lymphocytes, essential for eosinophil development and survival, being associated with allergic inflammation (Borish and Steinke, 2003; Commins et al., 2010; Tomaki et al., 2002).

IL4 is also a Th2 cytokine secreted by several cells such as Th2, NK, mast cells, eosinophils and basophils, being important for the immune response against parasitic worms. This anti-inflammatory cytokine is also able to inhibit proinflammatory cytokine production, being associated with allergy, asthma and autoimmunity (Koyanagi et al., 2010; Luzina et al., 2012; Paul, 2015; Pillai and Bix, 2011).

IL5 is a Th2 cytokine derived from T lymphocytes with major roles in allergic inflammation through eosinophil production, maturation, recruitment, differentiation, survival and activation (Borish and Steinke, 2003; Levine and Wenzel, 2010; Patterson et al., 2015).

IL6, also known as B cell differentiation factor is produced after stimulation of several different cells, mainly mononuclear phagocytes. With a large array of biological activities in humoral and cellular defense, it is responsible for the initiation of acute-phase response in hepatocytes, favoring Th1, Th2 and Th17 differentiation and suppressing Treg responses (Neurath and Finotto, 2011; Nishimoto and Kishimoto, 2006; Scheller et al., 2011).

IL7, also named lymphopoietin-1, is important for B lymphocyte maturation in bone marrow and for proliferation, differentiation and survival of B and T lymphocytes (Akdis et al., 2011; Fry and Mackall, 2002, 2005).

IL8, also called C-X-C chemokine ligand 8 (CXCL8), is a potent chemoattractant of neutrophils and T cells being associated with several chronic inflammatory conditions (Akdis et al., 2011; Harada et al., 1994; Qazi et al., 2011).

IL9 is produced by T cells and is an important stimulator of mast cells. This Th9 cytokine is essential for immunity and inflammation development and acts as both positive and negative regulator of immune responses, with unfavorable functions in allergy and autoimmunity and beneficial roles in parasitic infections (Goswami and Kaplan, 2011; Kaplan et al., 2015; Noelle and Nowak, 2010).

IL10 is a central player in the inflammatory response, protecting the host from excessive immune responses to microbial infections and in autoimmune diseases. This anti-inflammatory protein is able to inhibit the production and expression of several important cytokines, chemokines and receptors. IL10 has been shown to be important for the control of inflammatory reactions in viral infections (Couper et al., 2008; Sabat et al., 2010; Saraiva and O'Garra, 2010).

IL11, initially described as a stimulator factor for hematopoietic cells, plays a significant regulatory role in immunoglobulin and acute phase proteins production and also in lymphoid cell differentiation (Akdis et al., 2011; Curfs et al., 1997; Du and Williams, 1997).

IL12 and IL23 are heterodimers mostly produced by dendritic cells (DC) that share a common subunit (IL12p40). Both immunoregulatory proteins are important in T cell activation and influence the Th1 and Th17 responses, respectively (Akdis et al., 2011; Langrish et al., 2004; Teng et al., 2015; Watford et al., 2004). Along with IL18, IL12 induces lFNγ, and both present synergistic effects in the development of Th1 cells and antiviral activity. Furthermore, IL18 is important in the innate response against pathogens, establishing a link between the innate and the adaptive immune systems (Arend et al., 2008; Gasteiger and Rudensky, 2014; Hamza et al., 2010).

IL13 is mainly produced by Th2 CD4 cells and shares with IL4 all its functions and 25% of homology. In addition, IL13 is important for mucus secretion (Bao and Reinhardt, 2015; Gordon, 2003; Wills-Karp, 2004; Zhu, 2015).

Also known as alpha taxilin, IL14 is derived from normal T and B cells, inhibits Ig secretion of activated B cells and controls growth and proliferation of normal and abnormal B cells (Akdis et al., 2011; Shen et al., 2006).

IL15 is a T cell growth factor with important roles in defense against microorganisms and tumors. Produced by mononuclear phagocytes, this proinflammatory protein is crucial for the development, survival and differentiation of, for example, CD8 memory T cells, NK cells and Th2 cells (Akdis et al., 2011; Commins et al., 2010; Fehniger and Caligiuri, 2001; Stonier and Schluns, 2010).

IL16, also known as lymphocyte chemoattractant factor, is a proinflammatory cytokine secreted as a precursor (pro-IL16) that needs cleavage to acquire its mature form. Both forms are biological active, but with different functions. While pro-IL16 is a regulator of the cell cycle progression, the mature form is a chemoattractant and growth/differentiation factor for several hematopoietic cells (Commins et al., 2010; Cruikshank and Little, 2008; Richmond et al., 2014; Wilson et al., 2004). Studies on human IL16 revealed that its expression is associated with apoptosis, autoimmune and allergic diseases (Bowler et al., 2013; Elssner et al., 2004).

Mainly produced by Th17 cells, the IL17 family comprises six different structurally related proteins, IL17A-F, with IL17E renamed IL25 due to its specific functions. With distinct cellular sources, they are important regulators in host defense and inflammatory diseases with the ability to induce proinflammatory cytokines. IL17, also known as IL17A, is associated with progression of allergic and autoimmune diseases and with defense against external pathogens such as bacteria and fungi (Iwakura et al., 2011; Jin and Dong, 2013; Onishi and Gaffen, 2010; Witowski et al., 2004).

IL19 can be produced by both tissue and immune cells, being associated with antimicrobial immunity and tissue repair (Akdis et al., 2011; Fielding, 2012; Sabat, 2010).

IL20 shares with IL19 many of its sources and functions. Furthermore, it regulates keratinocyte functions, and contributes to the formation of lymphatic vessels (Akdis et al., 2011; Rutz et al., 2014).

IL21 is produced by CD4 cells and plays a role in Ig production, in T and NK cells stimulation, being also important for B cells proliferation, survival and differentiation. In addition, it has a crucial role on inflammatory responses associated with autoimmune and inflammatory diseases, assisting innate and adaptive responses regulation (Croce et al., 2015; Spolski and Leonard, 2008, 2014).

Produced by cells of the innate and adaptive immune systems, IL22 has both tissue protection and proinflammatory functions. Moreover, it is also essential for induction of acute-phase proteins by hepatocytes and protection against extracellular pathogens, participating in intestinal homeostasis and the regulation of commensal communities (Dudakov et al., 2015; Rutz et al., 2014).

IL24 is a proinflammatory cytokine produced by different immune cells and is a strong and specific inhibitor of cancer cell proliferation (Akdis et al., 2011; Sabat, 2010).

IL25, also known as IL17E, is a member of the IL17 family, associated with initiation, promotion and increase of Th2 cell-mediated immune response, mainly in helminthic infections and allergic inflammation (Iwakura et al., 2011; Jin and Dong, 2013; Monteleone et al., 2010).

IL26 is a Th17 cytokine important for elimination of extracellular bacteria and may also play a role in antiviral defense and inflammation (Donnelly et al., 2010; Tengvall et al., 2016).

IL27 is a heterodimeric cytokine composed by two subunits (IL27p28 – also known as IL30; and Epstein-Barr virus induced gene 3 - Ebi3) with important roles in development of NK and T cells and Th1 immunity. This immunosuppressive cytokine also promotes IFN-γ production and is an antagonist of different inflammatory cells with inhibitory effects on Th2 and Th17 responses (Hunter and Kastelein, 2012; Iwasaki et al., 2015; Yoshida and Hunter, 2015).

IL31 is predominantly produced by Th2 cells and is associated with proliferation of B and T cells. This protein is relevant for innate and adaptive immunity, being involved mostly in skin and allergic disorders (Castellani et al., 2010; Cornelissen et al., 2012; Zhang et al., 2008).

IL32, initially named NK cell transcript 4, is mainly found in NK and T cells. This proinflammatory protein has six variants (α, β, γ, δ, ε, ζ) and is important for the induction of other proinflammatory proteins. It has also been associated with infection, chronic inflammation and cancer (Dinarello and Kim, 2006; Heinhuis et al., 2012; Joosten et al., 2013; Kundu and Basu, 2006; Zhou and Zhu, 2015).

IL33 is an inducer of several chemokines and cytokines of the innate and adaptive immune systems and also of the Th2 response. Mostly expressed by cells with direct contact with the environment, it acts in the early inflammatory response, playing a role in cell activation, differentiation, polarization and chemotaxis (Martin and Martin, 2016; Miller, 2011; Palmer and Gabay, 2011).

IL34 increases monocyte viability and promotes macrophage colony formation from bone marrow cells, being expressed in a large number of tissues (Akdis et al., 2011; Commins et al., 2010; Masteller and Wong, 2014).

IL35 is a heterodimer composed by two subunits (IL12p35 and Ebi3) with immunosuppressive functions, mostly produced by Treg cells (Egwuagu et al., 2015; Olson et al., 2013; Sawant et al., 2015).

IL36 includes three agonistic cytokines: IL36α (IL1F6), IL36β (IL1F8) and IL36γ (IL1F9). Despite being expressed and activated by different cells, they share the same activities, being important for stimulation of both innate and adaptive immune responses (Gabay and Towne, 2015; Gresnigt and van de Veerdonk, 2013).

IL37, also known as IL1F7, is a suppressor of innate immune responses inhibiting the production of proinflammatory cytokines (Banchereau et al., 2012; Garlanda et al., 2013).

Firstly identified as IL1F10, IL38 is considered a negative regulator with activities in human inflammation and autoimmunity (Garlanda et al., 2013; Yuan et al., 2015).

Interferons (IFNs) are inducible cytokines that are able to protect cells from viral infections and may also regulate other processes of the immune response. There are three major classes of IFNs (type I, II and III). Type I IFN includes IFNα, β, ε, τ, κ, ω, δ and ζ, while type II and type III only include one protein, IFNγ and IFNλ, respectively. Type I IFNs are expressed in several different tissues with antiviral, antiproliferative and immunomodulatory effects. IFNγ is widely expressed, with important roles in regulation of the immune response, linking both innate and adaptive responses. Type III IFN includes IFNλ1 (IL29), IFNλ2 (IL28A), IFNλ3 (IL28B) and IFNλ4 and has antiviral activity (Ivashkiv and Donlin, 2014; McNab et al., 2015; Schneider et al., 2014).

Tumor necrosis factor (TNF) is a superfamily of cytokines with several functions in antimicrobial immunity. TNFα, the most well known member of this superfamily, is expressed by monocytes and T cells and is important for tumor growth control being a strong mediator of inflammation (Sedger and McDermott, 2014).

The European rabbit interleukins

From the proteins described above, only 32 are described for the European rabbit (Table 1.1). Some of these interleukins have been associated with rabbit immunity, namely in microbial infections and viral diseases such as rabbit hemorrhagic disease (RHD) and myxomatosis. The IL17 and IL22 are highly up-regulated in rabbit microbial infections (Schnupf and Sansonetti, 2012; Skyberg et al., 2013). In rabbit hemorrhagic disease virus (RHDV)-infected rabbits, different patterns of proinflammatory and anti-inflammatory proteins were observed at different time points post-infection and in different cells. The major differences were observed in IL1, IL2, IL6, IL8, IL10, TNFα and IFNα (Garcia-Lastra et al., 2010; Marques et al., 2012; Marques et al., 2014; Trzeciak-Ryczek et al., 2016; Tunon et al., 2011b).

Regarding myxoma virus infections, it has been shown that: IL4 enhances virus efficiency and virulence by inhibiting host viral clearance and decreasing genetic resistance (Kerr et al., 2004; Stanford and McFadden, 2005); IL12 may be useful as oncolytic virus, counteracting the myxoma virus infection in rabbits (Stanford et al., 2007); IL15 was able to stimulate the immune response leading to the elimination of the infection (Liu et al., 2009; Tosic et al., 2014); and IL18 was indirectly affected by poxviruses due to their ability to inhibit proteins important for IL18 activation (Johnston and McFadden, 2004; Vande Walle and Lamkanfi, 2011).

Furthermore, IL2 was associated with rabbit inflammatory processes (Perkins et al., 2000) while IL16 down-regulation leads to T cell activation and proliferation (Jacquier et al., 2015). Two transcripts have been described for rabbit IL7, one presents similar functions to human IL7, while the other has an opposite function being able to inhibit B lymphopoiesis (Siewe et al., 2010).

The list of functions described above is not exhaustive, as these proteins may have other roles. Indeed, they can have different functions depending on several factors such as target cells, cells where they are expressed, site of infection, pathogen, disease, etc. Moreover, many of these cytokines present complementary and conflicting roles in induction, regulation and function of the immune system and they can interact between themselves in order to obtain an optimal immune response.

2.1.2. Chemokines

Chemokines are cytokines capable to induce chemotaxis. Produced upon stimulation, these proteins only exert their roles by binding to specific transmembrane G protein-coupled receptors or extracellular matrix-associated glycosaminoglycans (Bryant and Slade, 2015). Biological roles associated with these proteins include coordination of leucocyte recruitment and trafficking, mediation of homeostatic migration and homing of several cells, mediation in vascular injury, antimicrobial activities, among others (Zlotnik and Yoshie, 2012).

Table 1. 1. Chromossomic location of the interleukins described for the European rabbit. The information is available on https://www.ncbi.nlm.nih.gov/gene and http://www.ensembl.org/index.html databases.

IL		Chromosomic location	IL		Chromosomic location
IL1α	Human	Chr 2: 112,773,915-112,784,590 reverse strand	**IL18**	Human	Chr 11: 112,143,251-112,164,117 reverse strand
	European rabbit	Chr 2: 97,559,306-97,569,595 reverse strand		European rabbit	Chr 1: 104,215,287-104-234,949 forward strand
IL1β	Human	Chr 2: 112,829,751-112,836,903 reverse strand	**IL19**	Human	Chr 1: 206,798,870-206,842,981 forward strand
	European rabbit	Chr 2: 97,614,851-97,618,656 reverse strand		European rabbit	Chr 16: 65,519,827-65,525,648 reverse strand.
IL2	Human	Chr 4: 122,451,470-122,456,725 reverse strand	**IL20**	Human	Chr 1: 206,865,354-206,869,223 forward strand.
	European rabbit	Chr 15: 98,312,429-98,317,579 reverse strand		European rabbit	Chr 16: 65,501,768-65,504,135 reverse strand.
IL4	Human	Chr 5: 132,673,986-132,682,676 forward strand	**IL21**	Human	Chr 4: 122,612,628-122,621,069 reverse strand
	European rabbit	Chr 3: 15,702,738-15,711,514 forward strand		European rabbit	Chr 15: 98,465,680-98,472,654 reverse strand.
IL5	Human	Chr 5: 132,541,444-132,556,838 reverse strand	**IL22**	Human	Chr 12: 68,248,242-68,253,607 reverse strand
	European rabbit	Chr 3: 15,573,697-15,575,837 reverse strand		European rabbit	Chr 4: 47,237,605-47,243,806 reverse strand
IL6	Human	Chr 7: 22,725,884-22,732,002 forward strand	**IL23A**	Human	Chr 12: 56,334,174-56,340,410 forward strand
	European rabbit	Chr 10: 77,715,96-77,764,08 forward strand		European rabbit	Chr 4: 39,913,661-39,916,220 forward strand
IL7	Human	Chr 8: 78,675,743-78,805,523 reverse strand	**IL25 (IL17E)**	Human	Chr 14: 23,372,809-23,376,403 forward strand
	European rabbit	Chr 3: 94,496,942-94,538,207 reverse strand		European rabbit	Chr 17: 43,569,196-43,571,703 forward strand
IL8 (CXCL8)	Human	Chr 4: 73,740,506-73,743,716 forward strand	**IL26**	Human	Chr 12: 68,201,351-68,225,821 reverse strand
	European rabbit	Chr 15: 76,368,976-76,371,927 reverse strand		European rabbit	Chr 4: 47,203,642-47,216,977 reverse strand
IL9	Human	Chr 5: 135,892,246-135,895,827 reverse strand	**IL27p28 (IL30)**	Human	Chr 16: 28,499,362-28,512,051 reverse strand
	European rabbit	Chr 3: 18,397,431-18,557,341 reverse strand		European rabbit	Chr 6: 18,601,896-18,605,035 forward strand
IL10	Human	Chr 1: 206,767,602-206,772,494 reverse strand	**IL28A (IFNλ2)**	Human	Chr 19: 39,268,514-39,270,092 forward strand
	European rabbit	Chr 16: 65,583,126-65,587,913 forward strand		European rabbit	Chr 5:680,445-683,763 reverse strand
IL12A	Human	Chr 3: 159,988,750-159,996,019 forward strand	**IL29 (IFNλ1)**	Human	Chr 19: 39,296,325-39,298,673 forward strand
	European rabbit	Chr 14: 54,457,643-54,464,672 forward strand		European rabbit	Chr 5: 657,858-658,987 reverse strand.
IL12B	Human	Chr 5: 159,314,783-159,330,887 reverse strand	**IL31**	Human	Chr 12: 122,172,030-122,174,199 reverse strand
	European rabbit	Chr 3: 40,710,276-40,721,669 reverse strand		European rabbit	Scaffold GL018824: 521,163-522,648 forward strand
IL13	Human	Chr 5: 132,656,263-132,661,110 forward strand	**IL32**	Human	Chr 16: 3,065,297-3,082,192 forward strand
	European rabbit	Chr 3: 15,690,388-15,692,118 forward strand		European rabbit	Scaffold NW_003159454.1: 3,003,26-3,031,27 forward strand
IL15	Human	Chr 4: 141,636,599-141,733,987 forward strand	**IL33**	Human	Chr 9: 6,215,786-6,257,983 forward strand
	European rabbit	Chr 15: 22,774,049-22,789,507 reverse strand		European rabbit	Chr 1: 47,919,299-47,960,442 reverse strand
IL16	Human	Chr 15: 81,159,575-81,314,058 forward strand	**IL34**	Human	Chr 16: 70,579,895-70,660,682 forward strand
	European rabbit	Scaffold GL018737: 1,518,058-1,683,475 reverse strand		European rabbit	Chr 5: 28,856,634-28,865,667 reverse strand
IL17A	Human	Chr 6: 52,186,387-52,190,638 forward strand	**IL36α (IL1F6)**	Human	Chr 2: 113,005,461-113,008,044 forward strand
	European rabbit	Chr 12: 41,692,982-41,696,102 forward strand		European rabbit	Chr 2: 97,773,226-97,777,970 forward strand
IL17B	Human	Chr 5: 149,371,324-149,404,202 reverse strand	**IL36γ (IL1F9)**	Human	Chr 2: 112,973,203-112,985,665 forward strand
	European rabbit	Chr 3: 31,217,355-31,221,924 reverse strand		European rabbit	Chr 2: 97,749,486-97,756,304 forward strand
IL17D (IL27)	Human	Chr 13: 20,702,127-20,723,098 forward strand	**IL37 (IL1F7)**	Human	Chr 2: 112,912,971-112,918,882 forward strand
	European rabbit	Chr 8: 43,865,579-43,885,936 reverse strand		European rabbit	Chr 2: 97,681,836-97,693,940 forward strand
IL17F	Human	Chr 6: 52,236,681-52,244,537 reverse strand	**IL38 (IL1F10)**	Human	Chr 2: 113,067,970-113,075,850 forward strand
	European rabbit	Chr 12: 41,737,327-41,744,062 reverse strand		European rabbit	Chr 2: 97,823,494-97,826,738 forward strand.

Currently, there are approximately 46 chemokines described that according to their N-terminus cysteine residues are classified into CC, CXC, CX3C, XC and CX (Bryant and Slade, 2015; Nomiyama et al., 2013). C-C chemokines present two adjacent cysteines, include 28 ligands and the encoding genes are located in different regions of genome. Indeed, C-C chemokine ligand 1 (CCL1), CCL2, CCL7, CCL8 and CCL11-CCL13 are located in the monocyte chemoattractant protein (MCP) region, while CCL3-CCL6, CCL9/CCL10, CCL14-CCL17 and CCL23 are located in the macrophage inflammatory protein (MIP) region (Figure 1.4). The remaining ligands (CCL18-CCL22 and CCL24-CCL28) are scattered throughout the genome (Nomiyama et al., 2013; Shibata et al., 2013; Zlotnik and Yoshie, 2012). CXC ligands have one amino acid between the cysteine residues and are encoded by 17 genes clustered in different regions (Nomiyama et al., 2013; Shibata et al., 2013; Zlotnik and Yoshie, 2012). CXCL1-CXCL8 are located in the growth related oncogene (GRO) region while CXCL9-CXCL11 and CXCL13 are found in the IFNγ-inducible protein 10 (IP10) region (Nomiyama et al., 2013; Shibata et al., 2013; Zlotnik and Yoshie, 2012). The CX3C family presents three residues between the first and the second cysteines and is composed by only one ligand (CX3CL1), also known as fractalkine (Nomiyama et al., 2013; Shibata et al., 2013; Zlotnik and Yoshie, 2012). In turn, in the XC subfamily the first and third cysteines are absent and it is composed by two ligands (XCL1 and XCL2) (Nomiyama et al., 2013; Shibata et al., 2013; Zlotnik and Yoshie, 2012). The last group of chemokines (CX) was recently described in zebrafish and includes only one ligand (CXL1) (Nomiyama et al., 2008). The number of chemokines varies between species: mouse – 46; human – 44 and rabbit – 35 (Shibata et al., 2013).

Chemokines belong to a multigene family with high diversity and promiscuity (Nomiyama et al., 2013). Besides their important roles in immunity and pathogen entry, the study of chemokines gained notoriety when chemokine receptors CCR5 and CXCR4 were described as co-receptors for human immunodeficiency virus (HIV) (Wilen et al., 2012). The interaction between these receptors and their ligands prevents HIV-1 entry into host cells (Munch et al., 2014). As for other mammals, in some leporid species CCR5 presents a gene conversion event with CCR2, at the second extracellular loop (Abrantes et al., 2011; Carmo et al., 2006; Esteves et al., 2007; Perelygin et al., 2008; Vazquez-Salat et al., 2007). Indeed, this gene conversion is present in the European,

riverine and Amami rabbits but not in the European brown hare (*Lepus europaeus*) or in the eastern cottontail (*Sylvilagus floridanus*) (Abrantes et al., 2011; Carmo et al., 2006), suggesting that the CCR5/CCR2 gene conversion occurred in the ancestor of the European, riverine and Amami rabbits, approximately 9 million years ago. CCR5 ligands include CCL3, CCL4, CCL5, CCL8, CCL11, CCL14 and CCL16 (Zlotnik and Yoshie, 2012). CCL3, CCL4 and CCL5 are functional genes in lagomorphs and present sites with signatures of purifying selection in regions important for receptor binding (de Matos et al., 2014). In mouse, rat and rabbit there is only one copy of CCL3 and CCL4, but in other rodents such as squirrel and guinea pig, as well as in humans, there are more (functional or inactivated) copies of these genes (Shibata et al., 2013; Zlotnik et al., 2006). In turn, CCL8 is pseudogenized in the European, Amami and riverine rabbits, whilst intact in hares and eastern cottontail (van der Loo et al., 2012; van der Loo et al., 2016). CXCR4 has only one ligand (CXCL2) (Zlotnik and Yoshie, 2012) and presents low genetic diversity. Indeed, only two amino acid mutations had been described in the European rabbit (Abrantes et al., 2008).

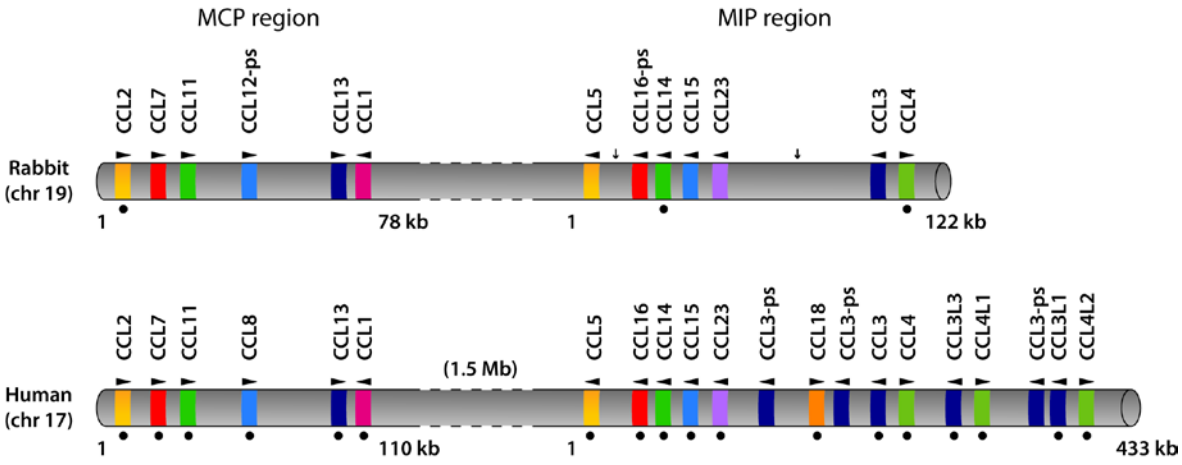


Figure 1. 4. Genomic organization of the C-C chemokine major cluster. The maps shown are adapted from Shibata et al, 2013.

2.2. Host-pathogen co-evolution

The main function of the immune system is to ensure host integrity in face of invading pathogens. Although there are several types of invaders, they all tend

to diminish host fitness (Muraille, 2013). The pathogen depends on the host to survive and consequently the host recognizes the pathogen antigens and develops an immune response to eliminate it. This leads to the development of new mechanisms by the pathogen in order to maintain this interaction with the host and to promote its own survival. In addition, advantageous mutations present in the host and pathogen populations that make them more able to survive and breed are likely to be selected and tend to be fixed in the populations. From this interaction there are three possible outcomes: i) the pathogen overcomes the host immune system leading to disease or death; ii) the host is able to develop new strategies and eliminate the pathogen; and iii) a cyclical process of adaptation and counter-adaptation that benefits some traits through evolutionary selective pressures, allowing the equilibrium between the host and the pathogen (Barreiro and Quintana-Murci, 2010). In 1973, Van Valen proposed the “Red Queen hypothesis” to describe this balance, stating that the species need to evolve continuously to keep up with the competition (Van Valen L., 1973). A good example of this host-pathogen co-evolution in natural conditions is the system myxoma virus and the European rabbit in Australia (Best and Kerr, 2000; Fenner, 1956).

The mammalian immune system genes are widely studied and are among the genes with faster evolution. Indeed, some of these genes show signatures of strong positive selection probably due to a constant need to adapt in this co-evolution process (Arbiza et al., 2006; Areal et al., 2011; Kosiol et al., 2008; O'Connell and McInerney, 2005; Vallender and Lahn, 2004; Zhang and Nei, 2000). At the sequence level, this might result in an increase of non-synonymous substitutions (d_N) over synonymous substitutions (d_S). Non-synonymous substitutions occur mainly at the first and second positions of the codon and lead to amino acid (aa) alteration, while synonymous substitutions involve the third position of the codon and do not change the aa. When the ratio (ω) d_N/d_S is significantly higher than 1, we assume that the codon is under positive (adaptive, Darwin or diversifying) selection (Kimura, 1977), as the frequency of mutations is high and they are rapidly fixed in the population providing a fitness advantage (Kelley and Swanson, 2008).

Genes of the immune system are examples of multigene families, i.e., a group of genes that have originated from a common ancestral gene. These genes

encode proteins with similar sequences and functions. The “birth and death” model is the most accepted model for the evolution of multigene families (Nei and Rooney, 2005) and proposes that new genes arise by gene duplication creating differentiated family members (Kaessmann, 2010; Ohno, 1970). This mechanism is also an evolutionary strategy by giving to the host the flexibility to rapidly evolve and adapt to invading pathogens (Duggal and Emerman, 2012). Genes resulting from duplication can be maintained in the genome through subfunctionalization and neofunctionalization or lost by pseudogenization. Pseudogenization leads to non-functional genes as a result of nucleotide mutations that disrupt the coding region. When subfunctionalization occurs, the two copies of the gene are maintained in the genome and share the ancestral functions. In neofunctionalization one copy retains the ancestral function while the other acquires new functions (Kaessmann, 2010; Levasseur and Pontarotti, 2011; Zhang, 2003).

3. THE EUROPEAN RABBIT

The European rabbit (*Oryctolagus cuniculus* – Linnaeus 1758) belongs to the Leporidae family of the order Lagomorpha, which along with rodents constitutes the Glires clade. It originated in the Iberian Peninsula in the middle Pleistocene period (7.8-1.3 Million years ago) (Lopez-Martinez, 2008) and nowadays it is spread worldwide. In Australia, the European rabbit is considered a pest since it has a severe negative impact in the Australian ecosystems (Cooke, 2012). In contrast, in the Iberian Peninsula it is considered a keystone species, crucial for conservation of highly endangered species (Iberian lynx and the Imperial eagle) (Delibes-Mateos et al., 2008; Monterroso et al., 2016). In addition, it is also considered a relevant game species and it is important for meat, wool and fur production and also for different biotechnological applications.

The European rabbit natural history, biology and the richly documented past of anthropogenic introductions, have made it a good model for several biological questions. Indeed, this species has been extensively used to study hybridization due to the existence of a narrow contact between two genetically and morphologically distinct subspecies, the domestication process and the adaptation to colonization of more than 800 islands. The European rabbit is also a good

model to understand the co-evolution between host and pathogens with two viral diseases affecting the natural populations, the myxomatosis and the rabbit hemorrhagic disease. Furthermore, the pioneer studies in basic immunology, rabies and syphilis, conducted in rabbits contributed for some of the major advances in biomedicine. Although replaced as a model by mouse in the 1980s, the European rabbit is still widely used as model in biomedical research. Recently it has been shown that the rabbit could be the elected model to study viral diseases that affect the humans, like HIV, hepatitis C and viral fulminant hepatic failure.

Hybridization model

The European rabbit is native to the Iberian Peninsula where two subspecies exist, *O. c. cuniculus* and *O. c. algirus* (Cabrera, 1914). These subspecies diverged between 1.8 and 2 million years ago (Branco et al., 2000; Carneiro et al., 2009) and while *O. c. algirus* originated in the southwest of the Iberian Peninsula with stony habitats and scrubs, the *O. c. cuniculus* originated in the northeast peninsula with open pasture and burrows for shelters (Biju-Duval et al., 1991; Branco et al., 2000; Carneiro et al., 2009). First described by Phoenicians during their navigations in the Mediterranean (Ferrand, 2008; Suckow, 2012), these subspecies have phenotypic (body size and cranial measurements) and genetic differences (Ferrand and Branco, 2007; Ferreira, 2012; Sharples et al., 1996; Villafuerte, 2002). They evolved allopatrically during the Pleistocene glaciation, however after this period, both expanded establishing a contact zone, along a northwest-southeast direction in the Iberian Peninsula (Alda and Doadrio, 2014; Branco et al., 2000; Branco et al., 2002; Carneiro et al., 2013). Furthermore, there is evidence of high rates of gene flow between the two subspecies (Carneiro et al., 2009). This hybrid zone constitutes a good model to study gene flow between subspecies and evaluate the hybrids fitness, as illustrated by the contrasting patterns observed in the mitochondrial DNA, Y chromosome and nuclear genes (Branco et al., 2000; Campos et al., 2007; Carneiro et al., 2013; Geraldes et al., 2008; Geraldes and Ferrand, 2006; Geraldes et al., 2005).

Island colonization

The majority of the European rabbit population expansions were human-mediated and the first records are associated with the Phoenicians during their path through the Mediterranean Sea. In the 13th and 14th centuries navigators were responsible for the introduction of the *O. cuniculus* spp. in several islands. The main objective was to obtain a source of meat. During 1300-1400 B.C., Portuguese and Spanish settlers released *O. c. cuniculus* in Balearics (Seixas et al., 2014) and in the Tunisian island of Zembra (Hardy et al., 1994). On the other hand, the *O. c. algirus* was introduced during the 15th century in the archipelagos of Azores, Madeira and Canary where a strong bottleneck effect is described (Ferrand, 2008). The introduction of the European rabbit into Australia is the most emblematic and well-documented introduction. Indeed, from a release of 20 individuals in 1890 in the Southwest of Australia, the absence of predators and competitors allowed the establishment and expansion of the European rabbits to all Australia. This colonization was so successful that nowadays the European rabbit is considered a major pest in Australia. Furthermore, the European rabbit is present in more than 800 islands all over the world which points out the success of this species in the colonization of new environments (Flux, 1994). The reasons for this success are particular features such as high growth and reproduction rate, an efficient food utilization (associated with coprophagy, i.e. the capacity to eat the faeces to increase nutritional efficiency), and the ability to shift between growth strategies ("r" or "K - selection") according to the carrying capacity of the environment. Other aspects are also fundamental in its success in islands, like the absence of competitors, the existence of few predators and the inexistence of diseases. These European rabbit populations could be seen as natural populations in nearly "laboratory" conditions, which can be very useful in conservational genetics issues, because they represent the ideal situation where consequences of an extreme bottleneck event can be studied. Indeed, the study of founder effects has become increasingly important in population genetics, speciation theory and conservation biology (reviewed in (Pinheiro et al., 2016)).

Domestication

The European rabbit is the only mammal species that has been domesticated in Western Europe (Monnerot et al., 1994), the species for which wild populations and domestic animals are more closely related (Carneiro et al., 2011). The first records that describe rabbits in captivity are associated with Romans during their occupation of the Iberian Peninsula in the first century BC. The main goal was to obtain rabbit meat and fur (Clutton-Brock et al., 1999). The true domestication process, however, with taming and selective breeding, is recent and was initiated in monasteries on the Champagne region in the northeast of France in the 16th century (Carneiro et al., 2014). Currently, there are more than 45 breeds recognized by the American Rabbit Breeders Association. In addition, the European rabbit is also frequently found as a pet animal.

All domestic rabbit breeds belong to the subspecies *O. c. cuniculus* (Carneiro et al., 2011). The expansion of wild populations belonging to the subspecies *O. c. cuniculus* from Iberia to France across the French Pyrenees occurred with a bottleneck that originated a reduction in the genetic diversity of the French populations. Then, when the domestication event occurred in France, an additional bottleneck affected the gene pool of domesticated European rabbits (Abrantes et al., 2013a; Alves et al., 2015; Carneiro et al., 2011; Esteves et al., 2004; Ferrand and Branco, 2007; Queney et al., 2001; Surridge et al., 2008; van der Loo et al., 1999). Indeed, the domestic breeds lost more than 70% of all the diversity observed in the Iberian Peninsula and 25% of the diversity observed in the wild French rabbit populations (Alves et al., 2015). The recent and geographically defined origin, resulting in the close relatedness between wild and domestic animals, makes the European rabbit a very good model to study the genetic changes that accompanied domestication. Several genetic studies showed associations between specific traits in the European rabbit breeds and genetic variability. For example, Fontanesi et al., 2006, demonstrated that different alleles of the melanocortin 1 receptor (MC1R) gene are associated with different coat colors (Fontanesi et al., 2006). Carneiro and co-workers (2014) showed that genes affecting brain and neuronal development have often been targeted during domestication and the tame behavior evolved by shifts in allele frequencies at many loci, rather than by critical changes at only a few domestication loci

(Carneiro et al., 2014). Candidate genes underlying heritable differences in reproductive seasonality between wild and domestic rabbits have been identified (Carneiro et al., 2016) and it was recently shown that Dwarfism and altered craniofacial development in the domestic rabbits is caused by a 12.1 kb deletion at the HMGA2 locus (Carneiro et al., 2016).

The European rabbit in biomedicine

The rabbit as a model to study host-pathogen interaction

In the last 60 years, the Iberian rabbit populations suffered a sharp decline due to hunting, habitat destruction and the emergence of two viral diseases, myxomatosis in the 1950s and rabbit hemorrhagic disease (RHD) in the 1980s (Delibes-Mateos et al., 2009). These diseases led to a contraction of the wild rabbit populations with serious ecological and economic consequences for the ecosystem. In Australia, where the European rabbit is considered a pest, the dynamics observed between the genetic resistance of the European rabbit natural populations against myxomatosis and the virulence grades of its causative agent, the myxoma virus, is one of the best text-book examples of co-evolution (Kerr, 2012). Regarding the RHD, caused by the rabbit hemorrhagic disease virus, the natural populations suffered a sharp decline in the end of 1980s and then recovered until reaching pre-RHDV effective sizes by 2008 (Elsworth and Cooke, 2008), but in 2011 the emergence of a new variant, genetically and antigenically highly distinct from the previous described strains, led to a new contraction of the natural populations' effective size (Abrantes et al., 2013b; Abrantes et al., 2014; Dalton et al., 2012; Le Gall-Recule et al., 2011; Lopes et al., 2014). This new variant shows frequent recombination and is still highly lethal (Lopes et al., 2015) which makes this situation an opportunity to understand how the host and the virus will co-evolve.

Myxomatosis

Myxomatosis is a viral disease caused by the myxoma virus (MYXV) of the family *Poxviridae*. The MYXV is naturally found in *Sylvilagus* spp., where it causes

a benign cutaneous fibroma, however in the European rabbit it proved to be a highly fatal pathogen (Fenner and Ratcliffe, 1965; Kerr and Donnelly, 2013; Kerr et al., 2015; Marshall and Regnery, 1960). The main symptoms of infection are swelling of the body, conjunctival inflammation and mucopurulent release from the nose and eyes which may occur 4 to 8 days after infection. Usually these symptoms become more severe leading to secondary infections and, ultimately, to death (Best and Kerr, 2000). First identified in 1896 in rabbits from a laboratory in South America (Sanarelli, 1898), MYXV was later deliberately introduced in Australia (1950) and in France (1952) in order to control wild rabbit populations (Ratcliffe et al., 1952). In Australia this killed approximately 99.8% of the infected rabbits while in France, MYXV killed over 90% of the wild population, affecting also 30-40% of the domestic rabbits (Kerr, 2012). After its introduction in France, the MYXV rapidly spread to other European countries arriving in 1953 to Britain (mortality over 99%) and to the Iberian Peninsula (mortality over 90%) (Barcena et al., 2000; Fenner and Ratcliffe, 1965; Fenner and Ross, 1994; Muñoz Goyanes, 1960). Currently, some rabbit populations still experience high mortality rates while in others, the disease has almost no impact (Marchandeau et al., 2014).

In Australia, after the first epizootics that decimated the European rabbit populations, rabbit numbers recovered to some extent (Fenner, 1953; Fenner et al., 1953). Over the next 30 years mildly attenuated MYXV strains that naturally emerged in the field out-competed the most virulent strain that had been initially introduced. These mildly attenuated strains allowed the infected rabbits to survive longer and, as such, were more readily transmitted by the mosquito vectors. At the same time rabbit populations were increasingly resistant to the disease as the breeding population became dominated by survivors to myxomatosis (reviewed in Kerr et al, 2015).

Rabbit hemorrhagic disease

Rabbit hemorrhagic disease is a viral disease caused by a *calicivirus*, the rabbit hemorrhagic disease virus (RHDV), that causes a severe, lethal and contagious disease that affects both domestic and wild populations (Abrantes et al., 2012). Clinical signs of RHD are inconspicuous, but a blood mucous discharge from the nose might be observed, and death occurs within 24-48 hours post

infection. The main histopathological alterations are observed upon necropsy with hemorrhages in several organs, mainly lungs and kidneys, discolored liver and splenomegaly (Cooke, 2002; Kerr and Donnelly, 2013). First described in China in 1984 following the importation of Angora rabbits from Germany, RHD had a tremendous impact in rabbitries killing approximately 140 million rabbits (Liu et al., 1984; Xu, 1991). The virus quickly spread to other countries in Asia (Park N Y et al., 1987), Europe (Cancellotti and Renzi, 1991), America (Gregg and House, 1989) and Africa (Morisse et al., 1991) in the following years. The first cases of RHD in the Iberian Peninsula occurred in 1988 (Spain) (Argüello et al., 1988) and 1989 (Portugal) (Anonymous, 1989), further reducing the small extant populations that were still recovering from myxomatosis. Such as with MYXV, RHDV was also deliberately introduced in Australia in 1995 as a biocontrol agent. The first trials began on the Wardang Island (5 kilometers of the South Australian coast), however it quickly spread to the mainland decreasing the rabbit population in ~60% (Cooke, 2002; Cooke, 2007, 2012).

Interestingly, young rabbits less than 2 months are not affected by RHD, probably due to their immunity acquired by vertical transmission (Cooke, 2014; Marques et al., 2012; Marques et al., 2014; Prieto et al., 2000), which allow them to develop an immune response that blocks the virus replication, while adult rabbits die before developing such an effective immune response. More recently, a new variant of RHDV (RHDV2 or RHDVb) that is able to kill young rabbits was identified in France (Le Gall-Recule et al., 2011). This new variant, which differs more than 15% of nucleotide diversity from the old RHDV strains, rapidly disseminated across Europe and caused mortalities similar to those observed in the first RHDV outbreaks (Abrantes et al., 2013b; Abrantes et al., 2014; Baily et al., 2014; Dalton et al., 2014; Le Gall-Recule et al., 2011). In Portugal, this new variant replaced the older strains and its evolution is associated with several recombination events (Lopes et al., 2015). The emergence of RHDV2 and even the origin of RHDV remain poorly understood. Two major hypotheses are debated and propose that the pathogenic virus either evolved from non-pathogenic virus strains circulating in European leporids or through a species jump from non-native leporids, possibly the American eastern cottontail, introduced into Europe in the 1960s-1970s (Esteves et al 2015).

Rabbit as a model

Laboratory animals have long been used as models to improve our understanding of several human conditions. The main goal is to artificially replicate in the laboratory animal the condition under study. The European rabbit was the first animal model used in several immunological studies, being crucial, for example, for the development of the rabies vaccine by Louis Pasteur in 1881 (Pasteur, 1885). Furthermore, it was the study of rabbit immunoglobulins that established much of what is known about immunoglobulins structure, function and regulated expression (Pinheiro et al., 2011; Pinheiro et al., 2016).

Despite being the only animal model for the study of molecular immunology in the late 1980s, in the following years rabbits were replaced by rodents (Burkholder et al., 2012). The main reasons for the use of rodents instead of rabbits are the low maintenance costs, smaller size, ease of breeding, short reproductive cycle, high number of progeny, lower inter-individual variability and a wide availability of commercial reagents (Mullane and Williams, 2014; Webb, 2014). However, rabbits have the advantage to be in-between the smaller animals as rodents and the larger ones as primates that are too costly (Fontanesi et al., 2016). In addition, in genes of the immune system, rabbits are more similar to primates than rodents (Graur et al., 1996; Neves et al., 2015; Perkins et al., 2000), they present a longer lifetime and similar organs' size to humans, and they are also carriers or reservoirs of several pathogens that can cause zoonotic diseases (e.g. tularemia, Lyme borreliosis, paratuberculosis, hemorrhagic and diarrheal diseases) (Fontanesi et al., 2016; Hill and Brown, 2011). Furthermore, some studies highlight the lower success rate of the mouse as a model in studying human diseases (James et al., 2005; Seok et al., 2013; Takao and Miyakawa, 2015). Indeed, the European rabbit is highly used as a laboratory model for immunological studies in different areas: atherosclerosis (Fan et al., 2015; Tian et al., 2012), intestinal immunity (Jimenez-Garcia et al., 2004), reproduction (Fischer et al., 2012), arthritis (Desando et al., 2013), cancer (Kang and Grossniklaus, 2011), Alzheimer's disease (Woodruff-Pak et al., 2007), viral infections (Cheng et al., 2012; Ma et al., 2010) and infectious diseases (Burkholder et al., 2012; Peng et al., 2015). One of the most important features is its ability to produce specific

and with high affinity antibodies that are largely used against several target antigens (Fontanesi et al., 2016; Jacquier et al., 2015).

Studies of Human immune-deficiency virus type 1 (HIV-1) pathogenesis and new possible treatments are a good example of the advantageous use of rabbit as a model as opposed to rodent and simian models. Though having allowed major breakthroughs in the understanding of HIV pathogenesis and treatment, both simian and rodent models have serious limitations that are overcome with the rabbit model. Tervo and Keppler (2010) demonstrated that rabbits show a natural HIV-1 permissivity and propose that it is possible to render this species fully permissive to infection by HIV-1 (Tervo and Keppler, 2010). It has been shown that the European rabbit can be used as an animal model in AIDS vaccine development (Chen et al., 2013; Pan et al., 2013). Indeed, the European rabbit antibodies can recognize immunogenic regions of gp120 and mimic the binding modes of human antibodies (Pan et al., 2013). The study of APOBEC1, an important anti-viral molecule, showed that the European rabbit is the species with the strongest activity to inhibit the infectivity of HIV-1 in 293T cell-based assays (Ikeda et al., 2008). Anchored in this observation, the creation of chimeric human/rabbit APOBEC1 with HIV-1 restrictions and DNA mutation activities showed that the C-terminal region of rabbit apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1 (APOBEC1) is involved in both its packaging into the HIV-1 virion and its deamination activity against both viral cDNA and genomic RNA (Ikeda et al., 2016).

Chemokines and their receptors play crucial roles in immune and inflammatory responses. C–C motif chemokine receptor 5 (CCR5) is one of the most well studied by the relevant role it has in the HIV infection process being involved in the HIV uptake into host cells (Barmania and Pepper, 2013). The CCR5 of European rabbit and two other related leporids, *Bunolagus* and *Pentalagus,* is unique among mammals. Indeed, the three genera suffered a change at the second extracellular loop resulting from a gene conversion event with the paralogous CCR2 (Abrantes et al., 2011; Carmo et al., 2006). This unique feature makes the European rabbit a suitable model to study the co-evolution between this receptor and its ligands. The study of its putative ligands showed different evolutionary patterns with a strong purifying selection in CCL3, CCL4 and CCL5 (de Matos et al., 2014) while CCL8 contrastingly showed a

pseudogenization specific to these three genera (van der Loo et al., 2012). The correlation observed between CCR5 and CCL8 evolution remains to be functionally evaluated.

The discovery of endogenous lentiviruses in rabbits (RELIK) showed that lentiviruses had the ability for germline integration (Katzourakis et al., 2007; Yap and Stoye, 2013). Furthermore, the comparison of their genomic structures showed high similarity with the modern-day lentiviruses (de Sousa-Pereira et al., 2016; Goldstone et al., 2010). Thus, the RELIK could provide important insights into lentivirus biology and host interactions.

Fragments homologous to Hepatitis C Virus were found to be present in the European rabbit (*Oryctolagus cuniculus*) and hare (*Lepus europaeus*) genomes and were successfully replicated in bovine cell cultures (Silva et al., 2012). Additionally, it was suggested that these genomic fragments can be internalized in the MDBK cells and together with the cell machinery initiate the replication and the generation of novel HCV-like virus (Silva et al., 2015). These results give important insights on the HCV evolution and diversity and open new avenues for therapeutic approaches.

The similarities of hepatic damage caused by RHDV infection and human fulminant hepatic failure (FHF) of viral origin (Mikami et al., 1999) suggested that RHDV infection would be a good model to study FHF. Tuñon and co-workers (2003) successfully characterized this new animal model for acute liver failure and have since been studying RHDV infection as to aid in finding new therapeutics to protect livers from acute liver failure. Their studies have focused on the effects of melatonin (Crespo et al., 2016; Laliena et al., 2012; San-Miguel et al., 2014; Tunon et al., 2013; Tunon et al., 2011a), but also on other signaling pathways involved in the autophagic response induced by RHDV (Garcia-Lastra et al., 2010; San-Miguel et al., 2006; Vallejo et al., 2014).

4. REFERENCES

Abrantes, J., Areal, H., Esteves, P.J., 2013a. **Insights into the European rabbit (*Oryctolagus cuniculus*) innate immune system: genetic diversity of the toll-like receptor 3 (TLR3) in wild populations and domestic breeds**. BMC Genet 14, 73.

Abrantes, J., Carmo, C.R., Matthee, C.A., Yamada, F., van der Loo, W., Esteves, P.J., 2011. **A shared unusual genetic change at the chemokine receptor type 5 between *Oryctolagus*, *Bunolagus* and *Pentalagus***. Conserv Genet 12, 325-330.

Abrantes, J., Esteves, P.J., Carmo, C.R., Muller, A., Thompson, G., van der Loo, W., 2008. **Genetic characterization of the chemokine receptor CXCR4 gene in lagomorphs: comparison between the families Ochotonidae and Leporidae**. Int J Immunogenet 35, 111-117.

Abrantes, J., Lopes, A.M., Dalton, K.P., Melo, P., Correia, J.J., Ramada, M., Alves, P.C., Parra, F., Esteves, P.J., 2013b. **New variant of rabbit hemorrhagic disease virus, Portugal, 2012-2013**. Emerg Infect Dis 19, 1900-1902.

Abrantes, J., Lopes, A.M., Dalton, K.P., Parra, F., Esteves, P.J., 2014. **Detection of RHDVa on the Iberian Peninsula: isolation of an RHDVa strain from a Spanish rabbitry**. Arch Virol 159, 321-326.

Abrantes, J., van der Loo, W., Le Pendu, J., Esteves, P.J., 2012. **Rabbit haemorrhagic disease (RHD) and rabbit haemorrhagic disease virus (RHDV): a review**. Vet Res 43, 12.

Afzal, N., Tahir, R., Jahan, S., 2012. **Cytokines: an ever expanding area** Biological and Biomedical Reports 2, 37-43.

Akdis, M., Burgler, S., Crameri, R., Eiwegger, T., Fujita, H., Gomez, E., Klunker, S., Meyer, N., O'Mahony, L., Palomares, O., Rhyner, C., Ouaked, N., Schaffartzik, A., Van De Veen, W., Zeller, S., Zimmermann, M., Akdis, C.A., 2011. **Interleukins, from 1 to 37, and interferon-gamma: receptors, functions, and roles in diseases**. J Allergy Clin Immunol 127, 701-721 e701-770.

Alda, F., Doadrio, I., 2014. **Spatial genetic structure across a hybrid zone between European rabbit subspecies**. PeerJ 2, e582.

Alves, J.M., Carneiro, M., Afonso, S., Lopes, S., Garreau, H., Boucher, S., Allain, D., Queney, G., Esteves, P.J., Bolet, G., Ferrand, N., 2015. **Levels and Patterns of Genetic Diversity and Population Structure in Domestic Rabbits**. PLoS One 10, e0144687.

Alves, P.C., Hacklander, K., 2008. **Lagomorph Species: Geographical Distribution and Conservation Status**, in: Alves, P.C., Ferrand, N., Hackländer, K. (Eds.), Lagomorph Biology: Evolution, Ecology and Conservation. Springer, pp. 388-405.

Anonymous, 1989. **Doença hemorrágica a vírus do Coelho em Portugal**. Re Port Ciênc Vet 84.

Arbiza, L., Dopazo, J., Dopazo, H., 2006. **Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome**. PLoS Comput Biol 2, e38.

Areal, H., Abrantes, J., Esteves, P.J., 2011. **Signatures of positive selection in Toll-like receptor (TLR) genes in mammals**. BMC Evol Biol 11, 368.

Arend, W.P., Palmer, G., Gabay, C., 2008. **IL-1, IL-18, and IL-33 families of cytokines**. Immunol Rev 223, 20-38.

Argüello, J.L., Llanos, A., Pérez, L.I., 1988. **Enfermedad hemorrágica del conejo en España**. Revue de Médecine Vétérinaire 5.

Baily, J.L., Dagleish, M.P., Graham, M., Maley, M., Rocchi, M.S., 2014. **RHDV variant 2 presence detected in Scotland**. Vet Rec 174, 411.

Banchereau, J., Pascual, V., O'Garra, A., 2012. **From IL-2 to IL-37: the expanding spectrum of anti-inflammatory cytokines**. Nat Immunol 13, 925-931.

Bao, K., Reinhardt, R.L., 2015. **The differential expression of IL-4 and IL-13 and its impact on type-2 immunity**. Cytokine 75, 25-37.

Barcena, J., Morales, M., Vazquez, B., Boga, J.A., Parra, F., Lucientes, J., Pages-Mante, A., Sanchez-Vizcaino, J.M., Blasco, R., Torres, J.M., 2000. **Horizontal transmissible protection against myxomatosis and rabbit hemorrhagic disease by using a recombinant myxoma virus**. J Virol 74, 1114-1123.

Barmania, F., Pepper, M.S., 2013. **C-C chemokine receptor type five (CCR5): An emerging target for the control of HIV infection**. Appl Transl Genom 2, 3-16.

Barreiro, L.B., Quintana-Murci, L., 2010. **From evolutionary genetics to human immunology: how selection shapes host defence genes**. Nature reviews. Genetics 11, 17-30.

Beadling, C., Slifka, M.K., 2006. **Regulation of innate and adaptive immune responses by the related cytokines IL-12, IL-23, and IL-27**. Arch Immunol Ther Exp (Warsz) 54, 15-24.

Best, S.M., Kerr, P.J., 2000. **Coevolution of host and virus: the pathogenesis of virulent and attenuated strains of myxoma virus in resistant and susceptible European rabbits**. Virology 267, 36-48.

Biju-Duval, C., Ennafaa, H., Dennebouy, N., Monnerot, M., Mignotte, F., Soriguer, R.C., El Gaaied, A., El Hili, A., Mounolou, J.C., 1991. **Mitochondrial DNA evolution in lagomorphs: origin of systematic heteroplasmy and organization of diversity in European rabbits**. Journal of molecular evolution 33, 92-102.

Borish, L.C., Steinke, J.W., 2003. **2. Cytokines and chemokines**. J Allergy Clin Immunol 111, S460-475.

Bowler, R.P., Bahr, T.M., Hughes, G., Lutz, S., Kim, Y.I., Coldren, C.D., Reisdorph, N., Kechris, K.J., 2013. **Integrative omics approach identifies interleukin-16 as a biomarker of emphysema**. OMICS 17, 619-626.

Boyman, O., Sprent, J., 2012. **The role of interleukin-2 during homeostasis and activation of the immune system**. Nat Rev Immunol 12, 180-190.

Branco, M., Ferrand, N., Monnerot, M., 2000. **Phylogeography of the European rabbit (Oryctolagus cuniculus) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome b gene**. Heredity (Edinb) 85 Pt 4, 307-317.

Branco, M., Monnerot, M., Ferrand, N., Templeton, A.R., 2002. **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 56, 792-803.

Brocker, C., Thompson, D., Matsumoto, A., Nebert, D.W., Vasiliou, V., 2010. **Evolutionary divergence and functions of the human interleukin (IL) gene family**. Hum Genomics 5, 30-55.

Bryant, V.L., Slade, C.A., 2015. **Chemokines, their receptors and human disease: the good, the bad and the itchy**. Immunol Cell Biol 93, 364-371.

Burkholder, T.H., Linton, G., F., R., Hoyt, J., Young, R., 2012. **The rabbit as an experimental model**, in: Suckow, M.A., Stevens, K.A., Wilson, R.P. (Eds.), The Laboratory Rabbit, Guinea Pig, Hamster, and Other Rodents. Elsevier.

Cabrera, A., 1914. **Fauna ibérica**, Madrid.

Campos, R., Branco, M., Weiss, S., Ferrand, N., 2007. **Patterns of hemoglobin polymorphism [α-globin (HBA) and β-globin (HBB)] across the contact zone of two distinct phylogeographic lineages of the European rabbit (Oryctolagus cuniculus)**, in: Weiss, S., Ferrand, N. (Eds.), Phylogeography of Southern European Refugia: Evolutionary perspectives on the origins and conservation of European biodiversity. Springer Netherlands, Dordrecht, pp. 237-255.

Cancellotti, F.M., Renzi, M., 1991. **Epidemiology and current situation of viral haemorrhagic disease of rabbits and the European brown hare syndrome in Italy**. Rev Sci Tech 10, 409-422.

Carmo, C.R., Esteves, P.J., Ferrand, N., van der Loo, W., 2006. **Genetic variation at chemokine receptor CCR5 in leporids: alteration at the 2nd extracellular domain by gene conversion with CCR2 in Oryctolagus, but not in Sylvilagus and Lepus species**. Immunogenetics 58, 494-501.

Carneiro, M., Afonso, S., Geraldes, A., Garreau, H., Bolet, G., Boucher, S., Tircazes, A., Queney, G., Nachman, M.W., Ferrand, N., 2011. **The genetic structure of domestic rabbits**. Mol Biol Evol 28, 1801-1816.

Carneiro, M., Baird, S.J., Afonso, S., Ramirez, E., Tarroso, P., Teotonio, H., Villafuerte, R., Nachman, M.W., Ferrand, N., 2013. **Steep clines within a highly permeable genome across a hybrid zone between two subspecies of the European rabbit**. Mol Ecol 22, 2511-2525.

Carneiro, M., Ferrand, N., Nachman, M.W., 2009. **Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (Oryctolagus cuniculus)**. Genetics 181, 593-606.

Carneiro, M., Hu, D., Archer, J., Feng, C., Afonso, S., Chen, C., Blanco-Aguiar, J.A., Garreau, H., Boucher, S., Ferreira, P.G., Ferrand, N., Rubin, C.J., Andersson, L., 2016. **Dwarfism and Altered Craniofacial Development in Rabbits Is Caused by a 12.1 kb Deletion at the HMGA2 Locus**. Genetics.

Carneiro, M., Rubin, C.J., Di Palma, F., Albert, F.W., Alfoldi, J., Barrio, A.M., Pielberg, G., Rafati, N., Sayyab, S., Turner-Maier, J., Younis, S., Afonso, S., Aken, B., Alves, J.M., Barrell, D., Bolet, G., Boucher, S., Burbano, H.A., Campos, R., Chang, J.L., Duranthon, V., Fontanesi, L., Garreau, H., Heiman, D., Johnson, J., Mage, R.G., Peng, Z., Queney, G., Rogel-Gaillard, C., Ruffier, M., Searle, S., Villafuerte, R., Xiong, A., Young, S., Forsberg-Nilsson, K., Good, J.M., Lander, E.S., Ferrand,

N., Lindblad-Toh, K., Andersson, L., 2014. **Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication**. Science 345, 1074-1079.

Castellani, M.L., Felaco, P., Galzio, R.J., Tripodi, D., Toniato, E., De Lutiis, M.A., Fulcheri, M., Caraffa, A., Antinolfi, P., Tete, S., Felaco, M., Conti, F., Pandolfi, F., Theoharides, T.C., Shaik-Dasthagirisaheb, Y.B., 2010. **IL-31 a Th2 cytokine involved in immunity and inflammation**. Int J Immunopathol Pharmacol 23, 709-713.

Chaplin, D.D., 2010. **Overview of the immune response**. J Allergy Clin Immunol 125, S3-23.

Chapman, J.A., Flux, J.E.C., 2008. **Introduction to the Lagomorpha**, in: Alves, P.C., Ferrand, N., Hackländer, K. (Eds.), Lagomorph Biology: Evolution, Ecology and Conservation. Springer, pp. 1-9.

Chapman, J.A., Flux, J.E.C., Group, I.U.f.C.o.N.a.N.R.L.S., 1990. **Rabbits, hares, and pikas : status survey and conservation action plan.**, Gland, Switzerland.

Chen, Y., Vaine, M., Wallace, A., Han, D., Wan, S., Seaman, M.S., Montefiori, D., Wang, S., Lu, S., 2013. **A novel rabbit monoclonal antibody platform to dissect the diverse repertoire of antibody epitopes for HIV-1 Env immunogen design**. J Virol 87, 10232-10243.

Cheng, X., Wang, S., Dai, X., Shi, C., Wen, Y., Zhu, M., Zhan, S., Meng, J., 2012. **Rabbit as a novel animal model for hepatitis E virus infection and vaccine evaluation**. PloS one 7.

Clutton-Brock, J., Clutton-Brock, J., Natural History Museum (London England), 1999. **A natural history of domesticated mammals**, 2nd ed. Cambridge University Press Cambridge,.

Commins, S.P., Borish, L., Steinke, J.W., 2010. **Immunologic messenger molecules: cytokines, interferons, and chemokines**. J Allergy Clin Immunol 125, S53-72.

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

Cooke, B.D., 2007. **A review of Rabbit Haemorrhagic Disease in Australia** Australian Wool Innovation and Meat and Livestock Australia.

Cooke, B.D., 2012. **Rabbits: manageable environmental pests or participants in new Australian ecosystems?** Wildlife Research 39, 279-289.

Cooke, B.D., 2014. **Australia's war against rabbits : the story of rabbit haemorrhagic disease**. CSIRO Publishing, Australia.

Cornelissen, C., Luscher-Firzlaff, J., Baron, J.M., Luscher, B., 2012. **Signaling by IL-31 and functional consequences**. Eur J Cell Biol 91, 552-566.

Couper, K.N., Blount, D.G., Riley, E.M., 2008. **IL-10: the master regulator of immunity to infection**. J Immunol 180, 5771-5777.

Crespo, I., San-Miguel, B., Sanchez, D.I., Gonzalez-Fernandez, B., Alvarez, M., Gonzalez-Gallego, J., Tunon, M.J., 2016. **Melatonin inhibits the sphingosine kinase 1/sphingosine-1-phosphate signaling pathway in rabbits with fulminant hepatitis of viral origin**. J Pineal Res 61, 168-176.

Croce, M., Rigo, V., Ferrini, S., 2015. **IL-21: a pleiotropic cytokine with potential applications in oncology**. J Immunol Res 2015, 696578.

Cruikshank, W., Little, F., 2008. **Interleukin-16: the ins and outs of regulating T-cell activation**. Crit Rev Immunol 28, 467-483.

Curfs, J.H., Meis, J.F., Hoogkamp-Korstanje, J.A., 1997. **A primer on cytokines: sources, receptors, effects, and inducers**. Clin Microbiol Rev 10, 742-780.

Dalton, K.P., Nicieza, I., Abrantes, J., Esteves, P.J., Parra, F., 2014. **Spread of new variant RHDV in domestic rabbits on the Iberian Peninsula**. Vet Microbiol 169, 67-73.

Dalton, K.P., Nicieza, I., Balseiro, A., Muguerza, M.A., Rosell, J.M., Casais, R., Alvarez, A.L., Parra, F., 2012. **Variant rabbit hemorrhagic disease virus in young rabbits, Spain**. Emerg Infect Dis 18, 2009-2012.

de Matos, A.L., Lanning, D.K., Esteves, P.J., 2014. **Genetic characterization of CCL3, CCL4 and CCL5 in leporid genera Oryctolagus, Sylvilagus and Lepus**. Int J Immunogenet 41, 154-158.

de Sousa-Pereira, P., Abrantes, J., Baldauf, H.M., Keppler, O.T., Esteves, P.J., 2016. **Evolutionary study of leporid CD4 reveals a hotspot of genetic variability within the D2 domain**. Immunogenetics 68, 477-482.

Delibes-Mateos, M., Delibes, M., Ferreras, P., Villafuerte, R., 2008. **Key role of European rabbits in the conservation of the Western Mediterranean basin hotspot**. Conserv Biol 22, 1106-1117.

Delibes-Mateos, M., Ferreras, P., Villafuerte, R., 2009. **European rabbit population trends and associated factors: a review of the situation in the Iberian Peninsula**. Mammal Rev 39, 124-140.

Deng, L., Luo, M., Velikovsky, A., Mariuzza, R.A., 2013. **Structural insights into the evolution of the adaptive immune system**. Annu Rev Biophys 42, 191-215.

Desando, G., Cavallo, C., Sartoni, F., Martini, L., Parrilli, A., Veronesi, F., Fini, M., Giardino, R., Facchini, A., Grigolo, B., 2013. **Intra-articular delivery of adipose derived stromal cells attenuates osteoarthritis progression in an experimental rabbit model**. Arthritis Res Ther 15, R22.

Dinarello, C.A., Kim, S.H., 2006. **IL-32, a novel cytokine with a possible role in disease**. Ann Rheum Dis 65 Suppl 3, iii61-64.

Donnelly, R.P., Sheikh, F., Dickensheets, H., Savan, R., Young, H.A., Walter, M.R., 2010. **Interleukin-26: an IL-10-related cytokine produced by Th17 cells**. Cytokine Growth Factor Rev 21, 393-401.

Douzery, E.J.P., Huchon, D., 2004. **Rabbits, if anything, are likely Glires**. Mol Phylogenet Evol 33, 922-935.

Du, X., Williams, D.A., 1997. **Interleukin-11: review of molecular, cell biology, and clinical use**. Blood 89, 3897-3908.

Dudakov, J.A., Hanash, A.M., van den Brink, M.R., 2015. **Interleukin-22: immunobiology and pathology**. Annu Rev Immunol 33, 747-785.

Duggal, N.K., Emerman, M., 2012. **Evolutionary conflicts between viruses and restriction factors shape immunity**. Nat Rev Immunol 12, 687-695.

Egwuagu, C.E., Yu, C.R., Sun, L., Wang, R., 2015. **Interleukin 35: Critical regulator of immunity and lymphocyte-mediated diseases**. Cytokine Growth Factor Rev 26, 587-593.

Elssner, A., Doseff, A.I., Duncan, M., Kotur, M., Wewers, M.D., 2004. **IL-16 is constitutively present in peripheral blood monocytes and spontaneously released during apoptosis**. J Immunol 172, 7721-7725.

Elsworth, P., Cooke, B., 2008. **Co-evolution of wild rabbits (Oryctolagus cuniculus) and RHDV: resistance and virulence**, 14th Australasian Vertebrate Pest Conference, Canberra, Australia.

Esteves, P.J., Abrantes, J., van der Loo, W., 2007. **Extensive gene conversion between CCR2 and CCR5 in domestic cat (Felis catus)**. Int J Immunogenet 34, 321-324.

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

Fan, J., Kitajima, S., Watanabe, T., Xu, J., Zhang, J., Liu, E., Chen, Y.E., 2015. **Rabbit models for the study of human atherosclerosis: from pathophysiological mechanisms to translational medicine**. Pharmacol Ther 146, 104-119.

Fehniger, T.A., Caligiuri, M.A., 2001. **Interleukin 15: biology and relevance to human disease**. Blood 97, 14-32.

Fenner, F., 1953. **Changes in the mortality-rate due to myxomatosis in the Australian wild rabbit**. Nature 172, 228-230.

Fenner, F., 1956. **Evolutionary aspects of Myxomatosis in Australia**. Mem Inst Oswaldo Cruz 54, 271-278.

Fenner, F., Marshall, I.D., Woodroofe, G.M., 1953. **Studies in the epidemiology of infectious myxomatosis of rabbits. I. Recovery of Australian wild rabbits (Oryctolagus cuniculus) from myxomatosis under field conditions**. J Hyg (Lond) 51, 225-244.

Fenner, F., Ratcliffe, F., 1965. **Myxomatosis**. University Press, Cambridge Eng.

Fenner, F., Ross, J., 1994. **Myxomatosis**, in: Thompson, H.V., King, C.M. (Eds.), The European rabbit: the history and biology of a successful colonizer. Oxford University Press, Oxford, pp. 205-240.

Ferrand, N., 2008. **Inferring the Evolutionary History of the European Rabbit (*Oryctolagus cuniculus*) from Molecular Markers**, in: Alves, P.C., Ferrand, N., Hackländer, K. (Eds.), Lagomorph Biology: Evolution, Ecology and Conservation. Springer, pp. 47-63.

Ferrand, N., Branco, M., 2007. **The evolutionary history of the European rabbit (Oryctolagus cuniculus): major patterns of population differentiation and geographic expansion inferred from protein polymorphism**, in: Weiss, S., Ferrand, N. (Eds.), Phylogeography of Southern European Refugia. Springer, pp. 207-235.

Ferreira, C., 2012. **European rabbit research in the Iberian Peninsula: state of the art and future perspectives**. Eur J Wildl Res 58, 885-895.

Fielding, C.A., 2012. **Interleukin-19: a new target to aim for?** Rheumatology (Oxford) 51, 399-400.

Fischer, B., Chavatte-Palmer, P., Viebahn, C., Navarrete Santos, A., Duranthon, V., 2012. **Rabbit as a reproductive model for human health**. Reproduction 144, 1-10.

Flux, J.E.C., 1994. **World distribution**, in: Thompson, H.V., King, C.M. (Eds.), The European rabbit. The history and biology of a successful colonizer. Oxford Science Publications, Oxford, pp. 8-21.

Fontanesi, L., Di Palma, F., Flicek, P., Smith, A.T., Thulin, C.G., Alves, P.C., Lagomorph Genomics, C., 2016. **LaGomiCs-Lagomorph Genomics Consortium: An International Collaborative Effort for Sequencing the Genomes of an Entire Mammalian Order**. J Hered 107, 295-308.

Fontanesi, L., Tazzoli, M., Beretti, F., Russo, V., 2006. **Mutations in the melanocortin 1 receptor (MC1R) gene are associated with coat colours in the domestic rabbit (Oryctolagus cuniculus)**. Animal genetics 37, 489-493.

Fontes, F.L., Pinheiro, D.M., Oliveira, A.H., Oliveira, R.K., Lajus, T.B., Agnez-Lima, L.F., 2015. **Role of DNA repair in host immune response and inflammation**. Mutat Res Rev Mutat Res 763, 246-257.

Fostowicz-Frelik, L., Frelik, G.J., Gasparik, M., 2010. **Morphological phylogeny of pikas (Lagomorpha: Ochotona), with a description of a new species from the Pliocene/Pleistocene transition of Hungary**. Proceedings of the Academy of Natural Sciences of Philadelphia 159, 97-118.

Fry, T.J., Mackall, C.L., 2002. **Interleukin-7: from bench to clinic**. Blood 99, 3892-3904.

Fry, T.J., Mackall, C.L., 2005. **The many faces of IL-7: from lymphopoiesis to peripheral T cell maintenance**. J Immunol 174, 6571-6576.

Fumagalli, M., Pozzoli, U., Cagliani, R., Comi, G.P., Riva, S., Clerici, M., Bresolin, N., Sironi, M., 2009. **Parasites represent a major selective force for interleukin genes and shape the genetic predisposition to autoimmune conditions**. J Exp Med 206, 1395-1408.

Gabay, C., Towne, J.E., 2015. **Regulation and function of interleukin-36 cytokines in homeostasis and pathological conditions**. J Leukoc Biol 97, 645-652.

Garcia-Lastra, R., San-Miguel, B., Crespo, I., Jorquera, F., Alvarez, M., Gonzalez-Gallego, J., Tunon, M.J., 2010. **Signaling pathways involved in liver injury and regeneration in rabbit hemorrhagic disease, an animal model of virally-induced fulminant hepatic failure**. Vet Res 41, 2.

Garlanda, C., Dinarello, C.A., Mantovani, A., 2013. **The interleukin-1 family: back to the future**. Immunity 39, 1003-1018.

Gasteiger, G., Rudensky, A.Y., 2014. **Interactions between innate and adaptive lymphocytes**. Nat Rev Immunol 14, 631-639.

Ge, D., Wen, Z., Xia, L., Zhang, Z., Erbajeva, M., Huang, C., Yang, Q., 2013. **Evolutionary history of lagomorphs in response to global environmental change**. PLoS One 8, e59668.

Geraldes, A., Carneiro, M., Delibes-Mateos, M., Villafuerte, R., Nachman, M.W., Ferrand, N., 2008. **Reduced introgression of the Y chromosome between subspecies of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula**. Mol Ecol 17, 4489-4499.

Geraldes, A., Ferrand, N., 2006. **A 7-bp insertion in the 3' untranslated region suggests the duplication and concerted evolution of the rabbit SRY gene**. Genet Sel Evol 38, 313-320.

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

Gidley, J.W., 1912. **The Lagomorphs an Independent Order**. Science 36, 285-286.

Goldstone, D.C., Yap, M.W., Robertson, L.E., Haire, L.F., Taylor, W.R., Katzourakis, A., Stoye, J.P., Taylor, I.A., 2010. **Structural and functional analysis of prehistoric lentiviruses uncovers an ancient molecular interface**. Cell host & microbe 8, 248-259.

Gordon, S., 2003. **Alternative activation of macrophages**. Nat Rev Immunol 3, 23-35.

Goswami, R., Kaplan, M.H., 2011. **A brief history of IL-9**. J Immunol 186, 3283-3288.

Graur, D., Duret, L., Gouy, M., 1996. **Phylogenetic position of the order Lagomorpha (rabbits, hares and allies)**. Nature 379, 333-335.

Gregg, D.A., House, C., 1989. **Necrotic hepatitis of rabbits in Mexico: a parvovirus**. Vet Rec 125, 603-604.

Gresnigt, M.S., van de Veerdonk, F.L., 2013. **Biology of IL-36 cytokines and their role in disease**. Semin Immunol 25, 458-465.

Hamza, T., Barnett, J.B., Li, B., 2010. **Interleukin 12 a key immunoregulatory cytokine in infection applications**. Int J Mol Sci 11, 789-806.

Harada, A., Sekido, N., Akahoshi, T., Wada, T., Mukaida, N., Matsushima, K., 1994. **Essential involvement of interleukin-8 (IL-8) in acute inflammation**. Journal of leukocyte biology 56, 559-564.

Hardy, C., Casane, D., Vigne, J.D., Callou, C., Dennebouy, N., Mounolou, J.C., Monnerot, M., 1994. **Ancient DNA from Bronze Age bones of European rabbit (Oryctolagus cuniculus)**. Experientia 50, 564-570.

Heinhuis, B., Netea, M.G., van den Berg, W.B., Dinarello, C.A., Joosten, L.A., 2012. **Interleukin-32: a predominantly intracellular proinflammatory mediator that controls cell activation and cell death**. Cytokine 60, 321-327.

Heinrich, P.C., Behrmann, I., Haan, S., Hermanns, H.M., Muller-Newen, G., Schaper, F., 2003. **Principles of interleukin (IL)-6-type cytokine signalling and its regulation**. The Biochemical journal 374, 1-20.

Hill, W.A., Brown, J.P., 2011. **Zoonoses of rabbits and rodents**. The veterinary clinics of North America. Exotic animal practice 14, 519-531, vii.

Hunter, C.A., Kastelein, R., 2012. **Interleukin-27: balancing protective and pathological immunity**. Immunity 37, 960-969.

Ikeda, T., Ohsugi, T., Kimura, T., Matsushita, S., Maeda, Y., Harada, S., Koito, A., 2008. **The antiretroviral potency of APOBEC1 deaminase from small animal species**. Nucleic Acids Res 36, 6859-6871.

Ikeda, T., Ong, E.B., Watanabe, N., Sakaguchi, N., Maeda, K., Koito, A., 2016. **Creation of chimeric human/rabbit APOBEC1 with HIV-1 restriction and DNA mutation activities**. Scientific reports 6, 19035.

Ishihara, K., Hirano, T., 2002. **Molecular basis of the cell specificity of cytokine action**. Biochim Biophys Acta 1592, 281-296.

Ivashkiv, L.B., Donlin, L.T., 2014. **Regulation of type I interferon responses**. Nat Rev Immunol 14, 36-49.

Iwakura, Y., Ishigame, H., Saijo, S., Nakae, S., 2011. **Functional specialization of interleukin-17 family members**. Immunity 34, 149-162.

Iwasaki, Y., Fujio, K., Okamura, T., Yamamoto, K., 2015. **Interleukin-27 in T cell immunity**. Int J Mol Sci 16, 2851-2863.

Jacquier, V., Estelle, J., Schmaltz-Panneau, B., Lecardonnel, J., Moroldo, M., Lemonnier, G., Turner-Maier, J., Duranthon, V., Oswald, I.P., Gidenne, T., Rogel-Gaillard, C., 2015. **Genome-wide immunity studies in the rabbit: transcriptome variations in peripheral blood mononuclear cells after in vitro stimulation by LPS or PMA-Ionomycin**. BMC genomics 16, 26.

James, J., Martin, L., Krenz, M., Quatman, C., Jones, F., Klevitsky, R., Gulick, J., Robbins, J., 2005. **Forced expression of alpha-myosin heavy chain in the rabbit ventricle results in cardioprotection under cardiomyopathic conditions**. Circulation 111, 2339-2346.

Jimenez-Garcia, A., Balongo-Garcia, R., Alconero, F.F., Araji, O.A., Martinez, G.J., Haba, M.G., Morales, L.C., Bevia, J.M., Martinez, J.C., 2004. **Intestinal wall damage in simple ileus in rabbits: immune-modulator role of somatostatin**. Hepato-gastroenterology 51, 1030-1036.

Jin, W., Dong, C., 2013. **IL-17 cytokines in immunity and inflammation**. Emerging microbes & infections 2, e60.

Johnston, J.B., McFadden, G., 2004. **Technical knockout: understanding poxvirus pathogenesis by selectively deleting viral immunomodulatory genes**. Cellular microbiology 6, 695-705.

Joosten, L.A., Heinhuis, B., Netea, M.G., Dinarello, C.A., 2013. **Novel insights into the biology of interleukin-32**. Cellular and molecular life sciences : CMLS 70, 3883-3892.

Kaessmann, H., 2010. **Origins, evolution, and phenotypic impact of new genes**. Genome research 20, 1313-1326.

Kaiser, P., Rothwell, L., Avery, S., Balu, S., 2004. **Evolution of the interleukins**. Developmental and comparative immunology 28, 375-394.

Kang, S.J., Grossniklaus, H.E., 2011. **Rabbit model of retinoblastoma**. Journal of biomedicine & biotechnology 2011, 394730.

Kaplan, M.H., Hufford, M.M., Olson, M.R., 2015. **The development and in vivo function of T helper 9 cells**. Nat Rev Immunol 15, 295-307.

Kasamatsu, J., 2013. **Evolution of innate and adaptive immune systems in jawless vertebrates**. Microbiology and immunology 57, 1-12.

Katzourakis, A., Tristem, M., Pybus, O.G., Gifford, R.J., 2007. **Discovery and analysis of the first endogenous lentivirus**. Proc Natl Acad Sci U S A 104, 6261-6265.

Kelley, J.L., Swanson, W.J., 2008. **Positive selection in the human genome: from genome scans to biological significance**. Annual review of genomics and human genetics 9, 143-160.

Kerr, P.J., 2012. **Myxomatosis in Australia and Europe: a model for emerging infectious diseases**. Antiviral Res 93, 387-415.

Kerr, P.J., Donnelly, T.M., 2013. **Viral infections of rabbits**. The veterinary clinics of North America. Exotic animal practice 16, 437-468.

Kerr, P.J., Liu, J., Cattadori, I., Ghedin, E., Read, A.F., Holmes, E.C., 2015. **Myxoma virus and the Leporipoxviruses: an evolutionary paradigm**. Viruses 7, 1020-1061.

Kerr, P.J., Perkins, H.D., Inglis, B., Stagg, R., McLaughlin, E., Collins, S.V., Van Leeuwen, B.H., 2004. **Expression of rabbit IL-4 by recombinant myxoma viruses enhances virulence and overcomes genetic resistance to myxomatosis**. Virology 324, 117-128.

Kimura, M., 1977. **Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution**. Nature 267, 275-276.

Kosiol, C., Vinar, T., da Fonseca, R.R., Hubisz, M.J., Bustamante, C.D., Nielsen, R., Siepel, A., 2008. **Patterns of positive selection in six Mammalian genomes**. PLoS genetics 4, e1000144.

Koyanagi, M., Kerns, J.A., Chung, L., Zhang, Y., Brown, S., Moldoveanu, T., Malik, H.S., Bix, M., 2010. **Diversifying selection and functional analysis of interleukin-4 suggests antagonism-driven evolution at receptor-binding interfaces**. BMC evolutionary biology 10, 223.

Kundu, M., Basu, J., 2006. **IL-32: an emerging player in the immune response network against tuberculosis?** PLoS Med 3, e274.

Laliena, A., San Miguel, B., Crespo, I., Alvarez, M., Gonzalez-Gallego, J., Tunon, M.J., 2012. **Melatonin attenuates inflammation and promotes regeneration in rabbits with fulminant hepatitis of viral origin**. J Pineal Res 53, 270-278.

Langrish, C.L., McKenzie, B.S., Wilson, N.J., de Waal Malefyt, R., Kastelein, R.A., Cua, D.J., 2004. **IL-12 and IL-23: master regulators of innate and adaptive immunity**. Immunol Rev 202, 96-105.

Le Gall-Recule, G., Zwingelstein, F., Boucher, S., Le Normand, B., Plassiart, G., Portejoie, Y., Decors, A., Bertagnoli, S., Guerin, J.L., Marchandeau, S., 2011. **Detection of a new variant of rabbit haemorrhagic disease virus in France**. Vet Rec 168, 137-138.

Levasseur, A., Pontarotti, P., 2011. **The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics**. Biology direct 6, 11.

Levine, S.J., Wenzel, S.E., 2010. **Narrative review: the role of Th2 immune pathway modulation in the treatment of severe asthma and its phenotypes**. Annals of internal medicine 152, 232-237.

Liao, W., Lin, J.X., Leonard, W.J., 2013. **Interleukin-2 at the crossroads of effector responses, tolerance, and immunotherapy**. Immunity 38, 13-25.

Lissovsky, A.A., 2014. **Taxonomic revision of pikas Ochotona (Lagomorpha, Mammalia) at the species level**. Mammalia 78, 199-216.

Liu, J., Wennier, S., Reinhard, M., Roy, E., MacNeill, A., McFadden, G., 2009. **Myxoma virus expressing interleukin-15 fails to cause lethal myxomatosis in European rabbits**. J Virol 83, 5933-5938.

Liu, S.J., Xue, H.P., Pu, B.Q., Qian, N.H., 1984. **A new viral disease in rabbit**. Anim Husb Vet Med 16, 253-255.

Lopes, A.M., Correia, J., Abrantes, J., Melo, P., Ramada, M., Magalhaes, M.J., Alves, P.C., Esteves, P.J., 2014. **Is the new variant RHDV replacing genogroup 1 in Portuguese wild rabbit populations?** Viruses 7, 27-36.

Lopes, A.M., Dalton, K.P., Magalhaes, M.J., Parra, F., Esteves, P.J., Holmes, E.C., Abrantes, J., 2015. **Full genomic analysis of new variant rabbit hemorrhagic disease virus revealed multiple recombination events**. The Journal of general virology 96, 1309-1319.

Lopez-Martinez, N., 2008. **The Lagomorph fossil record and the origin of the European rabbit**, in: Alves, P.C., Ferrand, N., Hackländer, K. (Eds.), Lagomorph Biology: Evolution, Ecology and Conservation. Springer, pp. 27-46.

Luzina, I.G., Keegan, A.D., Heller, N.M., Rook, G.A., Shea-Donohue, T., Atamas, S.P., 2012. **Regulation of inflammation by interleukin-4: a review of "alternatives"**. J Leukoc Biol 92, 753-764.

Ma, H., Zheng, L., Liu, Y., Zhao, C., Harrison, T.J., Ma, Y., Sun, S., Zhang, J., Wang, Y., 2010. **Experimental infection of rabbits with rabbit and genotypes 1 and 4 hepatitis E viruses**. PLoS One 5, e9160.

Malek, T.R., 2008. **The biology of interleukin-2**. Annu Rev Immunol 26, 453-479.

Marchandeau, S., Pontier, D., Guitton, J.S., Letty, J., Fouchet, D., Aubineau, J., Berger, F., Leonard, Y., Roobrouck, A., Gelfi, J., Peralta, B., Bertagnoli, S., 2014. **Early infections by myxoma virus of young rabbits (Oryctolagus cuniculus) protected by maternal antibodies activate their immune system and enhance herd immunity in wild populations**. Vet Res 45, 26.

Marques, R.M., Costa, E.S.A., Aguas, A.P., Teixeira, L., Ferreira, P.G., 2012. **Early inflammatory response of young rabbits attending natural resistance to calicivirus (RHDV) infection**. Veterinary immunology and immunopathology 150, 181-188.

Marques, R.M., Teixeira, L., Aguas, A.P., Ribeiro, J.C., Costa-e-Silva, A., Ferreira, P.G., 2014. **Immunosuppression abrogates resistance of young rabbits to Rabbit Haemorrhagic Disease (RHD)**. Vet Res 45, 14.

Marshall, I.D., Regnery, D.C., 1960. **Myxomatosis in a California brush rabbit (*Sylyilagus bachmani*)**. Nature 188, 73-74.

Martin, N.T., Martin, M.U., 2016. **Interleukin 33 is a guardian of barriers and a local alarmin**. Nat Immunol 17, 122-131.

Masteller, E.L., Wong, B.R., 2014. **Targeting IL-34 in chronic inflammation**. Drug discovery today 19, 1212-1216.

Matthee, C.A., van Vuuren, B.J., Bell, D., Robinson, T.J., 2004. **A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene**. Syst Biol 53, 433-447.

McNab, F., Mayer-Barber, K., Sher, A., Wack, A., O'Garra, A., 2015. **Type I interferons in infectious disease**. Nat Rev Immunol 15, 87-103.

Melo-Ferreira, J., Lemos de Matos, A., Areal, H., Lissovsky, A.A., Carneiro, M., Esteves, P.J., 2015. **The phylogeny of pikas (Ochotona) inferred from a multilocus coalescent approach**. Mol Phylogenet Evol 84, 240-244.

Mikami, O., Park, J.H., Kimura, T., Ochiai, K., Itakura, C., 1999. **Hepatic lesions in young rabbits experimentally infected with rabbit haemorrhagic disease virus**. Res Vet Sci 66, 237-242.

Miller, A.M., 2011. **Role of IL-33 in inflammation and disease**. J Inflamm (Lond) 8, 22.

Monnerot, M., Vigne, J.D., Biju-Duval, C., Casane, D., Callou, C., Hardy, C., Mougel, F., Soriguer, R., Dennebouy, N., Mounolou, J.C., 1994. **Rabbit and man: genetic and historic approach**. Genet Sel Evolution 26, 167-182.

Monteleone, G., Pallone, F., Macdonald, T.T., 2010. **Interleukin-25: a two-edged sword in the control of immune-inflammatory responses**. Cytokine Growth Factor Rev 21, 471-475.

Monterroso, P., Garrote, G., Serronha, A., Santos, E., Delibes-Mateos, M., Abrantes, J., Perez de Ayala, R., Silvestre, F., Carvalho, J., Vasco, I., Lopes, A.M., Maio, E., Magalhaes, M.J., Mills, L.S., Esteves, P.J., Simon, M.A., Alves, P.C., 2016. **Disease-mediated bottom-up regulation: An emergent virus affects a keystone prey, and alters the dynamics of trophic webs**. Scientific reports 6, 36072.

Morisse, J.P., Le Gall, G., Boilletot, E., 1991. **Hepatitis of viral origin in Leporidae: introduction and aetiological hypotheses**. Rev Sci Tech 10, 269-310.

Mullane, K., Williams, M., 2014. **Animal models of asthma: reprise or reboot?** Biochemical pharmacology 87, 131-139.

Munch, J., Standker, L., Forssmann, W.G., Kirchhoff, F., 2014. **Discovery of modulators of HIV-1 infection from the human peptidome**. Nature reviews. Microbiology 12, 715-722.

Muñoz Goyanes, G., 1960. **Anverso y reverso de la mixomatosis**, Madrid,.

Muraille, E., 2013. **Redefining the immune system as a social interface for cooperative processes**. PLoS pathogens 9, e1003203.

Nei, M., Rooney, A.P., 2005. **Concerted and birth-and-death evolution of multigene families**. Annual review of genetics 39, 121-152.

Netea, M.G., Joosten, L.A., Latz, E., Mills, K.H., Natoli, G., Stunnenberg, H.G., O'Neill, L.A., Xavier, R.J., 2016. **Trained immunity: A program of innate immune memory in health and disease**. Science 352, aaf1098.

Netea, M.G., Quintin, J., van der Meer, J.W., 2011. **Trained immunity: a memory for innate host defense**. Cell host & microbe 9, 355-361.

Neurath, M.F., Finotto, S., 2011. **IL-6 signaling in autoimmunity, chronic inflammation and inflammation-associated cancer**. Cytokine Growth Factor Rev 22, 83-89.

Neves, F., Abrantes, J., Almeida, T., de Matos, A.L., Costa, P.P., Esteves, P.J., 2015. **Genetic characterization of interleukins (IL-1alpha, IL-1beta, IL-2, IL-4, IL-8, IL-10, IL-12A, IL-12B, IL-15 and IL-18) with relevant biological roles in lagomorphs**. Innate immunity 21, 787-801.

Newton, K., Dixit, V.M., 2012. **Signaling in innate immunity and inflammation**. Cold Spring Harbor perspectives in biology 4.

Nishimoto, N., Kishimoto, T., 2006. **Interleukin 6: from bench to bedside**. Nature clinical practice. Rheumatology 2, 619-626.

Noelle, R.J., Nowak, E.C., 2010. **Cellular sources and immune functions of interleukin-9**. Nat Rev Immunol 10, 683-687.

Nomiyama, H., Hieshima, K., Osada, N., Kato-Unoki, Y., Otsuka-Ono, K., Takegawa, S., Izawa, T., Yoshizawa, A., Kikuchi, Y., Tanase, S., Miura, R., Kusuda, J., Nakao, M., Yoshie, O., 2008. **Extensive expansion and diversification of the chemokine gene family in zebrafish: identification of a novel chemokine subfamily CX**. BMC genomics 9, 222.

Nomiyama, H., Osada, N., Yoshie, O., 2013. **Systematic classification of vertebrate chemokines based on conserved synteny and evolutionary history**. Genes to cells : devoted to molecular & cellular mechanisms 18, 1-16.

O'Connell, M.J., McInerney, J.O., 2005. **Gamma chain receptor interleukins: evidence for positive selection driving the evolution of cell-to-cell communicators in the mammalian immune system**. Journal of molecular evolution 61, 608-619.

Ohno, S., 1970. **Evolution by gene duplication**. Allen & Unwin; Springer-Verlag, London, New York,.

Olson, B.M., Sullivan, J.A., Burlingham, W.J., 2013. **Interleukin 35: a key mediator of suppression and the propagation of infectious tolerance**. Front Immunol 4, 315.

Onishi, R.M., Gaffen, S.L., 2010. **Interleukin-17 and its target genes: mechanisms of interleukin-17 function in disease**. Immunology 129, 311-321.

Palmer, G., Gabay, C., 2011. **Interleukin-33 biology with potential insights into human diseases**. Nature reviews. Rheumatology 7, 321-329.

Pan, R., Sampson, J.M., Chen, Y., Vaine, M., Wang, S., Lu, S., Kong, X.P., 2013. **Rabbit anti-HIV-1 monoclonal antibodies raised by immunization can mimic the antigen-binding modes of antibodies derived from HIV-1-infected humans**. J Virol 87, 10221-10231.

Park N Y, Chong C Y, Kim J H, Cho S M, Cha Y H, Jung B T, Kim D S, Yoon J B, 1987. **An outbreak of viral haemorrhagic pneumonia (tentative name) of rabbits in Korea**. J Korean Vet Med Assoc 23, 603-610.

Pasteur, L., 1885. **Méthode pour prévenir la rage après morsure**. Comptes rendus hebdomadaires des séances de l'Académie des sciences (Paris) 101, 765-774.

Patterson, M.F., Borish, L., Kennedy, J.L., 2015. **The past, present, and future of monoclonal antibodies to IL-5 and eosinophilic asthma: a review**. Journal of asthma and allergy 8, 125-134.

Paul, W.E., 2015. **History of interleukin-4**. Cytokine 75, 3-7.

Peng, X., Knouse, J.A., Hernon, K.M., 2015. **Rabbit Models for Studying Human Infectious Diseases**. Comparative medicine 65, 499-507.

Perelygin, A.A., Zharkikh, A.A., Astakhova, N.M., Lear, T.L., Brinton, M.A., 2008. **Concerted evolution of vertebrate CCR2 and CCR5 genes and the origin of a recombinant equine CCR5/2 gene**. J Hered 99, 500-511.

Perkins, H.D., van Leeuwen, B.H., Hardy, C.M., Kerr, P.J., 2000. **The complete cDNA sequences of IL-2, IL-4, IL-6 AND IL-10 from the European rabbit (Oryctolagus cuniculus)**. Cytokine 12, 555-565.

Pillai, M.R., Bix, M., 2011. **Evolution of IL4 and pathogen antagonism**. Growth Factors 29, 153-160.

Pinheiro, A., Lanning, D., Alves, P.C., Mage, R.G., Knight, K.L., van der Loo, W., Esteves, P.J., 2011. **Molecular bases of genetic diversity and evolution of the immunoglobulin heavy chain variable region (IGHV) gene locus in leporids**. Immunogenetics 63, 397-408.

Pinheiro, A., Neves, F., Lemos de Matos, A., Abrantes, J., van der Loo, W., Mage, R., Esteves, P.J., 2016. **An overview of the lagomorph immune system and its genetic diversity**. Immunogenetics 68, 83-107.

Prieto, J.M., Fernandez, F., Alvarez, V., Espi, A., Garcia Marin, J.F., Alvarez, M., Martin, J.M., Parra, F., 2000. **Immunohistochemical localisation of rabbit haemorrhagic disease virus VP-60 antigen in early infection of young and adult rabbits**. Res Vet Sci 68, 181-187.

Qazi, B.S., Tang, K., Qazi, A., 2011. **Recent advances in underlying pathologies provide insight into interleukin-8 expression-mediated inflammation and angiogenesis**. International journal of inflammation 2011, 908468.

Queney, G., Ferrand, N., Weiss, S., Mougel, F., Monnerot, M., 2001. **Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (Oryctolagus cuniculus)**. Mol Biol Evol 18, 2169-2178.

Ratcliffe, F.N., Myers, K., Fennessy, B.V., Calaby, J.H., 1952. **Myxomatosis in Australia; a step towards the biological control of the rabbit**. Nature 170, 7-11.

Richmond, J., Tuzova, M., Cruikshank, W., Center, D., 2014. **Regulation of cellular processes by interleukin-16 in homeostasis and cancer**. Journal of cellular physiology 229, 139-147.

Rutz, S., Wang, X., Ouyang, W., 2014. **The IL-20 subfamily of cytokines--from host defence to tissue homeostasis**. Nat Rev Immunol 14, 783-795.

Sabat, R., 2010. **IL-10 family of cytokines**. Cytokine & growth factor reviews 21, 315-324.

Sabat, R., Grutz, G., Warszawska, K., Kirsch, S., Witte, E., Wolk, K., Geginat, J., 2010. **Biology of interleukin-10**. Cytokine Growth Factor Rev 21, 331-344.

San-Miguel, B., Alvarez, M., Culebras, J.M., Gonzalez-Gallego, J., Tunon, M.J., 2006. **N-acetyl-cysteine protects liver from apoptotic death in an animal model of fulminant hepatic failure**. Apoptosis 11, 1945-1957.

San-Miguel, B., Crespo, I., Vallejo, D., Alvarez, M., Prieto, J., Gonzalez-Gallego, J., Tunon, M.J., 2014. **Melatonin modulates the autophagic response in acute liver failure induced by the rabbit hemorrhagic disease virus**. J Pineal Res 56, 313-321.

Sanarelli, G., 1898. **Das myxomatogene Virus. Beitrag zum Stadium der Krankheitserreger ausserhalb des Sichtbaren.** . Zbl. Bakt. 23, 865-873.

Saraiva, M., O'Garra, A., 2010. **The regulation of IL-10 production by immune cells**. Nat Rev Immunol 10, 170-181.

Sawant, D.V., Hamilton, K., Vignali, D.A., 2015. **Interleukin-35: Expanding Its Job Profile**. J Interferon Cytokine Res 35, 499-512.

Scheller, J., Chalaris, A., Schmidt-Arras, D., Rose-John, S., 2011. **The pro- and anti-inflammatory properties of the cytokine interleukin-6**. Biochim Biophys Acta 1813, 878-888.

Schneider, W.M., Chevillotte, M.D., Rice, C.M., 2014. **Interferon-stimulated genes: a complex web of host defenses**. Annu Rev Immunol 32, 513-545.

Schnupf, P., Sansonetti, P.J., 2012. **Quantitative RT-PCR profiling of the rabbit immune response: assessment of acute Shigella flexneri infection**. PLoS One 7, e36446.

Schrader, J.W., 2003. **Interleukin is as interleukin does**. J Immunol Methods 276, 1-3.

Sedger, L.M., McDermott, M.F., 2014. **TNF and TNF-receptors: From mediators of cell death and inflammation to therapeutic giants - past, present and future**. Cytokine Growth Factor Rev 25, 453-472.

Seixas, F.A., Juste, J., Campos, P.F., Carneiro, M., Ferrand, N., Alves, P.C., Melo-Ferreira, J., 2014. **Colonization history of Mallorca Island by the European rabbit, *Oryctolagus cuniculus*, and the Iberian hare, *Lepus granatensis* (Lagomorpha: Leporidae)**. Biological Journal of the Linnean Society 111, 748-760.

Seok, J., Warren, H.S., Cuenca, A.G., Mindrinos, M.N., Baker, H.V., Xu, W., Richards, D.R., McDonald-Smith, G.P., Gao, H., Hennessy, L., Finnerty, C.C., Lopez, C.M., Honari, S., Moore, E.E., Minei, J.P., Cuschieri, J., Bankey, P.E., Johnson, J.L., Sperry, J., Nathens, A.B., Billiar, T.R., West, M.A., Jeschke, M.G., Klein, M.B., Gamelli, R.L., Gibran, N.S., Brownstein, B.H., Miller-Graziano, C., Calvano, S.E., Mason, P.H., Cobb, J.P., Rahme, L.G., Lowry, S.F., Maier, R.V., Moldawer, L.L., Herndon, D.N., Davis, R.W., Xiao, W., Tompkins, R.G., 2013. **Genomic responses in mouse models poorly mimic human inflammatory diseases**. Proc Natl Acad Sci U S A 110, 3507-3512.

Sharples, C.M., Fa, J.E., Bell, D.J., 1996. **Geographical variation in size in the European rabbit *Oryctolagus cuniculus* (Lagomorpha: Leporidae) in western Europe and North Africa**. Zoological Journal of the Linnean Society 117, 141-158.

Shen, L., Zhang, C., Wang, T., Brooks, S., Ford, R.J., Lin-Lee, Y.C., Kasianowicz, A., Kumar, V., Martin, L., Liang, P., Cowell, J., Ambrus, J.L., Jr., 2006. **Development of autoimmunity in IL-14alpha-transgenic mice**. Journal of immunology 177, 5676-5686.

Shibata, K., Nomiyama, H., Yoshie, O., Tanase, S., 2013. **Genome diversification mechanism of rodent and Lagomorpha chemokine genes**. Biomed Res Int 2013, 856265.

Siewe, B.T., Kalis, S.L., Esteves, P.J., Zhou, T., Knight, K.L., 2010. **A novel functional rabbit IL-7 isoform**. Dev Comp Immunol 34, 828-836.

Silva, E., Marques, S., Osorio, H., Carvalheira, J., Thompson, G., 2012. **Endogenous hepatitis C virus homolog fragments in European rabbit and hare genomes replicate in cell culture**. PLoS One 7, e49820.

Silva, E., Osorio, H., Thompson, G., 2015. **Hepatitis C-like viruses are produced in cells from rabbit and hare DNA**. Scientific reports 5, 14535.

Sims, J.E., Smith, D.E., 2010. **The IL-1 family: regulators of immunity**. Nat Rev Immunol 10, 89-102.

Skyberg, J.A., Rollins, M.F., Samuel, J.W., Sutherland, M.D., Belisle, J.T., Pascual, D.W., 2013. **Interleukin-17 protects against the Francisella tularensis live vaccine strain but not against a virulent F. tularensis type A strain**. Infection and immunity 81, 3099-3105.

Smith, A.T., 2008. **The world of pikas**, in: Alves, P.C., Ferrand, N., Hackländer, K. (Eds.), Lagomorph Biology: Evolution, Ecology and Conservation. Springer, pp. 89-102.

Spolski, R., Leonard, W.J., 2008. **Interleukin-21: basic biology and implications for cancer and autoimmunity**. Annu Rev Immunol 26, 57-79.

Spolski, R., Leonard, W.J., 2014. **Interleukin-21: a double-edged sword with therapeutic potential**. Nature reviews. Drug discovery 13, 379-395.

Stanford, M.M., Barrett, J.W., Gilbert, P.A., Bankert, R., McFadden, G., 2007. **Myxoma virus expressing human interleukin-12 does not induce myxomatosis in European rabbits**. J Virol 81, 12704-12708.

Stanford, M.M., McFadden, G., 2005. **The 'supervirus'? Lessons from IL-4-expressing poxviruses**. Trends in immunology 26, 339-345.

Stonier, S.W., Schluns, K.S., 2010. **Trans-presentation: a novel mechanism regulating IL-15 delivery and responses**. Immunology letters 127, 85-92.

Suckow, M.A., 2012. **The laboratory rabbit, guinea pig, hamster, and other rodents**, 1st ed. Academic Press, an imprint of Elsevier, Amsterdam ; Boston.

Sun, J.C., Ugolini, S., Vivier, E., 2014. **Immunological memory within the innate immune system**. The EMBO journal 33, 1295-1303.

Surridge, A.K., van der Loo, W., Abrantes, J., Carneiro, M., Hewitt, G.M., Esteves, P.J., 2008. **Diversity and evolutionary history of the MHC DQA gene in leporids**. Immunogenetics 60, 515-525.

Takao, K., Miyakawa, T., 2015. **Genomic responses in mouse models greatly mimic human inflammatory diseases**. Proc Natl Acad Sci U S A 112, 1167-1172.

Teng, M.W., Bowman, E.P., McElwee, J.J., Smyth, M.J., Casanova, J.L., Cooper, A.M., Cua, D.J., 2015. **IL-12 and IL-23 cytokines: from discovery to targeted therapies for immune-mediated inflammatory diseases**. Nature medicine 21, 719-729.

Tengvall, S., Che, K.F., Linden, A., 2016. **Interleukin-26: An Emerging Player in Host Defense and Inflammation**. Journal of innate immunity 8, 15-22.

Tervo, H.M., Keppler, O.T., 2010. **High natural permissivity of primary rabbit cells for HIV-1, with a virion infectivity defect in macrophages as the final replication barrier**. J Virol 84, 12300-12314.

Thaiss, C.A., Levy, M., Itav, S., Elinav, E., 2016. **Integration of Innate Immune Signaling**. Trends in immunology.

Tian, J., Hu, S., Sun, Y., Ban, X., Yu, H., Dong, N., Wu, J., Yu, B., 2012. **A novel model of atherosclerosis in rabbits using injury to arterial walls induced by ferric chloride as evaluated by optical coherence tomography as well as intravascular ultrasound and histology**. Journal of biomedicine & biotechnology 2012, 121867.

Tomaki, M., Zhao, L.L., Sjostrand, M., Linden, A., Ichinose, M., Lotvall, J., 2002. **Comparison of effects of anti-IL-3, IL-5 and GM-CSF treatments on eosinophilopoiesis and airway eosinophilia induced by allergen**. Pulmonary pharmacology & therapeutics 15, 161-168.

Tosic, V., Thomas, D.L., Kranz, D.M., Liu, J., McFadden, G., Shisler, J.L., MacNeill, A.L., Roy, E.J., 2014. **Myxoma virus expressing a fusion protein of interleukin-15 (IL15) and IL15 receptor alpha has enhanced antitumor activity**. PLoS One 9, e109801.

Trzeciak-Ryczek, A., Tokarz-Deptuła, B., Deptuła, W., 2016. **Expression of IL-1β, IL-2, IL-10, TNF-β and GM-CSF in peripheral blood leukocytes of rabbits experimentally infected with rabbit haemorrhagic disease virus** Veterinary Microbiology 186, 71-81.

Tunon, M.J., San-Miguel, B., Crespo, I., Laliena, A., Vallejo, D., Alvarez, M., Prieto, J., Gonzalez-Gallego, J., 2013. **Melatonin treatment reduces endoplasmic reticulum stress and modulates the unfolded protein response in rabbits with lethal fulminant hepatitis of viral origin**. J Pineal Res 55, 221-228.

Tunon, M.J., San Miguel, B., Crespo, I., Jorquera, F., Santamaria, E., Alvarez, M., Prieto, J., Gonzalez-Gallego, J., 2011a. **Melatonin attenuates apoptotic liver damage in fulminant hepatic failure induced by the rabbit hemorrhagic disease virus**. J Pineal Res 50, 38-45.

Tunon, M.J., San Miguel, B., Crespo, I., Riezu-Boj, J.I., Larrea, E., Alvarez, M., Gonzalez, I., Bustos, M., Gonzalez-Gallego, J., Prieto, J., 2011b. **Cardiotrophin-1 promotes a high survival rate in rabbits with lethal fulminant hepatitis of viral origin**. J Virol 85, 13124-13132.

Turvey, S.E., Broide, D.H., 2010. **Innate immunity**. J Allergy Clin Immunol 125, S24-32.

Vallejo, D., Crespo, I., San-Miguel, B., Alvarez, M., Prieto, J., Tunon, M.J., Gonzalez-Gallego, J., 2014. **Autophagic response in the Rabbit Hemorrhagic Disease, an animal model of virally-induced fulminant hepatic failure**. Vet Res 45, 15.

Vallender, E.J., Lahn, B.T., 2004. **Positive selection on the human genome**. Human molecular genetics 13 Spec No 2, R245-254.

van der Loo, W., Afonso, S., de Matos, A.L., Abrantes, J., Esteves, P.J., 2012. **Pseudogenization of the MCP-2/CCL8 chemokine gene in European rabbit (genus *Oryctolagus*), but not in species of Cottontail rabbit (*Sylvilagus*) and Hare (*Lepus*)**. BMC Genet 13, 72.

van der Loo, W., Magalhaes, M.J., de Matos, A.L., Abrantes, J., Yamada, F., Esteves, P.J., 2016. **Adaptive Gene Loss? Tracing Back the Pseudogenization of the Rabbit CCL8 Chemokine**. J Mol Evol 83, 12-25.

van der Loo, W., Mougel, F., Sanchez, M.S., Bouton, C., Castien, E., Fonseca, A., Ferrand, N., Soriguer, R., Monnerot, M., 1999. **Cytonuclear disequilibria in wild populations of rabbit (Oryctolagus cuniculus L.) suggest unequal allele turnover rates at the b locus (IGKC1)**. Immunogenetics 49, 629-643.

van der Meer, J.W., Joosten, L.A., Riksen, N., Netea, M.G., 2015. **Trained immunity: A smart way to enhance innate immune defence**. Molecular immunology 68, 40-44.

Van Valen L., 1973. **A new evolutionary law**. Evolutionary Theory 1, 1-30.

Vande Walle, L., Lamkanfi, M., 2011. **Inflammasomes: caspase-1-activating platforms with critical roles in host defense**. Frontiers in microbiology 2, 3.

Vaughan, T.A., Ryan, J.M., Czaplewski, N.J., 2011. **Mammalogy**, 5th ed. Jones and Bartlett Publishers, Sudbury, Mass.

Vazquez-Salat, N., Yuhki, N., Beck, T., O'Brien, S.J., Murphy, W.J., 2007. **Gene conversion between mammalian CCR2 and CCR5 chemokine receptor genes: a potential mechanism for receptor dimerization**. Genomics 90, 213-224.

Villafuerte, R., 2002. **Atlas de los Mamíferos Terrestres de España**, in: Palomo, L.J., Gisbert, J. (Eds.), Oryctolagus cuniculus. Dirección General de Conservación de la Naturaleza-SECEM-SECEMU, Madrid, pp. 464-466.

Watford, W.T., Hissong, B.D., Bream, J.H., Kanno, Y., Muul, L., O'Shea, J.J., 2004. **Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4**. Immunol Rev 202, 139-156.

Webb, D.R., 2014. **Animal models of human disease: inflammation**. Biochemical pharmacology 87, 121-130.

Wilen, C.B., Tilton, J.C., Doms, R.W., 2012. **HIV: cell binding and entry**. Cold Spring Harbor perspectives in medicine 2.

Wills-Karp, M., 2004. **Interleukin-13 in asthma pathogenesis**. Immunol Rev 202, 175-190.

Wilson, K.C., Center, D.M., Cruikshank, W.W., 2004. **The effect of interleukin-16 and its precursor on T lymphocyte activation and growth**. Growth Factors 22, 97-104.

Witowski, J., Ksiazek, K., Jorres, A., 2004. **Interleukin-17: a mediator of inflammatory responses**. Cellular and molecular life sciences : CMLS 61, 567-579.

Woodruff-Pak, D.S., Agelan, A., Del Valle, L., 2007. **A rabbit model of Alzheimer's disease: valid at neuropathological, cognitive, and therapeutic levels**. Journal of Alzheimer's disease : JAD 11, 371-383.

Xu, W.Y., 1991. **Viral haemorrhagic disease of rabbits in the People's Republic of China: epidemiology and virus characterisation**. Rev Sci Tech 10, 393-408.

Yap, M.W., Stoye, J.P., 2013. **Apparent effect of rabbit endogenous lentivirus type K acquisition on retrovirus restriction by lagomorph Trim5alphas**. Philos Trans R Soc Lond B Biol Sci 368, 20120498.

Yoshida, H., Hunter, C.A., 2015. **The immunobiology of interleukin-27**. Annu Rev Immunol 33, 417-443.

Yuan, X., Peng, X., Li, Y., Li, M., 2015. **Role of IL-38 and its related cytokines in inflammation**. Mediators of inflammation 2015, 807976.

Zelus, D., Robinson-Rechavi, M., Delacre, M., Auriault, C., Laudet, V., 2000. **Fast evolution of interleukin-2 in mammals and positive selection in ruminants**. Journal of molecular evolution 51, 234-244.

Zhang, J., 2003. **Evolution by gene duplication: an update**. TRENDS in Ecology and Evolution 18, 292-298.

Zhang, J., Nei, M., 2000. **Positive selection in the evolution of mammalian interleukin-2 genes**. Molecular biology and evolution 17, 1413-1416.

Zhang, Q., Putheti, P., Zhou, Q., Liu, Q., Gao, W., 2008. **Structures and biological functions of IL-31 and IL-31 receptors**. Cytokine & growth factor reviews 19, 347-356.

Zhou, Y., Zhu, Y., 2015. **Important Role of the IL-32 Inflammatory Network in the Host Response against Viral Infection**. Viruses 7, 3116-3129.

Zhu, H., Liu, C., 2003. **Interleukin-1 inhibits hepatitis C virus subgenomic RNA replication by activation of extracellular regulated kinase pathway**. J Virol 77, 5493-5498.

Zhu, J., 2015. **T helper 2 (Th2) cell differentiation, type 2 innate lymphoid cell (ILC2) development and regulation of interleukin-4 (IL-4) and IL-13 production**. Cytokine 75, 14-24.

Zlotnik, A., Yoshie, O., 2012. **The chemokine superfamily revisited**. Immunity 36, 705-716.

Zlotnik, A., Yoshie, O., Nomiyama, H., 2006. **The chemokine and chemokine receptor superfamilies and their molecular evolution**. Genome biology 7, 243.

CHAPTER 2

Aim and Objectives

Knowledge on interleukins and chemokine ligands in lagomorphs has been mostly restricted to the European rabbit, an animal model for the study of speciation, domestication, medical immunology research, virology and diseases. To fill in this gap in the Lagomorpha order, we performed a genomic characterization and studied the genetic diversity and evolution of interleukins and chemokine ligands with biological relevance. Thus, in order to unveil the scenarios driving the evolutionary patterns of these genes, the sequences obtained in this work were compared with other mammalian orthologous sequences.

This thesis is organized into five chapters and includes seven scientific papers, six already published in journals indexed in the Science Citation Index (SCI) and one in preparation.

In the first chapter, entitled *Chapter 1. General Introduction,* an overview of the lagomorph's phylogeny is given and the basic principles of innate immune system are described, with a main focus in the interleukins and chemokine ligands. Furthermore, the state of the art on these proteins in lagomorphs is revisited. Finally, the host-pathogen co-evolution theory is discussed. In addition this chapter describes the natural history of the European rabbit populations and highlights the importance of the European rabbit as an animal model. Indeed, this species has been used to study speciation events and to decipher the genes responsible for the genetic traits associated with the domestication process. The European rabbit is present in more than 800 islands with contrasting climatic conditions, making it a perfect model to study adaptations to new environments. Until the 1980s, the European rabbit was widely used as an animal model for human research, but it has been replaced by mouse (*Mus musculus*). However, the uniqueness of the immune system of the European rabbit makes it a valuable model for several fields, particularly for the study of human diseases, basic immunology and virology.

This *Chapter 2. Objectives* presents a brief description of all the chapters that compose this thesis as well as the pappers published.

Interleukins (ILs) are crucial in the regulation and determination of immune responses. Population genetic analyses and sequence analyses have been shown that ILs are evolving under positive selection. However, the amino acids presenting such signatures are almost unknown. In order to identify codons under positive selection in mammalian interleukins we used different codon-based maximum-likelihood (ML) approaches (paper II). Our results showed that the European rabbit IL6 gene differs from the other mammal species by having a 27 base pair extension. To infer the evolution of this gene in lagomorphs, we extended the genetic characterization to eight more lagomorph species. Additionally, it has been suggested that IL1, IL2, IL4, IL6, IL8, IL10, IL12, IL15 and IL18 are associated with rabbit susceptibility to rabbit hemorrhagic disease (RHD) and myxomatosis, and that IL17A might be related with Tularaemia. However, the genetic characterization of these genes in lagomorphs was limited. In order to improve the current knowledge, we sequenced these genes for seven lagomorph species. The resulting scientific manuscripts are compiled in *Chapter 3. Innate Immunity – Evolution of Interleukins in lagomorphs* in four scientific manuscripts:

Paper I. Neves F, Abrantes J, Steinke JW, Esteves PJ. (2014) Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals. Innate Immunity, 20(2): 184–191.

Paper II. Neves F, Abrantes J, Pinheiro A, Almeida T, Costa PP, Esteves PJ. (2014) Convergent evolution of IL-6 in two leporids (*Oryctolagus* and *Pentalagus*) originated an extended protein. Immunogenetics, 66(9):589–595.

Paper III. Neves F, Abrantes J, Almeida T, Costa PP, Esteves PJ. (2015) Evolutionary Insights into IL17A in Lagomorphs. Mediators of Inflammation, vol. 2015, Article ID 367670, 7 pages, 2015.

Paper IV. Neves F, Abrantes J, Almeida T, de Matos AL, Costa PP, Esteves PJ. (2015) Genetic characterization of interleukins (IL-1α, IL-1β, IL-2, IL-4, IL-8, IL-10, IL-12A, IL-12B, IL-15 and IL-18) with relevant biological roles in lagomorphs. Innate Immunity 21(8):787-801.

Previous studies have focused on the genetic variability of the C-C chemokine receptor 5 (CCR5) in lagomorphs showing significantly differences between them. Indeed, the European rabbit, riverine rabbit and Amami rabbit present a gene conversion of a CCR5 motif in the second extracellular loop with a motif that is characteristic of CCR2. Interestingly, this was not observed in the other lagomorphs. This observation impelled us to genetically characterize the CCR5 ligands CCL11, CCL14 and CCL16. The results obtained showed different evolution patterns, described in *Chapter 4. Innate immunity – Genetic aspects of CC motif chemokines in lagomorphs*. The results are presented in three scientific papers:

Paper V. Neves F, Abrantes J, Lissovsky AA, Esteves PJ. (2015) Pseudogenization of CCL14 in the Ochotonidae (pika) family. Innate Immunity 21(6):647-654.

Paper VI. Neves F, Abrantes J, Esteves PJ. (2016) Evolution of CCL11: genetic characterization in lagomorphs and evidence of positive and purifying selection in mammals. Innate immunity 22(5):336-343.

Paper VII. Neves F, Abrantes J, Lopes AM, Magalhães MJ, Esteves PJ. Evolution of CCL16 in Glires (Rodentia and Lagomorphs) shows an unusual random pseudogenization pattern. (*In preparation*)

In *Chapter 5. Final considerations*, the major findings and conclusion of this thesis are discussed and the implications for future research are presented.

CHAPTER 3

Innate Immunity – Evolution of Interleukins in lagomorphs

Paper I. Neves F, Abrantes J, Steinke JW, Esteves PJ. (2014) Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals. Innate Immunity, 20(2): 184–191.

Paper II. Neves F, Abrantes J, Pinheiro A, Almeida T, Costa PP, Esteves PJ. (2014) Convergent evolution of IL-6 in two leporids (*Oryctolagus* and *Pentalagus*) originated an extended protein. Immunogenetics, 66(9):589–595.

Paper III. Neves F, Abrantes J, Almeida T, Costa PP, Esteves PJ. (2015) Evolutionary Insights into IL17A in Lagomorphs. Mediators of Inflammation, vol. 2015, Article ID 367670, 7 pages, 2015.

Paper IV. Neves F, Abrantes J, Almeida T, de Matos AL, Costa PP, Esteves PJ. (2015) Genetic characterization of interleukins (IL-1α, IL-1β, IL-2, IL-4, IL-8, IL-10, IL-12A, IL-12B, IL-15 and IL-18) with relevant biological roles in lagomorphs. Innate Immunity 21(8):787-801.

MAXIMUM-LIKELIHOOD APPROACHES REVEALED SIGNATURES OF POSITIVE SELECTION IN INTERLEUKIN (IL) GENES IN MAMMALS

Fabiana Neves, Joana Abrantes, John W Steinke, Pedro J Esteves

1. ABSTRACT

Interleukins (ILs) are part of the immune system being involved in multiple biological activities. ILs have been shown to be evolving under positive selection, however little information exists regarding which codons are specifically selected. By using different codon-based maximum-likelihood (ML) approaches, signatures of positive selection in mammalian ILs were searched for.

Sequences of 46 ILs were retrieved from publicly available databases of mammalian genomes to detect signatures of positive selection in individual codons. Evolutionary analyses were conducted under two ML frameworks, the HyPhy package implemented in the Data Monkey Web Server and CODEML implemented in PAML.

Signatures of positive selection were found in 28 ILs: IL1A and B, IL2, IL4-IL10, IL12A and B, IL14-IL17A and C, IL18, IL20-IL22, IL25, IL26, IL27B, IL31, IL34, IL36A and G. Codons under positive selection varied between 1 and 15. No evidence of positive selection was detected in IL13, 17B and F, IL19, IL23, IL24, IL27A and IL29.

Most mammalian ILs have sites evolving under positive selection, which may be explained by the multitude of biological processes in which ILs are enrolled. The results obtained raise hypotheses concerning the ILs functions' that should be pursued by using mutagenesis and crystallographic approaches.

Keywords: Interleukins, interleukin receptors, maximum likelihood, positive selection

2. INTRODUCTION

Different cells of the immune system are able to secrete regulatory proteins, namely cytokines (cks) in response to a variety of stimuli. In mammals, cytokines produced as a part of the innate immune response have the ability to influence the extent and nature of the adaptive immunity response and are thus crucial for many aspects of the immune response (Brocker et al., 2010; Kaiser et al., 2004; O'Connell and McInerney, 2005; Poon et al., 2009; Zelus et al., 2000; Zhang and Nei, 2000). When cks are secreted by leukocytes and act on other cells they are called interleukins (ILs) (Akdis et al., 2011; Beadling and Slifka, 2006; Kaiser et al., 2004; Ozaki and Leonard, 2002; Schrader, 2003). Ils are polypeptides of low molecular weight involved in several biological activities, including immunity, inflammation, inflammatory diseases, hematopoiesis, oncogenesis, neurogenesis, fertility, among many others. Each interleukin is normally involved in a multiple biological processes (Afzal et al., 2012; Akdis et al., 2011; Heinrich et al., 2003; Ishihara and Hirano, 2002). Currently, 37 interleukins have been identified in mammals, some with a variable number of variants. Classification of ILs has been based on sequence homology, receptor chain similarities, activity, structural or functional features, leading to complex and intricate classifications (Akdis et al., 2011; Beadling and Slifka, 2006; Borish and Steinke, 2003; Commins et al., 2010; Krause and Pestka, 2005; Leonard and Lin, 2000; Sabat, 2010; Secombes et al., 2011; Steinke and Borish, 2006). The activities of ILs are possible through the binding of these polypeptides to specific cell-surface receptors, components of which can be shared by several interleukins, that are able to transmit intracellular signals through different signaling pathways (Afzal et al., 2012; Fumagalli et al., 2009; Ishihara and Hirano, 2002; O'Connell and McInerney, 2005). Sometimes, for optimal function, interaction between complementary interleukins is necessary (Commins et al., 2008). These interactions can be synergistic, additive or antagonistic and exhibit both negative and positive regulatory effects (Afzal et al.,

2012; Akdis et al., 2011; Commins et al., 2008; Feghali and Wright, 1997; Haddad, 2002).

Owing to the host-pathogen co-evolution, the immune system and its genes are constantly evolving, being under constant pressure and selection for adaptation and where advantageous mutations are highly favored and deleterious mutations are quickly eliminated (Barreiro and Quintana-Murci, 2010; Brocker et al., 2010). ILs such as IL2, IL3, IL4, IL5, IL13, IL23A, IL28A, IL28B and IL29 (Barreiro and Quintana-Murci, 2010; Brocker et al., 2010; Koyanagi et al., 2010; Makalowski and Boguski, 1998; Manry et al., 2011; O'Connell and McInerney, 2005; Pillai and Bix, 2011; Tindall and Hayes, 2010; Zelus et al., 2000; Zhang and Nei, 2000; Zhou et al., 2004) have been identified as some of the immune system genes under positive selection in different mammals. Despite all of these studies, most do not specify the codons positively selected (Makalowski and Boguski, 1998; Manry et al., 2011; Tindall and Hayes, 2010; Zelus et al., 2000; Zhang and Nei, 2000; Zhou et al., 2004), with the exception of IL4 (Koyanagi et al., 2010) in which 15 residues were detected as being under positive selection and located on sites responsible for binding to receptors. Regarding these results, we have extended the search for signatures of positive selection to include mammalian ILs 1-37 by using different maximum-likelihood approaches (ML).

3. MATERIALS AND METHODS

Sequences

The sequences of the mammalian ILs used in the analyses were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/), Ensembl (http://useast.ensembl.org/index.html) and UniProt (http://www.uniprot.org/). For each IL, amino acid residues were numbered from the first human methionine residue, with signal peptides and propeptide amino acids included in the numbering. The number of sequences retrieved for each IL ranged from 14-47 (some of the IL genes were not found in all mammalian genomes) and included representatives of most mammalians groups (e.g. Artiodactyla, Carnivores, Lagomorphs, Primates, Rodents, etc). The identification of the species used for each IL and the accession numbers are listed in Table 3.3 of Supplementary

material. Each of the 37 interleukins was aligned using ClustalW, as implemented in the software program BioEdit version 7.1.3 (Hall, 1999), and adjusted manually.

Codon based analyses of positive selection

According to Poon and collaborators (Poon et al., 2009), and in order to reliably identify codons under positive selection, only interleukins that were represented in at least 10 species were analyzed. Hence, IL17D, IL28 A and B IL32, IL36B and IL37 were not considered. IL3 and IL33 were also excluded since the alignments produced were not reliable and prone to affect the detection of positive selection by leading to false predictions (Majewski and Ott, 2003). Signatures of positive selection are inferred if the ratio of nonsynonymous substitutions per nonsynonymous substitutions site (d_N) over synonymous substitutions per synonymous sites (d_S) is statistically higher than the value observed under neutrality, 1. Here, to detect such signatures in individual codons of mammalian IL sequences, the different d_N/d_S ratios (ω) were compared using two Maximum-Likelihood (ML) frameworks, the HyPhy package implemented in the Data Monkey Web Server (http://www.datamonkey.org/) (Poon et al., 2009) and CODEML implemented in PAML version 4 (Majewski and Ott, 2003; Yang, 1997, 2007), being considered in the analysis the results where d_N/d_S ratios were significantly higher than 1.

In the Data Monkey Web Server, the best fitting nucleotide substitution model was first determined using the automated tool available on the server. Sequences of the IL genes were analyzed under three available models, Single Likelihood Ancestor Counting (SLAC), Fixed-Effect Likelihood (FEL) and Random Effect Likelihood (REL). The SLAC model is based on the reconstruction of ancestral sequences and counts the number of d_S and d_N changes at each codon position of the phylogeny. FEL estimates ratios of d_N to d_S changes for each site in an alignment. REL uses a flexible distribution and allows d_S and d_N to vary across sites independently (Kosakovsky Pond and Frost, 2005; Pond and Frost, 2005; Poon et al., 2009).

In CODEML two opposing models, M7 and M8, were compared using Likelihood Ratio Tests (LRT). While M7 assumes that ω ratios are distributed among sites according to a beta distribution allowing codons to evolve neutrally or

under purifying selection, M8 is an extension of the M7 model with an extra class of sites with an independent ω ratio freely estimated from the data allowing positive selection. M7 vs. M8 were compared by computing twice the difference in the natural logs of the likelihoods (2ΔlnL). The value obtained was used in a likelihood ratio test along with the degrees of freedom (2) and allowed the rejection of the null model ($p<0.05$). Amino acids detected as under positive selection in M8 were identified using the Bayes Empirical Bayes approach (BEB), with posterior probability > 95%. BEB is the preferred approach because it accounts sampling errors in the ML (Bielawski and Yang, 2003; Yang, 1997, 2002, 2007; Yang et al., 2000). For each gene, a neighbor-joining tree was constructed in MEGA5 (Tamura et al., 2011) as working topology with selected options p-distance as the substitution model and complete deletion to gaps/missing data treatment.

For a more conservative approach and as used previously (Areal et al., 2011; Wlasiuk and Nachman, 2010), only sites detected to be under positive selection in more than one ML method were considered.

4. RESULTS

Positive selection was found in 24 of the 46 ILs: IL1 (A and B), IL2, IL4, IL5, IL6, IL7, IL8, IL9, IL10, IL12 (A and B), IL14, IL15, IL16, IL17 (A and C), IL18, IL20, IL21, IL22, IL25, IL26, IL27B, IL31, IL34, IL36 (A and G), as shown in Table 3.1. Codons under positive selection detected by, at least, two ML approaches ranged between 1 and 15, with IL6 being the interleukin with the highest percentage of positively selected codons (2.28% with 15 codons under positive selection) followed by IL7 (2.04% and 11 codons under positive selection). The smallest percentages of positively selected codons were 0.06% in IL14 and 0.10% in IL12B, each with 1 codon. No positively selected codons were detected in IL13, 17B and F, IL19, IL23, IL24, IL27A and IL29.

Table 3. 1.Phylogenetic Tests of Positive Selection[a]

Gene	No. aa residues	No. of species	lnL M7	lnL M8	2ΔlnL	PAML	SLAC[b]	FEL[c]	REL[d]	Total no. of sites	% of sites
IL1 A	825	33	-10987,711	-10939,544	96,335*	104, 122, 132, 147, 148, 173, 174, 175, 178, 180, 190, 227, 230, 248, 258	173, 174	66, 173, 174, 180, 188, 215, 227, 230, 265	54, 99, 104, 116, 122, 125, 127, 129, 132, 139, 158, 162, 173, 174, 175, 179, 180, 215, 227, 230, 248, 250	11	1,333
IL1 B	801	37	-9578,964	-9565,064	27,800*	79, 150, 205	8, 84	8, 36, 71, 84, 93, 96, 104, 129, 153	84	2	0,25
IL2	513	47	-5429,515	-5417,492	24,045*	100,101, 124	-	28, 39, 42, 58, 69	43, 46, 66, 68, 91, 98, 100, 101, 102, 115, 122, 124, 152	3	0,585
IL4	465	44	-6436,977	-6424,860	24,235*	84, 105, 106, 120	84	5, 59, 105, 120, 128, 147	84, 105, 106, 120	4	0,86
IL5	408	36	-4996,583	-4994,044	5,078	-	-	4, 93, 96, 100	100	1	0,245
IL6	657	42	-9824,485	-9805,607	37,755*	46, 55, 62, 65, 75, 120, 141, 145, 149, 158, 161, 162, 177	8, 145	8, 162, 168, 174, 193	30, 38, 46, 55, 56, 62, 75, 76, 96, 113, 120, 141, 145, 149, 153, 158, 161, 162, 168, 171, 173, 174, 177	15	2,283
IL7	540	21	-3887,223	-3861,408	51,630*	47, 51, 54, 60, 92, 102, 113, 126, 153, 172, 174	92, 144	38, 43, 51, 53, 55, 92, 105, 106, 144	30, 38, 46, 47, 51, 53, 54, 55, 92, 101, 105, 106, 113, 143, 144	11	2,037
IL8	297	33	-3256,753	-3252,791	7,924*	71	-	22, 30	22, 30, 64, 71, 74	3	1,010
IL9	423	20	-4237,433	-4217,874	39,118*	5, 24, 25, 84, 86	84	79, 84, 106	5, 20, 22, 24, 25, 28, 31, 35, 49, 79, 82, 84, 86, 90, 98, 106, 131, 136, 140	7	1,655
IL10	534	44	-6188,435	-6184,527	7,815*	26	18	18, 63, 71, 121	18, 20, 26, 35, 63, 68, 71, 121	5	0,936
IL12A	669	27	-5853,066	-5847,159	11,813*	161	-	24, 35, 39, 91, 154	39, 161, 172	2	0,299
IL12B	1032	33	-9173,438	-9170,760	5,357	-	83	83, 147	21, 83, 182, 185, 188, 216, 289, 302, 306	1	0,097
IL14	1698	18	-8069,857	-8069,766	0,183	-	-	501	501	1	0,059
IL15	483	29	-4003,633	-3987,300	32,666*	51, 105, 106, 108	-	4, 59, 116	51, 52, 55, 59, 62, 96, 105, 106, 108, 116, 152	6	1,242

Gene	No. aa residues	No. of species	lnL M7	lnL M8	2ΔlnL	PAML	SLAC[b]	FEL[c]	REL[d]	Total no. of sites	
IL16	4095	17	-27560,331	-27552,931	14,800*	622	335	118, 188, 335, 633, 708, 734, 745, 801, 930, 1075	118, 188, 335, 705, 745, 801, 930, 1075	8	0,195
IL17A	483	27	-5623,567	-5618,970	9,193*	-	18, 98	18, 98	-	2	0,414
IL17C	600	18	-4833,872	-4829,855	8,034*	1	11, 17	11, 17	11, 17, 55	2	0,333
IL18	588	24	-4474,258	-4458,874	30,768*	4, 5, 29, 45, 75, 105, 176, 192	180	6, 32, 37, 45, 159, 180	4, 6, 45, 75, 82, 105, 180, 192	7	1,19
IL20	525	25	-4976,365	-4974,835	3,060	-	-	56, 62	56, 62	2	0,381
IL21	456	19	-3201,764	-3191,105	21,319*	98, 119	86, 112	86, 112	86, 98, 112, 119	4	0,877
IL22	543	25	-5167,215	-5159,805	14,821*	86, 173	53	8, 53, 84	3, 8, 53, 65, 84	3	0,552
IL25	510	20	-4452,843	-4441,182	23,323*	1, 2, 137	-	53, 137	36, 40, 55	1	0,196
IL26	504	20	-3693,866	-3687,927	11,877*	2	2	2	-	1	0,198
IL27B	708	16	-5613,682	-5608,707	9,949*	-	114	114, 139	-	1	0,141
IL31	489	14	-4093,760	-4063,715	60,091*	31, 33, 38, 67, 103, 133,	-	30, 33, 38, 43, 56, 67, 141	31, 33, 38, 56, 66, 98, 133	6	1,227
IL34	717	17	-4918,901	-4917,438	2,926	-	134	134	134	1	0,139
IL36A	477	18	-4805,698	-4766,449	78,496*	2, 4, 5, 8, 9, 12, 38, 133	-	8, 133	1, 2, 3, 5, 7, 8, 9, 12, 71, 133, 156	6	1,258
IL36G	471	21	-4966,855	-4916,882	99,946*	2, 3, 5, 6, 104, 140	2, 6, 104	2, 6, 104, 141	2, 5, 6, 18, 26, 41, 52, 55, 69, 73, 91, 104, 128, 144, 148	4	0,849

*p<0.05

[a]Codons identified by more than one ML method are underlined. Codons located in close vicinity of N-glycosylation sites are shaded and those codons related with disulphide bonds are boxed.

[b, c] Codons with P values <0.05

[d] Codons with Bayes factor >50

Alterations in coding regions of genes that lead to amino acids substitutions can induce changes in protein conformation. These changes may be conservative or radical and may alter physiochemical properties of the proteins such as charge and polarity (Majewski and Ott, 2003). From our physiochemical study of the amino acids under postive selection (see Table 3.4 in Supplementary material) we observed that for the majority of codons under positive selections changes in charge and polarity occur. For example, in IL36G, amino acid positions 2 and 5 correspond to 13 and 12 amino acid possibilities respectively, i.e., the different amino acids that can be found in all the species analyzed.

5. DISCUSSION

Mammalian genes involved in the immune response are among the most rapidly evolving genes (Barreiro and Quintana-Murci, 2010; Ferrer-Admetlla et al., 2008; Kosiol et al., 2008) as they include proteins with biological activities designed to protect the host (antibacterial, antiviral, antifungal or antiparasitic). Since positive selection is likely correlated with sites of important activity, an effort was made to try and verify if the positions under positive selection have any particular function that led to their evolution. For most ILs with sites under positive selection, a correlation between those sites and interaction with other molecules, such as receptors, pathogens, and binding proteins (BP), was found. It is known that sites where interaction of the interleukins with their receptors occurs are essential for their function (Koyanagi et al., 2010; Smith, 2006). To verify a cause-effect relationship between ILs and their receptors, we extended our positive selection analyses to interleukin receptors' (ILR), in particular focusing on sites which interact with interleukins at sites detected as being under positive selection (Table 3.2). From all interleukins where positive selection was detected, only in IL4 and IL18 was that correlation found. For IL4, and by focusing on the residues under positive selection known to interact with IL4 receptors, mainly IL4Rα and IL13Rα (see figure 3.1), our results are in agreement with previous studies (Hage et al., 1999; LaPorte et al., 2008; Zarlenga et al., 2004; Zhang et al., 2002).

Table 3. 2. Phylogenetic Tests of Positive Selection for receptors and binding proteins[a]

Gene	No. aa residues	No. of Species	lnL M7	lnL M8	2ΔlnL	PAML	SLAC[b]	FEL[c]	REL[d]	Total no. of sites	% of sites
IL4Rα	2529	10	-13357,029	-13346,832	20,393*	95	-	471, 479, 733, 764	67, 68, 93, 95, 529, 590, 677, 689, 764	2	0,079
IL13Rα	1296	19	-9056,699	-9038,203	36,992*	254, 276	254	83, 196, 225, 237, 254, 259	16, 19, 25, 31, 83, 171, 196, 225, 235, 237, 239, 242, 254, 259, 272, 274, 276, 289, 334, 335	7	0,540
IL18BP	573	21	-5034,333	-5016,838	34,991*	54, 60, 126, 131, 156, 176	11	9, 11, 24, 60, 61	11, 60, 63, 126, 131, 153, 156	5	0,873

*p<0.05

[a]Codons identified by more than one ML method are underlined

[b, c] Codons with P values <0.05

[d] Codons with Bayes factor >50

Indeed of the IL4 residues detected under positive selection, Glu84 is located between sites of interaction with IL4Rα and IL13Rα while Arg105 and Phe106 are located in interaction sites of IL4 with IL4Rα. No codons under positive selection were detected at the sites of IL4:γc interaction as suggested by others (Koyanagi et al., 2010; O'Connell and McInerney, 2005). Of the two residues of IL4Rα under positive selection, only Ser95 has a direct interaction with IL4 (Asn113). IL4 positively selected sites Arg105 and Phe106 are located near sites where IL4 interacts with IL4Rα (Koyanagi et al., 2010; LaPorte et al., 2008; Mueller et al., 2002). For IL13Rα, none of the 7 residues detected as under positive selection interacted with Glu84 of IL4, but the positively selected Ile254 from IL13Rα interacts with IL4. The sites detected are in accordance with those described by Koyanagi et al., 2010, but not with those detected by O'Connel et al., 2005. One explanation for this discordance is the number of species used which differs between studies. For IL18, of the 7 residues detected as under positive selection, only Leu45 is located near the site of interaction with IL18Rα (Lys44) and IL18BP (Glu42). The signatures of positive selection in these sites might also be involved in modulating signaling intensity (Koyanagi et al., 2010).

Figure 3. 1. Positively selected sites in the 3D structures of IL4-IL4Rα-IL13Rα (marked in yellow). IL4 appears in pink, IL4Rα in blue and IL13Rα in brown.

In their evolution, pathogens have evolved an arsenal of immune evading strategies which include antagonists for the host immune related proteins that target critical sites for these proteins' functions. Thus, to escape from antagonists, those targeting sites need to continuously evolve so that the IL-ILR interaction can still be functional. Supporting this concept, sites under positive selection for IL4 have been suggested to be associated with escaping from pathogen-encoded antagonistic proteins (Koyanagi et al., 2010).

Binding proteins (BP) are polypeptides that have the ability to interact with corresponding proteins and neutralize them. In the interleukins studied, there are many examples of such proteins. Indeed, IL18BP has the ability to bind to IL18 with high affinity preventing IL18-IL18R interaction, and therefore, neutralizing their biological activities. Some authors have demonstrated that IL18BP prevents lipopolysaccharide (LPS) induced IFN-γ production (Dinarello, 2000; He et al., 2008; Mallat et al., 2001). In some models, it has been described that blocking IL18 results in a reduction in disease severity in injuries related to IL18 increased production, namely injuries of heart, kidneys, liver, arthritis, etc (Dinarello, 2006). By studying IL18, Leu45 which was detected as under positive selection is located close to the site of interaction of IL18 (Glu42) with IL18BP. In addition, when we studied IL18BP, we detected that Cys131 is under positive selection, next the site of interaction of IL18BP (Lys130) with IL18.

Nevertheless, for some ILs, the evidence of positive selection is not so intelligible. When pathogens or stress are encountered, the initial response is via the innate immune system with activation of the inflammasome. Interestingly, when examined (Table 3.1) two of the genes with largest number of positively selected residues are IL1A and IL6. IL1A is the initial cytokine produced upon inflammasome activation which directly leads to increased expression of IL6 (Akdis et al., 2011). Given their critical importance in the early activation of the immune response, it is tempting to speculate that there has been selective pressure to maintain (rather than to change) structure and function of these molecules from their early precursors. However, the knowledge on the functional role of IL1A and IL6 derives mostly from studies in humans with almost nothing known for other mammals. It is possible that, in other species, these ILs might have evolved for other tasks, leading to the observed divergence in sequence across species.

Codons under positive selection were also detected within sites of N-glycosylation or where disulfide bonds exist. Glycosylation is considered to be important for protein folding, oligomerization, intrinsic stability, solubility, capacity to diffuse troughout the organism, interaction with cell surface receptors and subsequent biological activity (Shental-Bechor and Levy, 2008; Szabo et al., 2009; Waetzig et al., 2010). This modification is an effective way to generate diversity and modulate protein properties due to inherent structural variations of the glycans (Chamorey et al., 2002; Shental-Bechor and Levy, 2008; Waetzig et al., 2010). For IL1A, IL9, IL22, IL25 and IL31, residues under positive selection were found in close proximity to sites where N-glycosylation occurs (Table 3.1). Disulfide bonds play an important role in folding, stability and function of the protein, being, according to some authors (Fass, 2012; Li et al., 2011), associated with functional differentiation of the proteins. These bonds are thought to be well conserved in proteins (Li et al., 2011). Positive selection was detected at sites where disulfide bonds exist for IL2 and IL17A (Table 3.1) and may compromise their structure, stability and function. The changes in charge and polarity observed at the positively selected codons are likely to impose changes in the protein conformation with consequences at the protein function.

6. CONCLUSIONS

Positive selection may function to maintain a host response to pathogens, diseases, and environmental conditions. Here, we have detected positive selection in 28 interleukins with some of the identified codons associated with critical functions of these proteins. Several reasons might underlie our results, including the specific biological functions in which ILs are involved and the maintenance of critical structural features of the ILs. However, the limited knowledge of the role of specific codons in IL functions' for most mammalian species hampers the complete understanding of our observations. Further functional and structural studies by using mutagenesis and crystallographic approaches should be performed for a full comprehension of the role of the observed variation in mammalian ILs.

7. REFERENCES

Afzal, N., Tahir, R., Jahan, S., 2012. **Cytokines: an ever expanding area** Biological and Biomedical Reports 2, 37-43.

Akdis, M., Burgler, S., Crameri, R., Eiwegger, T., Fujita, H., Gomez, E., Klunker, S., Meyer, N., O'Mahony, L., Palomares, O., Rhyner, C., Ouaked, N., Schaffartzik, A., Van De Veen, W., Zeller, S., Zimmermann, M., Akdis, C.A., 2011. **Interleukins, from 1 to 37, and interferon-gamma: receptors, functions, and roles in diseases**. J Allergy Clin Immunol 127, 701-721 e701-770.

Areal, H., Abrantes, J., Esteves, P.J., 2011. **Signatures of positive selection in Toll-like receptor (TLR) genes in mammals**. BMC Evol Biol 11, 368.

Barreiro, L.B., Quintana-Murci, L., 2010. **From evolutionary genetics to human immunology: how selection shapes host defence genes**. Nature reviews. Genetics 11, 17-30.

Beadling, C., Slifka, M.K., 2006. **Regulation of innate and adaptive immune responses by the related cytokines IL-12, IL-23, and IL-27**. Arch Immunol Ther Exp (Warsz) 54, 15-24.

Bielawski, J.P., Yang, Z., 2003. **Maximum likelihood methods for detecting adaptive evolution after gene duplication**. J Struct Funct Genomics 3, 201-212.

Borish, L.C., Steinke, J.W., 2003. **2. Cytokines and chemokines**. J Allergy Clin Immunol 111, S460-475.

Brocker, C., Thompson, D., Matsumoto, A., Nebert, D.W., Vasiliou, V., 2010. **Evolutionary divergence and functions of the human interleukin (IL) gene family**. Hum Genomics 5, 30-55.

Chamorey, A.L., Magne, N., Pivot, X., Milano, G., 2002. **Impact of glycosylation on the effect of cytokines. A special focus on oncology**. Eur Cytokine Netw 13, 154-160.

Commins, S., Steinke, J.W., Borish, L., 2008. **The extended IL-10 superfamily: IL-10, IL-19, IL-20, IL-22, IL-24, IL-26, IL-28, and IL-29**. J Allergy Clin Immunol 121, 1108-1111.

Commins, S.P., Borish, L., Steinke, J.W., 2010. **Immunologic messenger molecules: cytokines, interferons, and chemokines**. J Allergy Clin Immunol 125, S53-72.

Dinarello, C.A., 2000. **Targeting interleukin 18 with interleukin 18 binding protein**. Ann Rheum Dis 59 Suppl 1, i17-20.

Dinarello, C.A., 2006. **Interleukin 1 and interleukin 18 as mediators of inflammation and the aging process**. Am J Clin Nutr 83, 447S-455S.

Fass, D., 2012. **Disulfide bonding in protein biophysics**. Annu Rev Biophys 41, 63-79.

Feghali, C.A., Wright, T.M., 1997. **Cytokines in acute and chronic inflammation**. Front Biosci 2, d12-26.

Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marques-Bonet, T., Ramirez-Soriano, A., Muntasell, A., Navarro, A., Lazarus, R., Calafell, F., Bertranpetit, J., Casals, F., 2008. **Balancing selection is the main force shaping the evolution of innate immunity genes**. J Immunol 181, 1315-1322.

Fumagalli, M., Pozzoli, U., Cagliani, R., Comi, G.P., Riva, S., Clerici, M., Bresolin, N., Sironi, M., 2009. **Parasites represent a major selective force for interleukin genes and shape the genetic predisposition to autoimmune conditions**. J Exp Med 206, 1395-1408.

Haddad, J.J., 2002. **Cytokines and related receptor-mediated signaling pathways**. Biochem Biophys Res Commun 297, 700-713.

Hage, T., Sebald, W., Reinemer, P., 1999. **Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface**. Cell 97, 271-281.

Hall, T.A., 1999. **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT**. Nucl. Acids. Symp. Ser. 41, 95-98.

He, Z., Lu, L., Altmann, C., Hoke, T.S., Ljubanovic, D., Jani, A., Dinarello, C.A., Faubel, S., Edelstein, C.L., 2008. **Interleukin-18 binding protein transgenic mice are protected against ischemic acute kidney injury**. American journal of physiology. Renal physiology 295, F1414-1421.

Heinrich, P.C., Behrmann, I., Haan, S., Hermanns, H.M., Muller-Newen, G., Schaper, F., 2003. **Principles of interleukin (IL)-6-type cytokine signalling and its regulation**. The Biochemical journal 374, 1-20.

Ishihara, K., Hirano, T., 2002. **Molecular basis of the cell specificity of cytokine action**. Biochim Biophys Acta 1592, 281-296.

Kaiser, P., Rothwell, L., Avery, S., Balu, S., 2004. **Evolution of the interleukins**. Developmental and comparative immunology 28, 375-394.

Kosakovsky Pond, S.L., Frost, S.D., 2005. **Not so different after all: a comparison of methods for detecting amino acid sites under selection**. Mol Biol Evol 22, 1208-1222.

Kosiol, C., Vinar, T., da Fonseca, R.R., Hubisz, M.J., Bustamante, C.D., Nielsen, R., Siepel, A., 2008. **Patterns of positive selection in six Mammalian genomes**. PLoS genetics 4, e1000144.

Koyanagi, M., Kerns, J.A., Chung, L., Zhang, Y., Brown, S., Moldoveanu, T., Malik, H.S., Bix, M., 2010. **Diversifying selection and functional analysis of interleukin-4 suggests antagonism-driven evolution at receptor-binding interfaces**. BMC evolutionary biology 10, 223.

Krause, C.D., Pestka, S., 2005. **Evolution of the Class 2 cytokines and receptors, and discovery of new friends and relatives**. Pharmacol Ther 106, 299-346.

LaPorte, S.L., Juo, Z.S., Vaclavikova, J., Colf, L.A., Qi, X., Heller, N.M., Keegan, A.D., Garcia, K.C., 2008. **Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system**. Cell 132, 259-272.

Leonard, W.J., Lin, J.X., 2000. **Cytokine receptor signaling pathways**. The Journal of allergy and clinical immunology 105, 877-888.

Li, X.Q., Zhang, T., Donnelly, D., 2011. **Selective loss of cysteine residues and disulphide bonds in a potato proteinase inhibitor II family**. PLoS One 6, e18615.

Majewski, J., Ott, J., 2003. **Amino acid substitutions in the human genome: evolutionary implications of single nucleotide polymorphisms**. Gene 305, 167-173.

Makalowski, W., Boguski, M.S., 1998. **Synonymous and nonsynonymous substitution distances are correlated in mouse and rat genes**. Journal of molecular evolution 47, 119-121.

Mallat, Z., Corbaz, A., Scoazec, A., Graber, P., Alouani, S., Esposito, B., Humbert, Y., Chvatchko, Y., Tedgui, A., 2001. **Interleukin-18/interleukin-18 binding protein signaling modulates atherosclerotic lesion development and stability**. Circulation research 89, E41-45.

Manry, J., Laval, G., Patin, E., Fornarino, S., Itan, Y., Fumagalli, M., Sironi, M., Tichit, M., Bouchier, C., Casanova, J.L., Barreiro, L.B., Quintana-Murci, L., 2011. **Evolutionary genetic dissection of human interferons**. J Exp Med 208, 2747-2759.

Mueller, T.D., Zhang, J.L., Sebald, W., Duschl, A., 2002. **Structure, binding, and antagonists in the IL-4/IL-13 receptor system**. Biochim Biophys Acta 1592, 237-250.

O'Connell, M.J., McInerney, J.O., 2005. **Gamma chain receptor interleukins: evidence for positive selection driving the evolution of cell-to-cell communicators in the mammalian immune system**. Journal of molecular evolution 61, 608-619.

Ozaki, K., Leonard, W.J., 2002. **Cytokine and cytokine receptor pleiotropy and redundancy**. The Journal of biological chemistry 277, 29355-29358.

Pillai, M.R., Bix, M., 2011. **Evolution of IL4 and pathogen antagonism**. Growth Factors 29, 153-160.

Pond, S.L., Frost, S.D., 2005. **Datamonkey: rapid detection of selective pressure on individual sites of codon alignments**. Bioinformatics 21, 2531-2533.

Poon, A.F., Frost, S.D., Pond, S.L., 2009. **Detecting signatures of selection from DNA sequences using Datamonkey**. Methods Mol Biol 537, 163-183.

Sabat, R., 2010. **IL-10 family of cytokines**. Cytokine & growth factor reviews 21, 315-324.

Schrader, J.W., 2003. **Interleukin is as interleukin does**. J Immunol Methods 276, 1-3.

Secombes, C.J., Wang, T., Bird, S., 2011. **The interleukins of fish**. Developmental and comparative immunology 35, 1336-1345.

Shental-Bechor, D., Levy, Y., 2008. **Effect of glycosylation on protein folding: a close look at thermodynamic stabilization**. Proc Natl Acad Sci U S A 105, 8256-8261.

Smith, K.A., 2006. **The structure of IL2 bound to the three chains of the IL2 receptor and how signaling occurs**. Med Immunol 5, 3.

Steinke, J.W., Borish, L., 2006. **3. Cytokines and chemokines**. The Journal of allergy and clinical immunology 117, S441-445.

Szabo, T.G., Palotai, R., Antal, P., Tokatly, I., Tothfalusi, L., Lund, O., Nagy, G., Falus, A., Buzas, E.I., 2009. **Critical role of glycosylation in determining the length and structure of T cell epitopes**. Immunome research 5, 4.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. **MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods**. Mol Biol Evol 28, 2731-2739.

Tindall, E.A., Hayes, V.M., 2010. **Comprehensive sequence analysis of the human IL23A gene defines new variation content and high rate of evolutionary conservation**. DNA research : an international journal for rapid publication of reports on genes and genomes 17, 117-122.

Waetzig, G.H., Chalaris, A., Rosenstiel, P., Suthaus, J., Holland, C., Karl, N., Valles Uriarte, L., Till, A., Scheller, J., Grotzinger, J., Schreiber, S., Rose-John, S., Seegert, D., 2010. **N-linked glycosylation is essential for the stability but not the signaling function of the interleukin-6 signal transducer glycoprotein 130**. J Biol Chem 285, 1781-1789.

Wlasiuk, G., Nachman, M.W., 2010. **Adaptation and constraint at Toll-like receptors in primates**. Mol Biol Evol 27, 2172-2186.

Yang, Z., 1997. **PAML: a program package for phylogenetic analysis by maximum likelihood**. Computer applications in the biosciences : CABIOS 13, 555-556.

Yang, Z., 2002. **Inference of selection from multiple species alignments**. Current opinion in genetics & development 12, 688-694.

Yang, Z., 2007. **PAML 4: phylogenetic analysis by maximum likelihood**. Mol Biol Evol 24, 1586-1591.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. **Codon-substitution models for heterogeneous selection pressure at amino acid sites**. Genetics 155, 431-449.

Zarlenga, D.S., Dawson, H., Kringel, H., Solano-Aguilar, G., Urban, J.F., Jr., 2004. **Molecular cloning of the swine IL-4 receptor alpha and IL-13 receptor 1-chains: effects of experimental Toxoplasma gondii, Ascaris suum and Trichuris suis infections on tissue mRNA levels**. Veterinary immunology and immunopathology 101, 223-234.

Zelus, D., Robinson-Rechavi, M., Delacre, M., Auriault, C., Laudet, V., 2000. **Fast evolution of interleukin-2 in mammals and positive selection in ruminants**. Journal of molecular evolution 51, 234-244.

Zhang, J., Nei, M., 2000. **Positive selection in the evolution of mammalian interleukin-2 genes**. Molecular biology and evolution 17, 1413-1416.

Zhang, J.L., Simeonowa, I., Wang, Y., Sebald, W., 2002. **The high-affinity interaction of human IL-4 and the receptor alpha chain is constituted by two independent binding clusters**. J Mol Biol 315, 399-407.

Zhou, G., Zhai, Y., Dong, X., Zhang, X., He, F., Zhou, K., Zhu, Y., Wei, H., Yao, Z., Zhong, S., Shen, Y., Qiang, B., 2004. **Haplotype structure and evidence for positive selection at the human IL13 locus**. Mol Biol Evol 21, 29-35.

8. SUPPLEMENTARY MATERIAL

Table 3. 3. List of the accession numbers of the sequences used for each IL alignment.

Species	Accession number
IL1A	
Ailuropoda melanoleuca	ENSAMEG0000001075
Bos taurus	NM174092.1
Bubalus bubalis	AB246786.1
Bubalus carabanensis	DQ188096.1
Callithrix jacchus	ENSCJAT00000031857
Canis lupus familiaris	NM001003157.2
Capra hircus	D63350.1
Cavia porcellus	ENSCPOT00000027229
Cercocebus torquatus	U19836.1
Echinops telfairi	ENSETET00000017640
Equus caballus	NM001082500.1
Felis catus	NM001009351.1
Homo sapiens	NM000575.3
Lama glama	AB107645.1
Loxodonta africana	ENSLAFT00000001246
Macaca fascicularis	AB000553.1
Macaca mulatta	NM001042757.1
Mus musculus	NM010554.4
Myotis lucifugus	ENSMLUT00000000165
Ochotona princeps	ENSOPRT00000012798
Oryctolagus cuniculus	NM001101684.1
Otolemur garnettii	ENSOGAT00000005383
Ovis aries	NM001009808.1
Pan troglodytes	XM525866.2
Procavia capensis	ENSPCAT00000005993
Pteropus vampyrus	ENSPVAT00000014232
Rattus norvegicus	NM017019.1
Sorex araneus	ENSSART00000013019
Spermophilus tridecemlineatus	ENSSTOT00000002284
Sus scrofa	NM214029.1
Tarsius syrichta	ENSTSYT00000012201
Tursiops truncatus	ENSTTRT00000011057
Vicugna pacos	ENSVPAT00000011121
IL1B	
Bos taurus	NM174093.1
Bubalus bubalis	AB246787.1
Bubalus carabanensis	AB246783.1
Callithrix jacchus	ENSCJAT00000031803
Canis lupus familiaris	NM001037971.1
Capra hircus	D63351.1
Cavia porcellus	NM001172968.1
Cercocebus torquatus	U19837.1
Cervus elaphus	U20500.1
Delphinapterus leucas	AF320322.1
Dipodomys ordii	ENSDORG00000009905
Echinops telfairi	ENSETEG00000008505
Equus caballus	NM001082526.1
Eumetopias jubatus	AY517546.1
Felis catus	NM001077414.1
Gorilla gorilla	ENSGGOG00000006387

Homo sapiens	NM000576.2
Lama glama	AB107644.1
Loxodonta africana	ENSLAFG00000004059
Macaca fascicularis	D63353.1
Macaca mulatta	NM001042756.1
Macaca nemestrina	U19853.1
Microcebus murinus	ENSMICG00000013334
Mus musculus	NM008361.3
Myotis lucifugus	ENSMLUG00000009941
Nomascus leucogenys	ENSNLEG00000001886
Ochotona princeps	ENSOPRG00000012817
Oryctolagus cuniculus	NM001082201.1
Ovis aries	NM001009465.2
Pan troglodytes	ENSPTRT00000022971
Phoca hispida	DQ358051.1
Phoca vitulina richardsi	AY578791.1
Pteropus vampyrus	ENSPVAG00000006483
Rattus norvegicus	ENSRNOG00000004649
Sus scrofa	NM214055.1
Tursiops truncatus	ENSTTRG00000011069
Vicugna pacos	ENSVPAG00000011126
IL2	
Ailuropoda melanoleuca	XM002923818.1
Aotus lemurinus	U88364.1
Aotus nancymaae	U88361.1
Aotus nigriceps	U88363.1
Aotus vociferans	U88362.1
Bos taurus	NM180997.2
Boselaphus tragocamelus	DQ017831.1
Bubalus bubalis	AF363786.1
Bubalus carabanensis	AB246271.1
Camelus bactrianus	AB246671.1
Camelus dromedarius	HM051105.1
Canis lupus familiaris	NM001003305.1
Capra aegagrus	EF056470.1
Capra falconeri	EF056471.1
Capra hircus	X76063.1
Cavia porcellus	NM001172837.1
Cercocebus torquatus	U19846.1
Dasypus novemcinctus	ENSDNOT00000009053
Delphinapterus leucas	AF072870.1
Equus caballus	NM001085433.1
Felis catus	NM001043337.1
Gorilla gorila	ENSGGOT00000028949
Homo sapiens	NM000586.3
Loxodonta africana	ENSLAFT00000004431
Macaca fascicularis	D63352.1
Macaca mulatta	NM001047130.1
Macaca nemestrina	U19852.1
Meriones unguiculatus	X68779.1
Mesocricetus auratus	EU729351.1
Mirounga angustirostris	U79187.1
Moschus berezovskii	AY840980.1
Mus musculus	NM008366.3
Myotis lucifugus	ENSMLUT00000011794
Nomascusleucogenys	K03174.1

Ochotonaprinceps	ENSOPRG00000002193:ENSOPRT00000002175
Oryctolagus cuniculus	NM001163180.1
Ovis aries	NM001009806.1
Pan troglodytes	XM003310465.1
Papio anubis	AY234220.1
Papio hamadryas	U88365.1
Rattus norvegicus	NM053836.1
Rousettus leschenaultii	AB472358.1
Saimiri sciureus	AF294755.1
Sorexaraneus	ENSSARG00000000201:ENSSART00000000204
Sus scrofa	NM213861.1
Syncerus caffer	AB571123.1
Vulpes vulpes	AJ621188.1
IL4	
Ailuropoda melanoleuca	XM002912899.1
Bison bonasus	EU000422.1
Bos taurus	NM173921.2
Boselaphus tragocamelus	AY939910.1
Bubalus bubalis	AY293620.1
Bubalus carabanensis	AB246275.1
Callithrix jacchus	XM002744606.1
Camelus bactrianus	AB246673.1
Camelus dromedarius	HM051106.1
Canis lupus familiaris	NM001003159.1
Capra hircus	U34273.1
Cercocebus torquatus	U19838.1
Elephas maximus	EU000424.1
Equus caballus	NM001082519.1
Equus zebra hartmannae	EU000427.1
Erinaceus europaeus	ENSEEUT00000010882
Felis catus	NM001043339.1
Giraffa camelopardalis	EU000423.1
Gorilla gorilla	ENSGGOT00000005697
Homo sapiens	NM000589.2
Lama glama	AB107648.1
Macaca fascicularis	AB000515.1
Macaca mulatta	NM001032904.1
Meriones unguiculatus	L37779.1
Microcebus murinus	ENSMICT00000012463
Mus musculus	NM021283.2
Mus musculus molossinus	AB174765.1
Mustela putorius furo	EF368210.1
Myotis lucifugus	ENSMLUT00000016579
Ochotona princeps	ENSOPRG00000011775:ENSOPRT00000011766
Oryctolagus cuniculus	NM001163177.1
Otolemur garnettii	ENSOGAT00000015510
Ovis aries	NM001009313.2
Pan troglodytes	NM001011714.1
Papio anubis	NM001112653.1
Peromyscus maniculatus	DQ446203.1
Pteropus vampyrus	ENSPVAT00000015176
Rattus norvegicus	NM201270.1
Rousettus leschenaultii	AB472359.1
Sigmodon hispidus	AF421390.1
Spermophilus tridecemlineatus	ENSSTOT00000007974
Sus scrofa	NM214123.1

Syncerus caffer	EU000421.1
Tursiops truncatus	AB020732.1
IL5	
Bison bison	EU915013.1
Bos gaurus	EU915001.1
Bos tauros	ENSBTAG00000031235:ENSBTAT00000044228
Bubalus bubalis	EF407851.1
Callithrix jacchus	DQ658152.1
Canis familiaris	ENSCAFT00000001335
Capra hircus	EU914965.1
Cavia porcellus	NM001172970.1
Cercocebus torquatus	L26033.1
Choloepus hoffmanni	ENSCHOT00000005746
Dasypus novemcinctus	ENSDNOT00000002731
Echinops telfairi	ENSETET00000007039
Equus caballus	ENSECAT00000017688
Felis catus	NM001009845.2
Gorilla gorila	ENSGGOT00000011623
Homosapiens	ENST00000231454
Loxodontaafricana	ENSLAFT00000009760
Macacamulatta	ENSMMUT00000004434
Macropus eugenii	AF064209.1
Meriones unguiculatus	L37780.1
Monodelphis domestica	ENSMODT00000017776
Mus musculus	ENSMUST00000048605
Myotis lucifugus	ENSMLUT00000016553
Nomascus leucogenys	ENSNLET00000016239
Ochotona princeps	ENSOPRT00000011721
Oryctolagus cuniculus	ENSOCUT00000004166
Ovis aries	NM001009783.1
Pantroglodytes	ENSPTRT00000031882
Pteropus vampyrus	ENSPVAT00000015170
Rattus norvegicus	ENSRNOT00000010729
Saimiri sciureus	AF294756.1
Sigmodon hispidus	AF148211.1
Sorex araneus	ENSSART00000011773
Sus scrofa	ENSSSCT00000015601
Tarsius syrichta	ENSTSYT00000010506
Vicugna pacos	ENSVPAT00000003840
IL6	
Ailuropoda melanoleuca	JN624856.1
Bos taurus	NM173923.2
Bubalus bubalis	AY347710.1
Bubalus carabanensis	AB246784.1
Callithrix jacchus	DQ658153.1
Camelus bactrianus	AB107656.1
Camelus dromedarius	HM051107.1
Canis lupus familiaris	NM001003301.1
Capra hircus	D86569.1
Cercocebus torquatus	L26032.1
Chlorocebus sabaeus	FJ194486.1
Delphinapterus leucas	AF076643.1
Dipodomys ordii	ENSDORT00000009373
Echinops telfairi	ENSETET00000001220
Equus caballus	NM001082496.1
Erinaceus europaeus	ENSEEUT00000005028

Felis catus	NM001009211.1
Gorilla gorila	ENSGGOT00000005362
Homo sapiens	NM000600.3
Lama glama	AB107647.1
Loxodonta africana	ENSLAFT00000015572
Macaca fascicularis	AB000554.1
Macaca mulatta	NM001042733.1
Macaca thibetana	AY849928.1
Microcebus murinus	ENSMICT00000006697
Myotis lucifugus	ENSMLUT00000015856
Mus musculus	NM031168.1
Mustela putorius furo	EF368209.1
Nomascus leucogenys	ENSNLET00000021390
Ochotona princeps	ENSOPRT00000003065
Oryctolagus cuniculus	NM001082064.1
Otolemur garnettii	ENSOGAT00000004179
Ovis aries	NM001009392.1
Papio anubis	NM001173536.1
Rattus norvegicus	NM012589.1
Rousettus leschenaultii	AB472360.1
Saimiri sciureus	AF294757.1
Sorex araneus	ENSSART00000012781
Sus scrofa	NM214399.1
Syncerus caffer	AB571652.1
Tursiops truncatus	ENSTTRT00000009040
Vulpes vulpes	AJ621189.1
IL7	
Bos taurus	NM173924.2
Canis lupus familiaris	NM001048138.1
Cavia porcellus	ENSCPOT00000010043
Chlorocebus sabaeus	FJ194487.1
Dasypus novemcinctus	ENSDNOT00000016289
Erinaceus europaeus	ENSEEUT00000014134
Gorilla gorila	ENSGGOT00000000478
Homo sapiens	NM000880.3
Macaca mulatta	NM001032846.1
Mus musculus	NM008371.4
Myotis lucifugus	ENSMLUT00000024474
Nomascus leucogenys	ENSNLET00000002660
Ovis aries	NM001009777.1
Pteropus vampyrus	ENSPVAT00000015872
Rattus norvegicus	NM013110.2
Sorex araneus	ENSSART00000003720
Sus scrofa	NM214135.1
Tarsius syrichta	ENSTSYT00000008272
Tupaia belangeri	ENSTBET00000007176
Tursiops truncatus	ENSTTRT00000000917
Vicugna pacos	ENSVPAT00000008415
IL8	
Ailuropoda melanoleuca	ENSAMET00000005297
Bos indicus	EU490318.1
Bos taurus	ENSBTAT00000026275
Canis familiaris	ENSCAFT00000004871
Cavia porcellus	NM001173399
Cercocebus torquatus	U19839.1
Chlorocebus sabaeus	FJ194488.1

Choloepus hoffmanni	ENSCHOT00000003064
Dasypus novemcinctus	ENSDNOT00000003797
Echinops telfairi	ENSETET00000016665
Equus caballus	ENSECAT00000016212
Felis catus	NM001009281
Gorilla gorila	ENSGGOT00000028757
Homo sapiens	ENST00000307407
Loxodonta africana	ENSLAFT00000000483
Macaca mulatta	ENSMMUT00000005432
Macaca nemestrina	U19851.1
Marmota monax	EU332349.1
Monodelphis domestica	ENSMODT00000024364
Myotis lucifugus	ENSMLUT00000004869
Nomascus leucogenys	ENSNLET00000010117
Ochotona prínceps	ENSOPRT00000002239
Oryctolagus cuniculus	NM001082293
Ovis aries	NM001009401
Pan troglodytes	ENSPTRT00000030085
Papio anubis	NM001173537
Pteropus vampyrus	ENSPVAT00000015367
Sorex araneus	ENSSART00000007042
Sus scrofa	ENSSSCT00000009807
Tarsius syrichta	ENSTSYT00000004113
Tupaia belangeri	ENSTBET00000010783
Tursiops truncatus	AB096002.1
Vicugna pacos	ENSVPAT00000003176
IL9	
Bos taurus	ENSBTAT00000024340
Cavia porcellus	ENSCPOT00000028361
Chlorocebus sabaeus	FJ194489.1
Choloepus hoffmanni	ENSCHOT00000010002
Dasypus novemcinctus	ENSDNOT00000005394
Felis catus	ENSFCAT00000013935
Gorilla gorilla	ENSGGOT00000014167
Homo sapiens	ENST00000274520
Loxodonta africana	ENSLAFT00000018403
Microcebus murinus	ENSMICT00000007280
Mus musculus	ENSMUST00000022019
Nomascus leucogenys	ENSNLET00000010589
Ochotona princeps	ENSOPRT00000010013
Otolemur garnettii	ENSOGAT00000009580
Pan troglodytes	ENSPTRT00000031950
Pteropus vampyrus	ENSPVAT00000010067
Rattus norvegicus	ENSRNOT00000016187
Sus scrofa	ENSSSCT00000015636
Tarsius syrichta	ENSTSYT00000014523
Tursiops truncatus	ENSTTRT00000014180
IL10	
Bos taurus	NM174088.1
Bubalus bubalis	AB246351.1
Bubalus carabanensis	AB246276.1
Callithrix jacchus	DQ658154.1
Camelus bactrianus	AB246674.1
Canis familiaris	ENSCAFT00000018154
Cavia porcellus	ENSCPOT00000009023
Cercocebus torquatus	L26030.1

Cervus elaphus	U11767.1
Dipodomys ordii	ENSDORT00000011288
Echinops telfairi	ENSETET00000011736
Equus caballus	NM001082490.1
Felis catus	NM001009209.1
Gorilla gorilla	ENSGGOT00000014347
Homo sapiens	NM000572.2
Lama glama	AB107649.1
Loxodonta africana	ENSLAFT00000021835
Macaca fascicularis	AB000514.1
Macaca mulatta	NM001044727.1
Macaca nemestrina	L26031.1
Macropus eugenii	ENSMEUT00000006795
Marmota himalayana	EF414445.1
Marmota monax	AY253723.1
Meriones unguiculatus	L37781.1
Microcebus murinus	ENSMICT00000008350
Monodelphis domestica	ENSMODT00000002607
Mus musculus	NM010548.2
Myotis lucifugus	ENSMLUT00000017719
Nomascus leucogenys	ENSNLET00000000987
Oryctolagus cuniculus	AF068058.1
Otolemur garnettii	ENSOGAT00000004080
Ovis aries	NM001009327.1
Pan troglodytes	NM001135620.2
Papio hamadryas	AY796417.1
Peromyscus maniculatus	AY251293.1
Procavia capensis	ENSPCAT00000008764
Rattus norvegicus	NM012854. 2
Rousettus leschenaultii	AB472361.1
Saimiri sciureus	AF294758.1
Sorex araneus	ENSSART00000004540
Sus Scrofa	MN214041.1
Syncerus caffer	AB571124.1
Trichosurus vulpecula	AF026277.1
Vulpes vulpes	AJ621190.1
IL12A	
Bos taurus	ENSBTAT00000020153
Bubalus carabanensis	AB246272.1
Canis familiaris	ENSCAFT00000022543
Capra hircus	AF003542.1
Cavia porcellus	NM 001172840.1
Cercocebus torquatus	U19835.1
Cervus elaphus	U57751.1
Equus caballus	ENSECAT00000026617
Erinaceus europaeus	ENSEEUT00000004467
Felis catus	NM 001009833.1
Gorilla gorilla	ENSGGOT00000006855
Homo sapiens	ENST00000305579
Lama glama	AB107653.1
Loxodonta africana	ENSLAFT00000001088
Macaca mulatta	ENSMMUT00000032477
Microcebus murinus	ENSMICT00000006469
Mus musculus	ENSMUST00000029345
Oryctolagus cuniculus	ENSOCUT00000010149
Otolemur garnettii/	ENSOGAG00000013000:ENSOGAT00000013001

Ovis aries	NM 001009736.1
Pan troglodytes	ENSPTRT00000029052
Papio anubis	NM 001112637.1
Pteropus vampyrus	ENSPVAT00000006081
Rattus norvegicus	ENSRNOT00000012831
Sigmodon hispidus	AF421396.1
Sus scrofa	ENSSSCT00000012840
Tursiops truncatus	ENSTTRG00000009289:ENSTTRT00000009287
IL12B	
Bos taurus	ENSBTAT00000006222
Bubalus bubalis	AY198121.1
Bubalus carabanensis	AB246273.1
Canis familiaris	ENSCAFT00000027382
Capra hircus	AF007576.1
Cavia porcellus	NM 001172967.1
Cercocebus torquatus	U19834.1
Cervus elaphus	U57752.1
Choloepus hoffmanni	ENSCHOT00000000421
Equus caballus	ENSECAT00000013594
Felis catus	NM 001077413.1
Gorilla gorilla	ENSGGOT00000004439
Homo sapiens	ENST00000231228
Lama glama	AB107654.1
Loxodonta africana	ENSLAFT00000015755
Macaca mulatta	ENSMMUT00000030533
Marmota monax	X97019.1
Mesocricetus auratus	AB085792.1
Mus musculus	ENSMUST00000170513
Myotis lucifugus	ENSMLUT00000006353
Nomascus leucogenys	ENSNLET00000005482
Ochotona princeps	ENSOPRT00000013152
Oryctolagus cuniculus	ENSOCUT00000016318
Otolemur garnettii	ENSOGAT00000014013
Ovis aries	NM 001009438.1
Pan troglodytes	ENSPTRT00000032327
Papio anubis	AY234218.1
Pongo abelii	ENSPPYT00000018623
Pteropus vampyrus	ENSPVAT00000006918
Rattus norvegicus	ENSRNOT00000067620
Sigmodon hispidus	AF421395.1
Sus scrofa	ENSSSCT00000018557
Tupaia belangeri	ENSTBET00000009324
IL14	
Ailuropoda melanoleuca	XM 002921887.1
Bos taurus	XM 002685068.1
Callithrix jacchus	XM 002750558.1
Canis lupus familiaris	XM 544439.2
Cricetulus griseus	XM 003500640.1
Equus caballus	XM 001503841.2
Gorilla gorilla	ENSGGOG00000001571:ENSGGOT00000001581
Homo Sapiens	NM 175852.3
Loxodonta africana	XM 003415281.1
Mus musculus	NM 001005506.3
Myotis lucifugus	ENSMLUG00000017480:ENSMLUT00000017484
Nomascus leucogenys	XM 003276358.1
Oryctolagus cuniculus	ENSOCUT00000007529

Otolemur garnettii	ENSOGAG00000009458:ENSOGAT00000009463
Rattus norvegicus	NM 001127633.1
Sus scrofa	XM 003127774.3
Tarsius syrichta	ENSTSYG00000000340:ENSTSYT00000000340
Tursiops truncatus	ENSTTRG00000009711:ENSTTRT00000009712
IL15	
Ailuropoda melanoleuca	EU375450.1
Bos taurus	NM 174090.1
Bubalus bubalis	AJ891036.1
Cavia porcellus	NM 001172829.1
Cercopithecus aethiops	U03099.1
Chlorocebus sabaeus	FJ194485.1
Dipodomys ordii	ENSDORT00000014153
Equus caballus	NM 001163986.1
Felis catus	NM 001009207.1
Gorilla gorilla	ENSGGOT00000011327
Homo sapiens	NM 000585.4
Loxodonta africana	ENSLAFT00000008638
Macaca mulatta	NM 001044731.1
Macaca thibetana	DQ021912.1
Marmota himalayana	EF414446.1
Marmota monax	AY426605.1
Microcebus murinus	ENSMICT00000003152
Mus musculus	NM 008357.1
Myotis lucifugus	ENSMLUT00000007614
Nomascus leucogenys	ENSNLEG00000005257:ENSNLET00000006706
Oryctolagus cuniculus	NM 001082216.1
Otolemur garnettii	ENSOGAG00000011090:ENSOGAT00000011094
Ovis aries	NM 001009734.1
Pan troglodytes	ENSPTRT00000030607
Procavia capensis	ENSPCAT00000013218
Pteropus vampyrus	ENSPVAT00000006435
Rattus norvegicus	NM 013129.2
Sus scrofa	NM 214390.1
Tursiops truncatus	ENSTTRG00000000131:ENSTTRT00000000131
IL16	
Bos taurus	ENSBTAT00000014556
Callithrix jacchus	ENSCJAT00000015521
Canis familiaris	ENSCAFT00000021964
Cavia porcellus	ENSCPOT00000015641
Equus caballus	ENSECAT00000024507
Homo sapiens	ENST00000394660
Loxodonta africana	ENSLAFT00000002339
Macaca mulatta	ENSMMUT00000016182
Mus musculus	ENSMUST00000001792
Nomascus leucogenys	ENSNLET00000001717
Ochotona princeps	ENSOPRT00000002093
Oryctolagus cuniculus	ENSOCUT00000010388
Pan troglodytes	ENSPTRG00000007369:ENSPTRT00000013586
Pongo abelii	ENSPPYT00000007936
Pteropus vampyrus	ENSPVAT00000016174
Rattus norvegicus	ENSRNOT00000016289
Sus scrofa	ENSSSCT00000002002
IL17A	
Bos taurus	NM 001008412.1
Canis lupus familiaris	NM 001165878.1

Capra hircus	GU269912.1
Cavia porcellus	ENSCPOG00000010505:ENSCPOT00000010600
Dipodomys ordii	ENSDORG00000008932:ENSDORT00000008930
Echinops telfairi	ETEG00000002530:ENSETET00000002529
Equus caballus	NM 001143792.1
Erinaceus europaeus	ENSEEUG00000006605:ENSEEUT00000006571
Felis catus	ENSFCAG00000010067:ENSFCAT00000010070
Gorilla gorilla	ENSGGOG00000009736:ENSGGOT00000009771
Homo sapiens	ENST00000340057
Loxodonta africana	ENSLAFG00000001544:ENSLAFT00000001544
Macaca mulatta	XM 001106391.2
Macropus eugenii	ENSMEUG00000005305:ENSMEUT00000005320
Microcebus murinus	ENSMICG00000008438:ENSMICT00000008434
Mus musculus	ENSMUST00000027061
Myotis lucifugus	ENSMLUG00000001736:ENSMLUT00000001737
Nomascus leucogenys	ENSNLEG00000003884:ENSNLET00000004937
Oryctolagus cuniculus	ENSOCUG00000004270:ENSOCUT00000004268
Pan troglodytes	XM 527408.2
Procavia capensis	ENSPCAG00000011830:ENSPCAT00000011752
Pteropus vampyrus	ENSPVAG00000010174:ENSPVAT00000010174
Rattus norvegicus	ENSRNOT00000016664
Sorex araneus	ENSSARG00000007196:ENSSART00000007184
Sus scrofa	NM 001005729.1
Tursiops truncatus	ENSTTRG00000004836:ENSTTRT00000004832
Vicugna pacos	ENSVPAG00000000545:ENSVPAT00000000545
IL17C	
Bos taurus	ENSBTAG00000047128:ENSBTAT00000065977
Canis familiaris	ENSCAFG00000019883:ENSCAFT00000031634
Cavia porcellus	ENSCPOG00000001017:ENSCPOT00000001025
Equus caballus	ENSECAG00000009737:ENSECAT00000009944
Gorilla gorilla	ENSGGOG00000012533:ENSGGOT00000012570
Homo sapiens	NM 013278.3
Microcebus murinus	ENSMICG00000006713:ENSMICT00000006706
Mus musculus	NM 145834.3
Myotis lucifugus	ENSMLUG00000008861:ENSMLUT00000008843
Nomascus leucogenys	ENSNLEG00000000566:ENSNLET00000000696
Ochotona princeps	ENSOPRG00000006630:ENSOPRT00000006620
Otolemur garnettii	ENSOGAG00000000416:ENSOGAT00000000417
Pan troglodytes	ENSPTRG00000008457:ENSPTRT00000015582
Procavia capensis	ENSPCAG00000010041:ENSPCAT00000009979
Pteropus vampyrus	ENSPVAG00000017351:ENSPVAT00000017350
Rattus norvegicus	ENSRNOG00000013007:ENSRNOT00000017358
Sorex araneus	ENSSARG00000001623:ENSSART00000001615
Tursiops truncatus	ENSTTRG00000006251:ENSTTRT00000006245
IL18	
Bos taurus	NM174091.2
Boselaphus tragocamelus	AY842499.1
Bubalus bubalis	AY394479.1
Canis lupus familiaris	NM001003169.1
Capra hircus	AY605263.1
Cavia porcellus	AB025722.1
Choloepus hoffmanni	ENSCHOT00000003599
Equus caballus	NM001082512.1
Felis catus	NM001009213.2
Gorilla gorilla	ENSGGOT00000015920
Homo sapiens	NM001243211.1

Loxodonta africana	ENSLAFT00000026165
Macaca mulatta	NM001032834.1
Meriones unguiculatus	AY095932.1
Mus musculus	NM008360.1
Nomascus leucogenys	ENSNLET00000009316
Nyctereutes procyonoides	EU200969.1
Oryctolagus cuniculus	NM001122940.1
Ovis aries	NM001009263.1
Pan troglodytes	ENSPTRT00000007973
Rattus norvegicus	NM019165.1
Sus scrofa	NM213997.1
Vicugna pacos	ENSVPAT00000000057
Vulpes vulpes	EF581889.1
IL20	
Bos taurus	ENSBTAT00000003574
Canis familiaris	ENSCAFT00000018180
Cavia porcellus	ENSCPOT00000005092
Choloepus hoffmanni	ENSCHOT00000010825
Dasypus novemcinctus	ENSDNOT00000014500
Dipodomys ordii	ENSDORT00000008856
Equus caballus	ENSECAT00000009853
Felis catus	ENSFCAG00000013827:ENSFCAT00000013831
Gorilla gorilla	ENSGGOT00000013318
Homo sapiens	ENST00000367096
Loxodonta africana	ENSLAFT00000018715
Macaca mulatta	ENSMMUT00000033157
Microcebus murinus	ENSMICT00000008363
Monodelphis domestica	ENSMODG00000002115:ENSMODT00000002626
Mus musculus	ENSMUST00000027673
Nomascus leucogenys	ENSNLET00000001030
Ochotona princeps	ENSOPRT00000011337
Oryctolagus cuniculus	ENSOCUT00000014282
Otolemur garnettii	ENSOGAT00000004083
Pan troglodytes	ENSPTRT00000003492
Rattus norvegicus	DQ229286.1
Sus scrofa	ENSSSCG00000015654:ENSSSCT00000017051
Tarsius syrichta	ENSTSYT00000008788
Tupaia belangeri	ENSTBET00000012959
Tursiops truncatus	ENSTTRT00000012109
IL21	
Bos taurus	ENSBTAT00000016413
Canis familiaris	ENSCAFT00000006415
Cavia porcellus	ENSCPOT00000012550
Echinops telfairi	ENSETET00000013163
Homo sapiens	ENST00000264497
Loxodonta africana	ENSLAFT00000006023
Macaca mulatta	ENSMMUT00000004728
Mus musculus	ENSMUST00000029273
Myotis lucifugus	ENSMLUT00000010925
Nomascus leucogenys	ENSNLET00000012063
Oryctolagus cuniculus	ENSOCUT00000015297
Pan troglodytes	ENSPTRT00000030528
Pteropus vampyrus	ENSPVAT00000016467
Rattus norvegicus	ENSRNOT00000023348
Sus scrofa	ENSSSCT00000009953
Tarsius syrichta	ENSTSYT00000002221

Tupaia belangeri	ENSTBET00000015850
Tursiops truncatus	ENSTTRT00000014492
Vicugna pacos	ENSVPAT00000010534
IL22	
Bos taurus	ENSBTAT00000020476
Canis familiaris	ENSCAFT00000000647
Capra hircus	HM542482.1
Cavia porcellus	ENSCPOT00000010011
Dasypus novemcinctus	ENSDNOT00000005432
Equus caballus	XM 001491754.2
Erinaceus europaeus	ENSEEUT00000004607
Gorilla gorilla	ENSGGOT00000008988
Homo sapiens	ENST00000538666
Macaca mulatta	ENSMMUT00000026319
Microcebus murinus	ENSMICT00000003649
Mus musculus	ENSMUST00000096691
Myotis lucifugus	ENSMLUT00000002706
Nomascus leucogenys	ENSNLET00000004491
Ochotona princeps	ENSOPRT00000017262
Oryctolagus cuniculus	ENSOCUT00000001663
Otolemur garnettii	ENSOGAT0000000431
Ovis aries	HE617662.1
Pan troglodytes	ENSPTRG00000005195:ENSPTRT0000000952
Procavia capensis	ENSPCAT00000004576
Pteropus vampyrus	ENSPVAT00000015053
Rattus norvegicus	ENSRNOT00000009776
Spermophilus tridecemlineatus	ENSSTOT00000012769
Sus scrofa	ENSSSCT00000000520
Tursiops truncatus	ENSTTRT00000006507
IL25	
Bos taurus	ENSBTAT00000008806
Canis familiaris	ENSCAFT00000018418
Cavia porcellus	ENSCPOT00000007094
Erinaceus europaeus	ENSEEUT00000013993
Gorilla gorilla	ENSGGOT00000017179
Homo sapiens	ENST00000329715
Loxodonta africana	ENSLAFT00000016968
Macropus eugenii	ENSMEUT00000008676
Mus musculus	ENSMUST00000037863
Myotis lucifugus	ENSMLUT00000017213
Ochotona princeps	ENSOPRT00000016503
Oryctolagus cuniculus	ENSOCUT00000011024
Otolemur garnettii	ENSOGAT00000008716
Ovis aries	NM 001195219.1
Pan troglodytes	ENSPTRT00000011316
Pteropus vampyrus	ENSPVAT00000005164
Rattus norvegicus	ENSRNOT00000022563
Spermophilus tridecemlineatus	ENSSTOT00000002552
Sus scrofa	ENSSSCT00000002267
Tupaia belangeri	ENSTBET00000000948
IL26	
Canis familiaris	ENSCAFT00000000644
Choloepus hoffmanni	ENSCHOT00000007234
Echinops telfairi	ENSETET00000005718
Felis catus	ENSFCAT00000010577
Gorilla gorilla	ENSGGOT00000011227

Homo sapiens	ENST00000229134
Macaca mulatta	ENSMMUT00000006205
Microcebus murinus	ENSMICT00000003642
Myotis lucifugus	ENSMLUT00000002700
Nomascus leucogenys	ENSNLET00000004457
Oryctolagus cuniculus	ENSOCUT00000000495
Otolemur garnettii	ENSOGAT00000028397
Pan troglodytes	ENSPTRT00000009541
Procavia capensis	ENSPCAT00000005542
Pteropus vampyrus	ENSPVAT00000015051
Sorex araneus	ENSSART00000002471
Sus scrofa	ENSSSCT00000000518
Tupaia belangeri	ENSTBET00000000479
Tursiops truncatus	ENSTTRT00000006497
Vicugna pacos	ENSVPAT00000000924
IL27B	
Bos taurus	ENSBTAG00000012829:ENSBTAT00000017046
Canis familiaris	ENSCAFG00000019114:ENSCAFT00000030356
Cavia porcellus	ENSCPOG00000000706:ENSCPOT00000000711
Equus caballus	ENSECAG00000018667:ENSECAT00000019767
Gorilla gorilla	ENSGGOG00000011932:ENSGGOT00000011974
Homo sapiens	ENSG00000105246:ENST00000221847
Loxodonta africana	ENSLAFG00000022621:ENSLAFT00000022044
Macaca mulatta	ENSMMUG00000010099:ENSMMUT00000014097
Mus musculus	ENSMUSG00000003206:ENSMUST00000003274
Myotis lucifugus	ENSMLUG00000008782:ENSMLUT00000008775
Nomascus leucogenys	ENSNLET00000003939
Otolemur garnettii	ENSOGAG00000001661:ENSOGAT00000001662
Pan troglodytes	ENSPTRG00000010298:ENSPTRT00000018919
Pteropus vampyrus	ENSPVAG00000016926:ENSPVAT00000016926
Rattus norvegicus	NM 001109421.1
Sus scrofa	ENSSSCG00000013498:ENSSSCT00000014744
IL31	
Canis lupus familiaris	NM 001165914.1
Echinops telfairi	ENSETEG00000012673:ENSETET00000012673
Gorilla gorilla	ENSGGOG00000005159:ENSGGOT00000005183
Homo sapiens	ENST00000377035
Loxodonta africana	ENSLAFG00000016441:ENSLAFT00000016441
Macaca mulatta	XM 001096743.2
Microcebus murinus	ENSMICG00000001165:ENSMICT00000001164
Nomascus leucogenys	ENSNLEG00000001440:ENSNLET00000001788
Otolemur garnettii	ENSOGAG00000032826:ENSOGAT00000029496
Pan troglodytes	XM 001166732.2
Pteropus vampyrus	ENSPVAG00000001268:ENSPVAT00000001268
Tarsius syrichta	ENSTSYG00000009008:ENSTSYT00000009006
Tupaia belangeri	ENSTBEG00000003295:ENSTBET00000003277
Tursiops truncatus	ENSTTRG00000015785:ENSTTRT00000015781
IL34	
Bos taurus	ENSBTAT00000049652
Canis familiaris	ENSCAFT00000032088
Dipodomys ordii	ENSDORT00000009844
Equus caballus	XM 001500794.1
Erinaceus europaeus	ENSEEUT00000005513
Gorilla gorilla	ENSGGOT00000032414
Homo sapiens	ENST00000288098
Loxodonta africana	ENSLAFT00000009189

Macaca mulatta	ENSMMUT00000028714
Mus musculus	ENSMUST00000076846
Nomascus leucogenys	ENSNLET00000020565
Ochotona princeps	ENSOPRT00000009881
Otolemur garnettii	ENSOGAT00000000060
Pan troglodytes	ENSPTRT00000015330
Rattus norvegicus	ENSRNOT00000023823
Spermophilus tridecemlineatus	ENSSTOT00000015368
Sus scrofa	EU872447.1
IL36A	
Bos taurus	ENSBTAG00000002087:ENSBTAT00000002694
Cavia porcellus	ENSCPOG00000015601:ENSCPOT00000015754
Echinops telfairi	ENSETEG00000011917:ENSETET00000011916
Erinaceus europaeus	ENSEEUG00000006257:ENSEEUT00000006228
Gorilla gorilla	ENSGGOG00000015991:ENSGGOT00000016039
Homo sapiens	ENSG00000136694:ENST00000259211
Loxodonta africana	ENSLAFG00000032446:ENSLAFT00000027734
Macaca mulatta	ENSMMUG00000000988:ENSMMUT00000001409
Mus musculus	ENSMUSG00000026984:ENSMUST00000028361
Ochotona princeps	ENSOPRG00000012842:ENSOPRT00000012832
Oryctolagus cuniculus	ENSOCUG00000014998:ENSOCUT00000014993
Otolemur garnettii	ENSOGAG00000016767:ENSOGAT00000016770
Pan troglodytes	ENSPTRG00000012375:ENSPTRT00000022974
Procavia capensis	ENSPCAG00000007974:ENSPCAT00000007935
Rattus norvegicus	ENSRNOG00000005722:ENSRNOT00000007540
Sorex araneus	ENSSARG00000003704:ENSSART00000003683
Spermophilus tridecemlineatus	ENSSTOG00000001732:ENSSTOT00000001728
Tarsius syrichta	ENSTSYG00000013789:ENSTSYT00000013787
IL36G	
Bos taurus	ENSBTAG00000002085:ENSBTAT00000045402
Callithrix jacchus	ENSCJAG00000013959:ENSCJAT00000027116
Canis familiaris	ENSCAFG00000007267:ENSCAFT00000011633
Dipodomys ordii	ENSDORG00000013003:ENSDORT00000013003
Echinops telfairi	ENSETEG00000011919:ENSETET00000011919
Erinaceus europaeus	ENSEEUG00000006174:ENSEEUT00000006151
Gorilla gorilla	ENSGGOG00000015985:ENSGGOT00000016036
Homo sapiens	NM 019618.2
Loxodonta africana	ENSLAFG00000022694:ENSLAFT00000022158
Macaca mulatta	ENSMMUG00000000989:ENSMMUT00000043106
Microcebus murinus	ENSMICG00000016744:ENSMICT00000016740
Mus musculus	AY071843.1
Nomascus leucogenys	ENSNLEG00000001973:ENSNLET00000002475
Ochotona princeps	ENSOPRG00000016759:ENSOPRT00000016757
Oryctolagus cuniculus	ENSOCUG00000006660:ENSOCUT00000006659
Otolemur garnettii	ENSOGAG00000016763:ENSOGAT00000016767
Pan troglodytes	ENSPTRG00000012374:ENSPTRT00000022973
Pteropus vampyrus	ENSPVAG00000010752:ENSPVAT00000010752
Rattus norvegicus	ENSRNOG00000005701:ENSRNOT00000007512
Spermophilus tridecemlineatus	ENSSTOG00000001727:ENSSTOT00000001723
Tursiops truncatus	ENSTTRG00000011809:ENSTTRT00000011811

Table 3. 4. Characterization of the amino acids possibilities for each residue identified under positive selection for each IL.

Position	Amino acid
IL1A	
104	T*, P#, I#, A#, V#, L#, S*, K*+
122	V#, M#, A#, E*-, R*+, T*, G#, L#
132	H*+, Q*, G#, S*, R*+, K*+
173	S*, V#, L#, T*, P#, K*+, I#
174	S*, A#, E*-, K*+, P#, R*+, G#
175	K*+, E*-, D*-, G#, S*, A#, L#, N*, T*
180	I#, V#, L#, D*-, F#, R*+, Y*
215	T*, K*+, R*+, P#, I#, M#
227	T*, K*+, N*, R*+, S*, I#
230	T*, S*+, N*, I#, D*-
248	Q*, H*+, G#, E*-, D*-, N*, S*, P#, V#, M#
IL1B	
8	A#, T*, I#, S*, N*, P#, F#
84	T*, P#, I#, V#, A#, S*, Y*, G#, D*-
IL2	
100	L#, P#, T*, S*, K*+, G#
101	R*+, S*, N*, A#, K*+, T*, I#, E*-, Q*, G#
124	T*, R*+, G#, S*, A#, K*+
IL4	
84	E*-, Q*, K*+, R*+, D*-, C*, N*, A#, T*, M#, G#
105	R*+, K*+, E*-, Q*, S*, D*-, M#, G#
106	F#, S*, L#, Y*, I#, C*, H*+, P#, E*-, D*-, A#
120	L#, S*, A#, M#, K*+, G#, R*+, D*-, T*, Q*
IL5	
100	G#, R*+, L#, Q*, S*, A#
IL6	
8	A#, I#, T*, D*-
46	P#, L#, T*, V#, Q*, S*
55	K*+, D*-, G#, E*-, A#, Q*, N*, I#
62	D*-, G#, R*+, W#, S*, L#, A#, K*+, M#, E*-
75	S*, Y*, H*+, N*, F#, D*-, G#, E*-
120	L#, S*, V#, M#, Q*, A#
141	R*+, N*, G#, M#, I#, E*-, K*+
145	M#, F#, D*-, T*, I#, S*, K*+, R*+, N*, Q*, L#
149	V#, G#, A#, N*, T*, I#, H*+, Y*, S*
158	A#, V#, N*, E*-, G#, M#, Q*, I#, P#, L#
161	L#, Q*, K*+, P#, A#, E*-, S*, R*+, T*
162	D*-, S*, G#, Y*, H*+, N*, V#, E*-, T*
168	D*-, E*-, N*, T*, V#, S*, K*+
174	S*, G#, N*, T*, D*-, L#, K*+, E*-,
177	T*, S*, K*+, D*-, N*, A#, E*-, P#

IL7	
38	E^{*-}, Q^{*}, K^{*+}, $G^{\#}$, S^{*}
47	Q^{*}, D^{*-}, Y^{*}, E^{*-}, $I^{\#}$
51	S^{*}, R^{*+}, N^{*}, $I^{\#}$, D^{*-},K^{*+}
53	K^{*+}, $I^{\#}$, $L^{\#}$, R^{*+}, T^{*}, N^{*}
54	E^{*-}, D^{*-}, N^{*}, K^{*+}, H^{*+}, $G^{\#}$
55	$I^{\#}$, $F^{\#}$, T^{*}, N^{*}, $V^{\#}$
92	$L^{\#}$, $F^{\#}$, $V^{\#}$, S^{*}, N^{*}
105	$L^{\#}$, S^{*}, T^{*}, E^{*-}
106	K^{*+}, T^{*}, R^{*+}, $I^{\#}$, $V^{\#}$
113	$I^{\#}$, T^{*}, Q^{*}, S^{*}, K^{*+}
IL8	
22	$G^{\#}$, D^{*-}, $A^{\#}$
30	K^{*+}, S^{*}, $G^{\#}$, $A^{\#}$, T^{*}
71	S^{*}, Q^{*}, $V^{\#}$, $I^{\#}$, $F^{\#}$, $A^{\#}$, K^{*+}, T^{*}, E^{*-}
IL9	
5	$M^{\#}$, $V^{\#}$, T^{*}, $A^{\#}$, R^{*+}, Y^{*}
24	$L^{\#}$, $F^{\#}$, H^{*+}, S^{*}, R^{*+}, T^{*}
25	$A^{\#}$, $M^{\#}$, T^{*}, $L^{\#}$, $A^{\#}$, Q^{*}, S^{*}, $W^{\#}$, $V^{\#}$, $F^{\#}$
79	T^{*}, S^{*}, $V^{\#}$, K^{*+}, $A^{\#}$
84	R^{*+}, E^{*-}, T^{*}, S^{*}, K^{*+}, $G^{\#}$
86	$P^{\#}$, $I^{\#}$, H^{*+}, S^{*}, T^{*}, $L^{\#}$
106	Y^{*}, $F^{\#}$, S^{*}, $V^{\#}$, T^{*}, $L^{\#}$, D^{*-}, N^{*}
IL10	
18	$A^{\#}$, $V^{\#}$, T^{*}, $I^{\#}$, $F^{\#}$, $L^{\#}$, S^{*}, $P^{\#}$
26	S^{*}, $L^{\#}$, $P^{\#}$, E^{*-}, $G^{\#}$, $A^{\#}$, Y^{*}, N^{*}
63	N^{*}, S^{*}, D^{*-}, T^{*}
71	$L^{\#}$, $M^{\#}$, Q^{*}
121	$L^{\#}$, $V^{\#}$, S^{*}, $M^{\#}$, $I^{\#}$, R^{*+}
IL12A	
39	H^{*+}, N^{*}, S^{*}, Q^{*}, D^{*-}, $A^{\#}$
161	$A^{\#}$, S^{*}, $G^{\#}$, T^{*}, $V^{\#}$, $M^{\#}$
IL12B	
83	$G^{\#}$, T^{*}, E^{*-}, $A^{\#}$, S^{*}, $L^{\#}$
IL14	
501	$P^{\#}$, S^{*}, T^{*}
IL15	
51	$V^{\#}$, H^{*+}, $I^{\#}$, Q^{*}, C^{*}, E^{*-}
59	K^{*+}, $I^{\#}$, R^{*+}, T^{*}
105	$A^{\#}$, T^{*}, E^{*-}, $M^{\#}$, S^{*}, $V^{\#}$, $G^{\#}$, H^{*+}
106	S^{*}, D^{*-}, T^{*}, $P^{\#}$, $A^{\#}$, N^{*}, $V^{\#}$
108	H^{*+}, N^{*}, K^{*+}, Y^{*}, S^{*}, Q^{*}, $V^{\#}$, R^{*+}, D^{*-}, E^{*-}
116	$I^{\#}$, $F^{\#}$, Y^{*}, $M^{\#}$, $L^{\#}$
IL16	
118	$L^{\#}$, $A^{\#}$, S^{*}, $P^{\#}$
188	$A^{\#}$, $V^{\#}$, $M^{\#}$, S^{*}, $L^{\#}$, T^{*}
335	$A^{\#}$, $G^{\#}$, $L^{\#}$, T^{*}, $V^{\#}$, $I^{\#}$, S^{*}
708	T^{*}, $I^{\#}$, S^{*}, $A^{\#}$
745	N^{*}, S^{*}, $G^{\#}$, $V^{\#}$, H^{*+}, R^{*+}

801	N^{*}, T^{*}, $G^{\#}$, H^{*+}, S^{*}
930	N^{*}, S^{*}, T^{*}, D^{*-}, $G^{\#}$, $A^{\#}$, R^{*+}
1075	$L^{\#}$, H^{*+}, $P^{\#}$, S^{*}, $G^{\#}$
IL17A	
18	E^{*-}, $V^{\#}$, $M^{\#}$, $A^{\#}$, $G^{\#}$, $L^{\#}$
98	$G^{\#}$, R^{*+}, S^{*}, N^{*}, $F^{\#}$
IL17C	
11	T^{*}, $P^{\#}$, $I^{\#}$, $L^{\#}$, $F^{\#}$, $A^{\#}$, $G^{\#}$
17	$L^{\#}$, $V^{\#}$, T^{*}, $I^{\#}$, $M^{\#}$
IL18	
4	E^{*-}, N^{*}, $I^{\#}$, S^{*}, $G^{\#}$, $A^{\#}$, Q^{*}
6	$V^{\#}$, $A^{\#}$, $I^{\#}$, S^{*}, $P^{\#}$
45	$L^{\#}$, $V^{\#}$, $F^{\#}$, $A^{\#}$, T^{*}
75	$P^{\#}$, $A^{\#}$, R^{*+}, S^{*}, T^{*}, $I^{\#}$, N^{*}, E^{*-}
105	E^{*-}, K^{*+}, N^{*}, S^{*}, T^{*}, $G^{\#}$, R^{*+}
180	$L^{\#}$, D^{*-}, N^{*}, C^{*}, K^{*+}, Y^{*}, $I^{\#}$, $F^{\#}$
192	E^{*-}, $P^{\#}$, Q^{*}, K^{*+}, N^{*}, $L^{\#}$
IL20	
56	K^{*+}, S^{*}, R^{*+}, E^{*-}, Q^{*}, N^{*}
62	$I^{\#}$, $V^{\#}$, $F^{\#}$, $M^{\#}$, $L^{\#}$, K^{*+}
IL21	
86	S^{*}, $P^{\#}$, $A^{\#}$, $L^{\#}$
98	$V^{\#}$, $L^{\#}$, $I^{\#}$, $M^{\#}$, S^{*}, D^{*-}
112	$A^{\#}$, T^{*}, $G^{\#}$, E^{*-}, S^{*}
119	R^{*+}, E^{*-}, S^{*}, K^{*+}, $M^{\#}$, $I^{\#}$, $G^{\#}$
IL22	
8	$V^{\#}$, $L^{\#}$, $A^{\#}$, $M^{\#}$, $G^{\#}$
53	T^{*}, $I^{\#}$, $F^{\#}$, $A^{\#}$, N^{*}, $V^{\#}$
86	S^{*}, N^{*}, R^{*+}, T^{*}, $G^{\#}$, K^{*+}, $M^{\#}$
IL25	
137	K^{*+}, $G^{\#}$, E^{*-}, S^{*}, H^{*+}, D^{*-}, R^{*+}, Q^{*}
IL26	
2	$L^{\#}$, Q^{*}, $W^{\#}$, R^{*+}
IL27B	
114	S^{*}, $A^{\#}$, $V^{\#}$, T^{*}, $I^{\#}$
IL31	
31	$L^{\#}$, $F^{\#}$, Q^{*}, $P^{\#}$, S^{*}, R^{*+}, $W^{\#}$, $G^{\#}$, $V^{\#}$
33	R^{*+}, Q^{*}, $G^{\#}$, $P^{\#}$, D^{*-}, S^{*}
38	$V^{\#}$, $I^{\#}$, R^{*+}, $L^{\#}$, E^{*-}, Q^{*}, $W^{\#}$
56	$V^{\#}$, Q^{*}, R^{*+}, K^{*+},
67	N^{*}, S^{*}, R^{*+}, E^{*-}, K^{*+}, $P^{\#}$, Q^{*}
133	C^{*}, N^{*}, Q^{*}, K^{*+}, R^{*+}, $W^{\#}$, H^{*+}, $L^{\#}$
IL34	
31	T^{*}, $G^{\#}$, R^{*+}, $L^{\#}$, $M^{\#}$, K^{*+}, $A^{\#}$, $P^{\#}$
IL36A	
2	E^{*-}, S^{*}, N^{*}, $A^{\#}$, $F^{\#}$, Q^{*}, D^{*-}
5	$L^{\#}$, S^{*}, $V^{\#}$, $F^{\#}$, K^{*+}, H^{*+}
8	D^{*-}, E^{*-}, T^{*}, K^{*+} $G^{\#}$ $I^{\#}$, $A^{\#}$, $P^{\#}$, Y^{*}
9	T^{*}, $M^{\#}$, Y^{*}, E^{*-}, K^{*+}, H^{*+}, S^{*}, $A^{\#}$, Q^{*}, $V^{\#}$

12	Q^{*}, $W^{\#}$, $P^{\#}$, R^{*+}, K^{*+}, $L^{\#}$, H^{*+}, $F^{\#}$
133	E^{*-}, K^{*+}, Q^{*}, C^{*}, T^{*}, S^{*}, $A^{\#}$
IL36G	
2	C^{*}, Y^{*}, R^{*+}, $G^{\#}$, N^{*}, D^{*-}, $A^{\#}$, S^{*}, $P^{\#}$, E^{*-}, H^{*+}, Q^{*}, $L^{\#}$
5	$I^{\#}$, $G^{\#}$, Q^{*}, H^{*+}, R^{*+}, $W^{\#}$, $V^{\#}$, Y^{*}, D^{*-}, $I^{\#}$, E^{*-}, K^{*+}
6	T^{*}, N^{*}, $V^{\#}$, Y^{*}, $L^{\#}$, S^{*}, $F^{\#}$, $P^{\#}$
104	$A^{\#}$, $V^{\#}$, $L^{\#}$, D^{*-}, S^{*}, Q^{*}, $G^{\#}$, $M^{\#}$, T^{*}

Polarity: *hydrophilic; $^{\#}$hydrophobic

Charge: $^{+}$positive; $^{-}$negative

CONVERGENT EVOLUTION OF IL6 IN TWO LEPORIDS (*ORYCTOLAGUS* AND *PENTALAGUS*) ORIGINATED AN EXTENDED PROTEIN

Fabiana Neves, Joana Abrantes, Ana Pinheiro, Tereza Almeida, Paulo P Costa, Pedro J Esteves

1. ABSTRACT

Interleukin 6 (IL6) is a class-I helical cytokine with a broad spectrum of biological activities and a gene structure conserved throughout vertebrates, with 5 coding exons. IL6 from European rabbits belonging to the subspecies *Oryctolagus cuniculus cuniculus,* was previously shown to differ from the other mammals by extending an additional 27 amino acids. However, in other leporids (*Sylvilagus* spp. and *Lepus* spp.) that diverged from the European rabbit ~12 million years ago this mutation was not present. In this study we extended the study of IL6 for the *O. c. algirus* subspecies and five additional lagomorphs' genera (*Brachylagus, Bunolagus, Pentalagus, Romerolagus* and *Ochotona*). We confirmed the presence of the mutated stop codon in both *O. c. cuniculus* and *O. c. algirus*. We found that the typical stop codon is present in *S. bachmani* and *L. europaeus*, in agreement with previous reports, but also in *Bunolagus, Brachylagus* and *Ochotona.* Remarkably, in *Pentalagus* we detected a deletion of the stop codon causing an extension of IL6 for 17 extra residues. Our results indicate that the IL6 extension in those species occurred by two independent events: one occurred between 2 and 8 million years ago in the ancestral of the *Oryctolagus* subspecies, and the other occurred in a *Pentalagus* ancestral at a maximum of 9 million years ago. The absence of this IL6 extension in *Bunolagus*, sister genus of *Oryctolagus*, shows that this evolutionary event happened by convergence suggesting some functional relevance.

Keywords: Interleukin 6 (IL6), Lagomorphs, Convergent evolution, *Oryctolagus*, *Pentalagus*

2. INTRODUCTION

Interleukins are secreted proteins involved in multiple biological functions in the immune system that exert their function by binding to specific receptors. Interleukin 6 (IL6), also known as interferon beta 2, is a class-I helical cytokine with a broad spectrum of biological activities in immune regulation, hematopoiesis, inflammation and oncogenesis, among others (Akdis et al., 2011; Huising et al., 2006; Kishimoto, 2010). This protein is produced after stimulation of several differents cells (T, B, smooth muscle cells, eosinophils, granolocytes, etc) with mononuclear phagocytes as the main source (Akdis et al., 2011; Huising et al., 2006; Tanabe et al., 1988). Essential for Th17 differentiation, IL6 has a gene structure conserved throughout vertebrates, with 5 coding exons (Huising et al., 2006; Tanabe et al., 1988). Characterized by 4 long α helices (A, B, C and D), this immunoregulatory cytokine is able to activate a cell surface signaling assembly composed by IL6, IL6Rα (which interacts with low affinity with the D helix of IL6) and the shared signaling receptor gp130 that interacts with high affinity with IL6 in sites located in the A and C helices. IL6 first interacts with the extracelular region of IL6Rα, and then, this complex is able to associate with gp130 and initiate signaling (Akdis et al., 2011; Kalai et al., 1997; Savino et al., 1994). Furthermore, it has been shown that in mammals IL6 has been evolving under strong positive selection (Neves et al., 2014).

In the European rabbit, IL6 is involved in the immune response against a highly fatal and contagious viral disease, the rabbit hemorrhagic disease (RHD) (Abrantes et al., 2012; Marques et al., 2012; Teixeira et al., 2012). Previously, IL6 from European rabbits belonging to the subspecies *Oryctolagus cuniculus cuniculus,* was shown to differ from other mammals IL6 by extending an additional 27 amino acids (Perkins et al., 2000). A mutation in the typical stop codon, conserved in other mammals, into a glutamate encoding codon caused the elongation of IL6. However, in other leporids (*Sylvilagus* spp. and *Lepus* spp.) that diverged from the European rabbit ~12 million years ago (mya) (Esteves et al., 2005; Matthee et al., 2004), this mutation was not present. The family Leporidae, (Figure 3.2) includes 11 genera (*Brachylagus*, *Bunolagus*, *Caprolagus*, *Lepus*, *Nesolagus*, *Oryctolagus*, *Pentalagus*, *Poelagus*,

Pronolagus, *Romerolagus* and *Sylvilagus*) of which three are closely related to the *Oryctolagus* genus: *Bunolagus* is thought to have diverged from *Oryctolagus* at ~7mya, *Caprolagus* at ~8 mya and *Pentalagus* at ~9 mya (Matthee et al., 2004; Pinheiro et al., 2011).

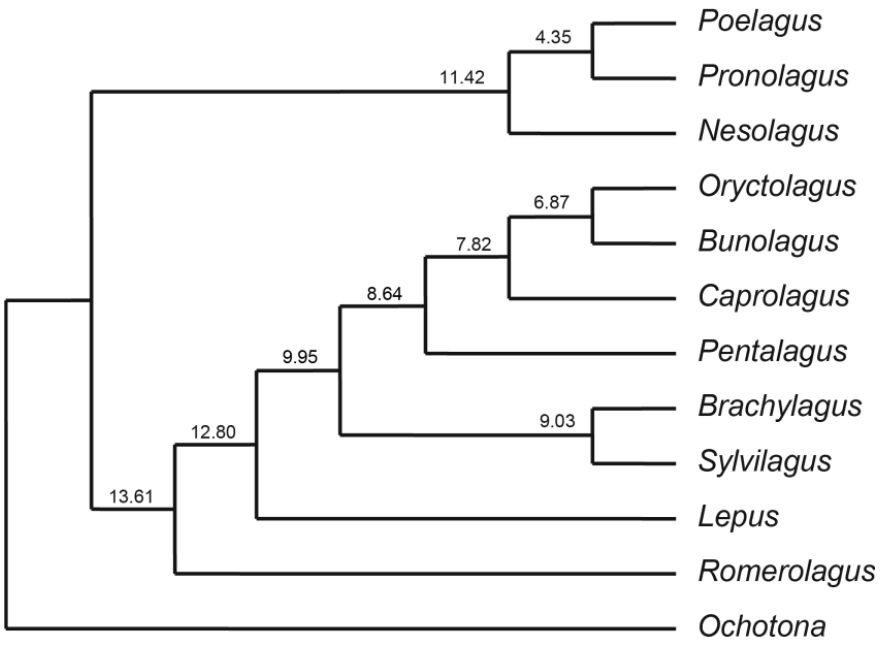


Figure 3. 2. Evolutionary topology showing the relationships within the order Lagomorpha based on a molecular super matrix. Adapted from Matthee et al., 2004.

The aim of this study was to determine the evolution of IL6 in other lagomorphs, *Oryctolagus cuniculus algirus, Brachylagus idahoensis, Sylvilagus bachmani*, *Lepus europaeus, Bunolagus monticularis, Pentalagus furnessi, Romerolagus diazi* and *Ochotona princeps.*

3. MATERIALS AND METHODS

The IL6 gene was PCR-amplified with primers designed (Table 3.5) according to the European rabbit and *O. princeps* IL6 sequences available in public databases and sequenced for eight lagomorphs (conditions are available upon request): *Oryctolagus cuniculus* (subspecies *O. c. cuniculus, and O. c. algirus), Brachylagus idahoensis, Bunolagus monticularis, Pentalagus furnessi, Sylvilagus bachmani*, *Lepus europaeus*, *Romerolagus diazi* and *Ochotona*

princeps. All obtained sequences were submitted to GenBank and the following accession numbers provided: KJ939512-KJ939527.

The obtained IL6 sequences were translated and compared with other mammalian IL6 sequences retrieved from public databases (GenBank, Ensembl and Uniprot). Sequences were aligned using ClustalW implemented in the software program BioEdit version 7.1.3 (Hall, 1999) and adjusted manually.

For secondary structure prediction we used PsiPred (http://bioinf.cs.ucl.ac.uk/psipred/) (Buchan et al., 2013; Jones, 1999) and DiAminoacid Neural Network Application (DiANNA) (http://clavius.bc.edu/~clotelab/DiANNA/) (Ferre and Clote, 2006). These methods predict protein cysteines that create permanent structural disulfide bonds. Psipred uses Position Specific Iterated – BLAST (PSI-BLAST) searches of the non-redundant protein sequence database to obtain evolutionary information used to predict the secondary structure of the query protein. DiANNA is a neural network trained to recognize cysteines in an oxidized state (sulfur covalently bonded) from those in a reduced state.

Putative *N*-glycosylation sites were predicted using NetNGlyc 1.0 server available at http://www.cbs.dtu.dk/services/NetNGlyc/ (Gupta et al., 2004).

4. RESULTS/DISCUSSION

In this study we amplified and sequenced the IL6 gene for eight different lagomorphs belonging to two families, *Leporidae* (rabbits and hares) and *Ochotonidae* (pikas), that separated ~35 mya (Matthee et al., 2004). We further compared our sequences to sequences of IL6 from representatives of most mammalian groups available in online databases (e.g. Artiodactyla, Carnivores, Chiroptera, Primates, Rodents, etc).

Table 3. 5. List of primer pairs used by using a) the genomic DNA (gDNA) as template (positions are according to OryCun2.0:10:7771292:7776103:1 for Leporids and pika:scaffold_6618:68359:73623:1n for *O. princeps*), b) cDNA as template (positions are according to NM_001082064.2 for Leporids and XM_004582561.1 for *O. princeps*.

a)

	Primers sequence (5'- 3')	Primer name	Position in gDNA (base pair)	Exons amplified	Species amplified
Leporids	CGACGTCACATTGCACAATC	IL6OrcuF1	134-153	Exon1 + exon2	*O.c.cuniculus, O.c.algirus, L. europaeus, S. bachmani, , B.idahoensis, P. furnessi, R.diazi*
	CAGGGTGTTCCTCTCTTCTG	IL6OrcuR1	849-868		
	CGACTATGAACTCCTTCAC	IL6OrcuF1.1	352-371		
	GATTCTCACGTCTCGCTTCTC	IL6OrcuR1.1	941-961		
	CTTCAGCATGAGCGGTTCTG	IL6OrcuF2	1675-1694	Exon 3 + exon 4	*O.c.cuniculus, O.c.algirus, B.idahoensis, S. bachmani, L. europaeus, P. furnessi, B. monticularis, R.diazi*
	GCTGAGGAGAGTTCAGTTG	IL6OrcuR2	3024-3042		
	GCCCTCTAGTGGTGTTTC	IL6OrcuF2a	1646-1663		
	CAAAGGTGGGTGCTCCTTS	IL6OrcuR2a	2925-2943		
	CTTCCACTGAGAGGATGTGTG	IL6OrcuF3	4479-4499	Exon 5	*O.c.cuniculus, O.c.algirus, B.idahoensis, S. bachmani, L. europaeus, P. furnessi, B. monticularis, R.diazi*
	CTAGGAAGATGAGCGTTAGGAC	IL6OrcuR3	4807-4828		
O.princeps	GCTAAGGCTCATTCTGCCTCT	IL6OcprF1	6-25	Exon1 + exon 2	*O.princeps*
	CTTCTGCAGTTCTCTGCAGC	IL6OcprR1	658-677		
	CTTCTAGCTGAGCATGAACTGG	IL6OcprF2	1474-1495	Exon 3 + exon 4	
	CTAGAGGAGCTCAATCGTCAG	IL6OcprR2	2812-2832		
	CACTGCAGAACTCACCCTAAC	IL6OcprF3	3670-3690	Exon 5	
	GAATACGCCATCCTGTCCAC	IL6OcprR3	4029-4048		

b)

Primers sequence (5'- 3')	Primer name	Position in cDNA (base pair)	Species amplified
ATGAACTCCTTCACAAGCG	IL6OrcuFcDNA	1-19	Leporids
CTGTCCATTGGACACATAA	IL6OrcuRcDNA	720-738	
GCCGTGGGCCTGATGTAG	IL6RcDNA	640-657	
ATGAACTCCGTCTTCGCAAG	IL6OcprFcDNA	1-20	*O.princeps*
GCGGTGCGCCTGATGTAG	IL6OcprRcDNA	634-651	

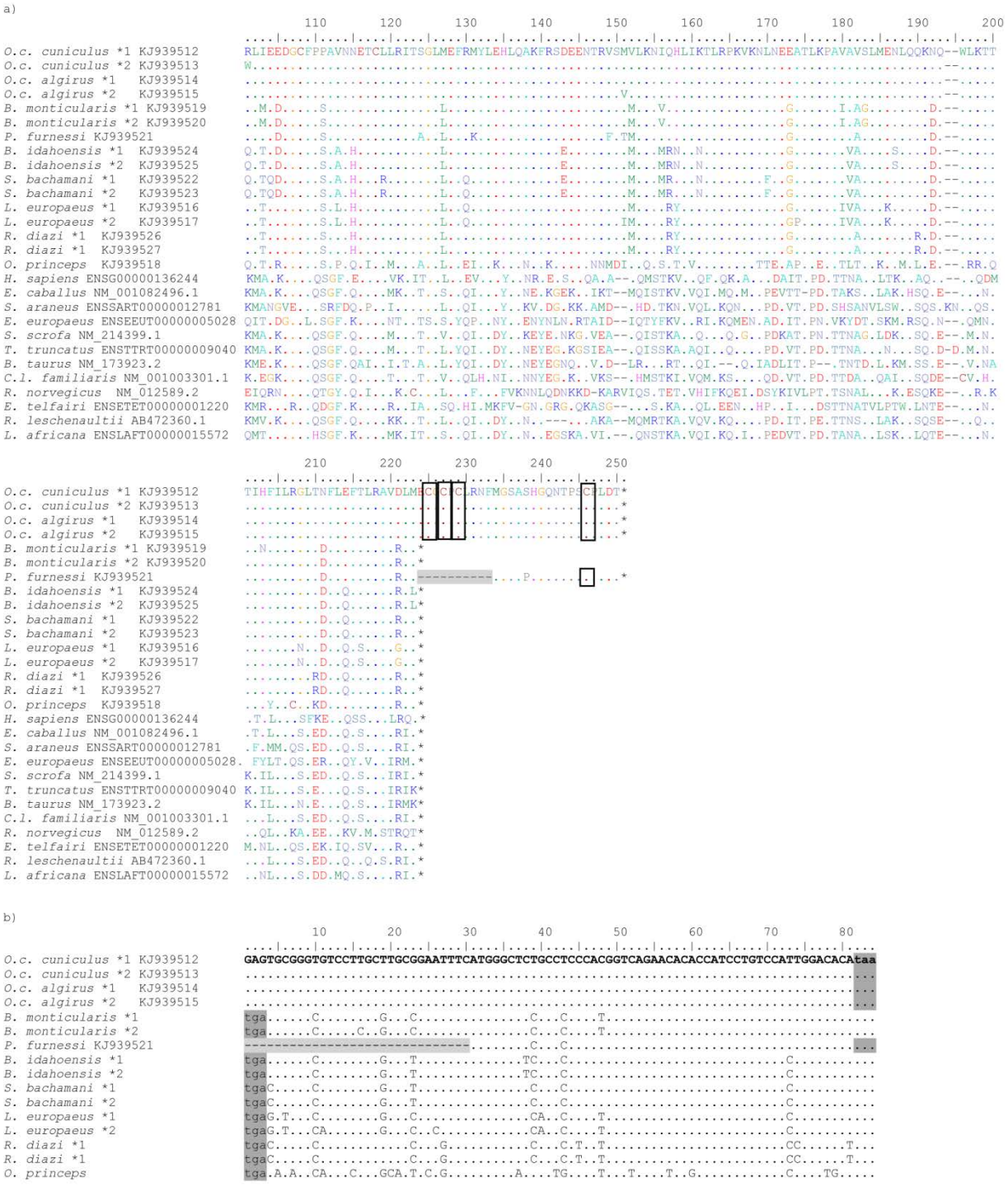


Figure 3. 3. **a)** Alignment of IL-6 for the different mammalian species (GenBank, Ensembl and Uniprot reference numbers are given for the retrieved sequences). The extra cysteines in the extended IL-6 are boxed in black. (*) represent stop codons **b)** Alignment of the 3'UTR of *IL-6* for the different lagomorphs studied. The stop codons are in lower-case letters and boxed in dark grey. *1 and *2 represent alleles; (.) represent identity with the reference sequence (*O. c. cuniculus* *1); (-) represent deletions; *P. furnessi* deletion is boxed in grey.

Our results confirmed the presence of the mutated stop codon in both *O. cuniculus* subspecies (*O. c. cuniculus* and *O. c. algirus*) causing an extension of 27 amino acids in the IL6 protein (Figure 3.3). In agreement with what has been

reported by Perkins et al., 2000, we observed that the typical stop codon is present in *Sylvilagus* and *Lepus*. We further observed the presence of the expected stop codon for *Brachylagus idahoensis, Bunolagus monticularis*, *Romerolagus diazi* and *Ochotona princeps.* Interestingly, for *Pentalagus furnessi* we found a 10 amino acid deletion starting at the typical stop codon.

This deletion extends the IL6 coding sequence for 17 extra amino acids. Therefore, *P. furnessi* as in *Oryctolagus* spp. sequences translation of IL6 continues into the sequence corresponding to the 3' UTR in other leporids and stops in the next in-frame stop codon (81 and 51 nucleotides downstream for *Oryctolagus* spp. and *P. furnessi*, respectively). The extension of IL6 in *Oryctolagus* and *Pentalagus,* but not in *Bunolagus,* shows that two independent events led to a longer IL6 protein: i) in *Oryctolagus* spp. this extension is due to a mutation of the stop codon into a glutamate residue (E), and ii) in *Pentalagus* is due to the deletion of the stop codon. The mutation of the stop codon into a glutamate residue (E) occurred at least 2 mya in the ancestral of the two *Oryctolagus* subspecies (Carneiro et al., 2009) and at most at 7 mya after *Oryctolagus* diverged from all other leporids (Matthee et al., 2004). The deletion of the stop codon occurred at a maximum of ~9 million years ago when *P. furnessi* diverged from the other leporids (Matthee et al., 2004). Interestingly, in lagomorphs the IL6 3'UTR is well conserved, as had already been described for other species (Dassi et al., 2013; Duret et al., 1993; Paschoud et al., 2006). 3'UTR is known to be important for translation, stabilization and localization of mRNA, and, alterations in this region, namely mutations in the stop codon, can cause translation de-regulation and may be associated with several diseases and their progressions (Chatterjee and Pal, 2009; Lau et al., 2010; Soifer et al., 2007; Szostak and Gebauer, 2013). This region is also described to be important for the interaction between mRNA and related-proteins, and therefore, alteration in this region may alter these interactions leading to changes in structure and function of the gene (Chatterjee and Pal, 2009; Lau et al., 2010; Soifer et al., 2007; Szostak and Gebauer, 2013). Thus, while mutation of the stop codon and extension of IL6 might interfere with its functions, these are not expected to be negatively affected as they appeared independently in two leporid species.

In addition to these alterations, some specific differences were also observed: for *O. princeps*, 3 amino acid insertions (Val4, Ala31 and Pro32) and 1 deletion at position 53 were detected, and for *L. europaeus* and *B. idahoensis* there were 3 shared amino acid insertions (Tyr65, Ile66, Leu67).

IL6 has 4 conserved cysteines (Cys) that form two disulfide bonds, Cys72-Cys78 and Cys101-Cys111 (amino acid residues were numbered from the first human methionine residue, with signal peptide included in the numbering). The first bond is known to contribute minimally to IL6 activity, while the second is considered to be essential for human IL6 activity (Boulanger et al., 2003; Breton et al., 1995; Huising et al., 2006; Savino et al., 1993; Simpson et al., 1997; Somers et al., 1997; Xu et al., 1997). Additional cysteine residues are found in other mammalian species: *Rousettus leschenaultia* Cys5, *Dipodomys ordii* Cys10, *Rattus norvegicus* Cys116, *Canis lupus familiaris* Cys185 and *Saimiri sciureus* Cys204. For *O. princeps* we observed two extra cysteines: Cys81 and Cys196. Moreover, all studied leporids share a cysteine, Cys14, located in the signal peptide. Additionally, in the 27 amino acid extension *O. cuniculus* has four extra cysteines, Cys214, 216, 218 and 235, and in *P. furnessi* there are two extra cysteine residues, Cys48 and Cys235 (located in the 17 amino acid extension). The results of the analysis of protein secondary structure, using the PsiPred (Buchan et al., 2013; Jones, 1999) and DiANNA (Ferre and Clote, 2006) and the comparison with the human IL6 sequence, revealed a predicted extra helix in the extension of *O. cuniculus* (from Cys218 to Phr222 included) and that the nine cysteine residues that exist in the coding sequence would form four predicted bonds, three of which do not match the disulfide bonds present in the human IL6 (Cys14-Cys78, Cys72-Cys216, Cys101-Cys111 and Cys214-Cys235). As for *Pentalagus*, the 17 amino acid extension increases the length of the 3' terminal coil and the seven cysteine residues are predicted to form the following bonds: Cys14-Cys48, Cys72-Cys78 and Cys101-Cys235. For the regular IL6 in *O. princeps* three cysteine bonds are predicted, the two links present in the human sequence and another between Cys81-Cys196. Disulfide bonds are considered to be the strongest link in proteins, being important for protein folding and stability (Petersen et al., 1999; Simpson et al., 1997). According to some authors these bonds are also associated with functional differentiation of the proteins (Fass, 2012; Li et al.,

2011). Although the biological implications of the predicted secondary structure remain to be assessed, the four additional cysteines present in the 3' sequence of *O. cuniculus* may have some functional relevance, with implications in the IL6 protein conformation and its interaction with receptors.

Glycosylation is also important for protein structure and function. Indeed, glycosylation is involved in diverse biological processes ranging from protein folding to immune response (Helenius and Aebi, 2004; Rudd et al., 2001). Our search for putative N-glycosylation (Asn-X-Ser/Thr motifs where X can be any amino acid except proline) sites in Lagomorphs IL6 showed that, unlike human IL6 that has 2 potential N-glycosylation sites at Asn73 and Asn172 (Breton et al., 1995; Simpson et al., 1997; Xu et al., 1997), some Lagomorphs have only 1 potential N-glycosylation site (*O. princeps* – Asn89, *P. furnessi* - Asn108, *O. cuniculus* and *P. furnessi* - Asn108; *L. europaeus* – Asn197), while others (*Sylvilagus bachmani* and *Brachylagus idahoensis*) have none.

5. CONCLUSIONS

Our results show the extension of IL6 in two different lagomorph genera: in *Oryctolagus* in which the extension of 27 amino acids is due to a mutation of the stop codon, and in *Pentalagus* in which the extension of 17 amino acids is due to a deletion of the typical stop codon. These events occurred independently in these two lineages and are the outcome of a process of convergent evolution to have a longer IL6. Biological implications of this extension remain to be assessed but its occurrence in two independent genera might suggest some functional relevance. Further functional and structural studies should be performed to fully understand the impact of these alterations in the IL6 of *Oryctolagus* and *Pentalagus*.

6. REFERENCES

Abrantes, J., van der Loo, W., Le Pendu, J., Esteves, P.J., 2012. **Rabbit haemorrhagic disease (RHD) and rabbit haemorrhagic disease virus (RHDV): a review**. Vet Res 43, 12.

Akdis, M., Burgler, S., Crameri, R., Eiwegger, T., Fujita, H., Gomez, E., Klunker, S., Meyer, N., O'Mahony, L., Palomares, O., Rhyner, C., Ouaked, N., Schaffartzik, A., Van De Veen, W., Zeller, S., Zimmermann, M., Akdis, C.A., 2011. **Interleukins, from 1 to 37, and interferon-gamma: receptors, functions, and roles in diseases**. J Allergy Clin Immunol 127, 701-721 e701-770.

Boulanger, M.J., Chow, D.C., Brevnova, E.E., Garcia, K.C., 2003. **Hexameric structure and assembly of the interleukin-6/IL-6 alpha-receptor/gp130 complex**. Science 300, 2101-2104.

Breton, J., La Fiura, A., Bertolero, F., Orsini, G., Valsasina, B., Ziliotto, R., De Filippis, V., Polverino de Laureto, P., Fontana, A., 1995. **Structure, stability and biological properties of a N-terminally truncated form of recombinant human interleukin-6 containing a single disulfide bond**. Eur J Biochem 227, 573-581.

Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K., Jones, D.T., 2013. **Scalable web services for the PSIPRED Protein Analysis Workbench**. Nucleic Acids Res 41, W349-357.

Carneiro, M., Ferrand, N., Nachman, M.W., 2009. **Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (Oryctolagus cuniculus)**. Genetics 181, 593-606.

Chatterjee, S., Pal, J.K., 2009. **Role of 5'- and 3'-untranslated regions of mRNAs in human diseases**. Biology of the cell/under the auspices of the European Cell Biology Organization 101, 251-262.

Dassi, E., Zuccotti, P., Leo, S., Provenzani, A., Assfalg, M., D'Onofrio, M., Riva, P., Quattrone, A., 2013. **Hyper conserved elements in vertebrate mRNA 3'-UTRs reveal a translational network of RNA-binding proteins controlled by HuR**. Nucleic Acids Res 41, 3201-3216.

Duret, L., Dorkeld, F., Gautier, C., 1993. **Strong conservation of non-coding sequences during vertebrates evolution: potential involvement in post-transcriptional regulation of gene expression**. Nucleic Acids Res 21, 2315-2322.

Esteves, P.J., Lanning, D., Ferrand, N., Knight, K.L., Zhai, S.K., van der Loo, W., 2005. **The evolution of the immunoglobulin heavy chain variable region (IgVH) in Leporids: an unusual case of transspecies polymorphism**. Immunogenetics 57, 874-882.

Fass, D., 2012. **Disulfide bonding in protein biophysics**. Annu Rev Biophys 41, 63-79.

Ferre, F., Clote, P., 2006. **DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification**. Nucleic Acids Res 34, W182-185.

Gupta, R., Jung, E., Brunak, S., 2004. **Prediction of N-glycosylation sites in human proteins**. in press.

Hall, T.A., 1999. **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT**. Nucl. Acids. Symp. Ser. 41, 95-98.

Helenius, A., Aebi, M., 2004. **Roles of N-linked glycans in the endoplasmic reticulum**. Annual review of biochemistry 73, 1019-1049.

Huising, M.O., Kruiswijk, C.P., Flik, G., 2006. **Phylogeny and evolution of class-I helical cytokines**. The Journal of endocrinology 189, 1-25.

Jones, D.T., 1999. **Protein secondary structure prediction based on position-specific scoring matrices**. J Mol Biol 292, 195-202.

Kalai, M., Montero-Julian, F.A., Grotzinger, J., Fontaine, V., Vandenbussche, P., Deschuyteneer, R., Wollmer, A., Brailly, H., Content, J., 1997. **Analysis of the human interleukin-6/human interleukin-6 receptor binding interface at the amino acid level: proposed mechanism of interaction**. Blood 89, 1319-1333.

Kishimoto, T., 2010. **IL-6: from its discovery to clinical applications**. International immunology 22, 347-352.

Lau, A.G., Irier, H.A., Gu, J., Tian, D., Ku, L., Liu, G., Xia, M., Fritsch, B., Zheng, J.Q., Dingledine, R., Xu, B., Lu, B., Feng, Y., 2010. **Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF)**. Proc Natl Acad Sci U S A 107, 15945-15950.

Li, X.Q., Zhang, T., Donnelly, D., 2011. **Selective loss of cysteine residues and disulphide bonds in a potato proteinase inhibitor II family**. PLoS One 6, e18615.

Marques, R.M., Costa, E.S.A., Aguas, A.P., Teixeira, L., Ferreira, P.G., 2012. **Early inflammatory response of young rabbits attending natural resistance to calicivirus (RHDV) infection**. Veterinary immunology and immunopathology 150, 181-188.

Matthee, C.A., van Vuuren, B.J., Bell, D., Robinson, T.J., 2004. **A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene**. Syst Biol 53, 433-447.

Neves, F., Abrantes, J., Steinke, J.W., Esteves, P.J., 2014. **Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals**. Innate immunity 20, 184-191.

Paschoud, S., Dogar, A.M., Kuntz, C., Grisoni-Neupert, B., Richman, L., Kuhn, L.C., 2006. **Destabilization of interleukin-6 mRNA requires a putative RNA stem-loop structure, an AU-rich element, and the RNA-binding protein AUF1**. Molecular and cellular biology 26, 8228-8241.

Perkins, H.D., van Leeuwen, B.H., Hardy, C.M., Kerr, P.J., 2000. **The complete cDNA sequences of IL-2, IL-4, IL-6 AND IL-10 from the European rabbit (Oryctolagus cuniculus)**. Cytokine 12, 555-565.

Petersen, M.T., Jonson, P.H., Petersen, S.B., 1999. **Amino acid neighbours and detailed conformational analysis of cysteines in proteins**. Protein engineering 12, 535-548.

Pinheiro, A., Lanning, D., Alves, P.C., Mage, R.G., Knight, K.L., van der Loo, W., Esteves, P.J., 2011. **Molecular bases of genetic diversity and evolution of the immunoglobulin heavy chain variable region (IGHV) gene locus in leporids**. Immunogenetics 63, 397-408.

Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A., Dwek, R.A., 2001. **Glycosylation and the immune system**. Science 291, 2370-2376.

Savino, R., Lahm, A., Giorgio, M., Cabibbo, A., Tramontano, A., Ciliberto, G., 1993. **Saturation mutagenesis of the human interleukin 6 receptor-binding site: implications for its three-dimensional structure**. Proc Natl Acad Sci U S A 90, 4067-4071.

Savino, R., Lahm, A., Salvati, A.L., Ciapponi, L., Sporeno, E., Altamura, S., Paonessa, G., Toniatti, C., Ciliberto, G., 1994. **Generation of interleukin-6 receptor antagonists by molecular-modeling guided mutagenesis of residues important for gp130 activation**. The EMBO journal 13, 1357-1367.

Simpson, R.J., Hammacher, A., Smith, D.K., Matthews, J.M., Ward, L.D., 1997. **Interleukin-6: structure-function relationships**. Protein Sci 6, 929-955.

Soifer, H.S., Rossi, J.J., Saetrom, P., 2007. **MicroRNAs in disease and potential therapeutic applications**. Molecular therapy : the journal of the American Society of Gene Therapy 15, 2070-2079.

Somers, W., Stahl, M., Seehra, J.S., 1997. **1.9 A crystal structure of interleukin 6: implications for a novel mode of receptor dimerization and signaling**. The EMBO journal 16, 989-997.

Szostak, E., Gebauer, F., 2013. **Translational control by 3'-UTR-binding proteins**. Briefings in functional genomics 12, 58-65.

Tanabe, O., Akira, S., Kamiya, T., Wong, G.G., Hirano, T., Kishimoto, T., 1988. **Genomic structure of the murine IL-6 gene. High degree conservation of potential regulatory sequences between mouse and human**. J Immunol 141, 3875-3881.

Teixeira, L., Marques, R.M., Aguas, A.P., Ferreira, P.G., 2012. **Regulatory T cells are decreased in acute RHDV lethal infection of adult rabbits**. Veterinary immunology and immunopathology 148, 343-347.

Xu, G.Y., Yu, H.A., Hong, J., Stahl, M., McDonagh, T., Kay, L.E., Cumming, D.A., 1997. **Solution structure of recombinant human interleukin-6**. J Mol Biol 268, 468-481.

Fabiana Neves, Joana Abrantes, Tereza Almeida, Paulo P Costa, Pedro J Esteves

1. ABSTRACT

In leporids, IL17A have been implicated in the host defense against extracellular pathogens, such as *Francisella tularensis* that infects hares and rabbits and causes the zoonotic disease tularemia. Here, we studied IL17A from five lagomorphs, European rabbit, pygmy rabbit, brush rabbit, European brown hare and American pika. We observed that this protein is highly conserved between these species, with a similarity of 97-99% in leporids and ~88% between leporids and American pika. The exon/intron structure, N-glycosylation sites and cysteine residues are conserved between lagomorphs. However, at codon 88, one of the interaction sites between IL17A and it receptor IL17RA there is an Arg>Pro mutation that only occurs in European rabbit and European brown hare. This could induce critical alterations in the IL17A structure, conformation and consequently modify its function. The differences observed between leporids and humans or rodents might also represent important alterations in protein structure and function. In addition, as for other interleukins, IL17A sequences of human and European rabbit are more closely related than the sequences of human and mouse or European rabbit and mouse. This study gives further support to the hypothesis that European rabbit might be a more suitable animal model for studies on human IL17.

Keywords: Immune system, IL17, lagomorphs

2. INTRODUCTION

Interleukin 17, first known as cytotoxic T lymphocyte associated antigen (CTLA)-8, is originated from a T-cell derived factor with cytokine-like activity (Rouvier et al., 1993; Yao et al., 1995). With a ubiquitous expression in different

tissues, this protein, nowadays known as IL17A, has a sequence composition different from all the other cytokine families (Gaffen, 2008; Yao et al., 1995). IL17A, along with five functional homodimers (IL17B-F), one heterodimer (IL17A/F) and 5 receptors (IL17RA-RE), compose the IL17 family, which is important to adaptive immunity responses, namely as mediator of chronic inflammation and autoimmune diseases (Ely et al., 2009; Gaffen, 2008; Gerhardt et al., 2009; Ouyang et al., 2008). There is a wide range of genes that are targeted by IL17, such as pro-inflammatory and hematopoietic cytokines, genes associated with acute phase response and anti-microbial substances (Gaffen, 2008; Shen and Gaffen, 2008). This protein is also part of a subset of CD4 T helper (Th) cells known as Th17 which are able to establish a connection between innate and adaptive immune responses, being a complement to Th1 and Th2 defense mechanisms (Hemdan et al., 2010). Furthermore, the production of IL17A is important for host defense against extracellular pathogens (fungi, viruses, bacteria and parasites) assisting in neutrophils recruitment and activation and also promoting antimicrobial peptides (Bettelli et al., 2007; Hemdan et al., 2010; Ishigame et al., 2009; Iwakura et al., 2011; van de Veerdonk et al., 2009). Studies in mice (Huang et al., 2004; Ishigame et al., 2009; Khader and Gopal, 2010; Skyberg et al., 2013) and humans (Paranavitana et al., 2010; Tesmer et al., 2008; Zhu and Qian, 2012) highlighted the importance of IL17 expressing cells for immunity against several diseases, and low expression levels of IL17 and IL17RA make organisms more susceptible to disease, including those caused by extracellular pathogens such as *Francisella tularensis.*

F. tularensis is a highly pathogenic Gram negative intracellular bacteria included by the Center of Disease Control and Prevention (CDC) into the category A of bioterrorism (http://emergency.cdc.gov/agent/agentlist.asp). Able to cause the zoonotic disease tularemia, this microorganism has several known hosts, from mammals to protozoans; however transmission to humans is normally associated with direct contact with lagomorphs, rodents and some arthropods (Carvalho et al., 2014; Kubelkova and A., 2015; Skyberg et al., 2013; Valentino et al., 2010). In lagomorphs and rodents, *F. tularensis* has the ability to cause septicemia while in humans the outcome of infection is a multi-system organ failure (Kim et al., 2010). There are several reports of *F.*

tularensis infections in leporids, mainly in rabbits (European rabbit and cottontails) (Lepitzki et al., 1990; Mailles and Vaillant, 2014; Wobeser et al., 2009) and hares (Mailles and Vaillant, 2014; Rijks et al., 2013; Wobeser et al., 2009) and despite an apparent period of stasis (2006-2010), there were some recently documented outbreaks of tularemia in Europe (Carvalho et al., 2014; Hestvik et al., 2015).

The order Lagomorpha includes two families, Leporidae (rabbits and hares) with eleven genera and Ochotonidae (pikas) with only one genus, *Ochotona* (Matthee et al., 2004). Together with rodents, lagomorphs form the clade Glires, a sister group of Euarchonta that includes primates (Horner et al., 2007; Murphy et al., 2001). Along with mouse, the European rabbit had been used as a research model for several human diseases, development of therapeutics and vaccines (Schnupf and Sansonetti, 2012). Several studies have suggested that the European rabbit may be a better research model than mouse (Fischer et al., 2012; Neves et al., 2015; Pinheiro et al., 2016; Vaure and Liu, 2014; Vuillaumier et al., 1997). With the exception of humans and mouse, there is a big gap of information on IL17A in other mammalian groups, including leporids. Thus, considering the important biological role of the European rabbit immune response against several diseases, including tularemia, we performed a genetic characterization of IL17A in four leporid genera (*Oryctolagus*, *Brachylagus*, *Sylvilagus*, *Lepus*).

3. MATERIALS AND METHODS

Samples of European rabbit (*Oryctolagus cuniculus cuniculus and Oryctolagus cuniculus algirus),* pygmy rabbit (*Brachylagus idahoensis),* brush rabbit (*Sylvilagus bachmani*) and European brown hare (*Lepus europaeus)* were provided by the CIBIO Lagomorpha tissue collection. Genomic DNA (gDNA) was extracted using the EasySpin Genomic DNA Minipreps Tissue Kit (Citomed, Torun, Poland) according to the manufacturer's instructions. Total RNA was extracted by using the RNeasy Mini Kit also according to the manufacturer's instructions (Qiagen, Hilden, Germany) from one specimen of European rabbit and one of European brown hare. Complementary DNA (cDNA) was synthesized using oligo(dT) as primers and SuperScript III reverse

transcriptase (Invitrogen, Carlsbad, CA, USA). The European rabbit and American pika IL17A sequences were retrieved from public databases (accession numbers are given in bold in Figure 1). PCR amplification was performed with the Multiplex PCR Kit (Qiagen) by using two pairs of primers designed according to the retrieved sequences (for genomic DNA F1- CGTCCAACCTCAGTTGATC + R1- CACTGTACCATCTATCCTGC and F2- CCTTCATTTACTCCCATTCG + R2- CATCCATCACATGGCCTAA; for cDNA the combination of primers F1+R2 was used). Sequencing was performed on an ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems, Foster City, CA, USA) and PCR products were sequenced in both directions. The sequences obtained were submitted to GenBank with the following accession numbers: KU163611-KU163619.

Haplotype phases of the sequences obtained were reconstructed with the program PHASE, built into the software DnaSP (Librado and Rozas, 2009). Multiple Sequence Comparison by Log-Expectation (MUSCLE; http://www.ebi.ac.uk/) (Edgar, 2004) was used for sequence alignment. The putative N-glycosylation sites were predicted using NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) (Gupta et al., 2004).

The number of nucleotide differences per site between sequences was estimated in MEGA6 (Tamura et al., 2013) with the following options: bootstrap method (1000 replicates), p-distance as model and pairwise deletion for gaps/missing data treatment. A Maximum Likelihood approach was used to estimate the phylogenetic relationships between the IL17A nucleotide sequences by using MEGA6; the best-fit nucleotide substitution model was predicted by the same software and 1000 bootstrap replicates were used.

The secondary structure of IL17A was predicted using PsiPred (http://bioinf.cs.ucl.ac.uk/psipred/) (Buchan et al., 2013; Jones, 1999) and DiAminoacid Neural Network Application (DiANNA) (http://clavius.bc.edu/~clotelab/DiANNA/) (Ferre and Clote, 2006). Both methods predict protein cysteines that create disulfide bonds, but while PsiPred uses Position Specific Iterated – BLAST (PSI-BLAST) to obtain evolutionary information used to predict the secondary structure of the query protein, DiANNA is a neural network that recognize cysteines in an oxidized state (sulfur covalently bonded) distinguishing them from those in a reduced state.

4. RESULTS AND DISCUSSION

In this study we amplified and sequenced the IL17A gene for four leporids species (European rabbit, European brown hare, brush rabbit and pygmy rabbit). For European rabbit (*O. c. cuniculus*) and European brown hare, both genomic and cDNA sequences were identical and only one of the sequences is presented; however both sequences have been assigned different accession numbers. These sequences were further compared to sequences of IL17A from another lagomorph, American pika (*Ochotona princeps*), and from representatives of the most relevant mammalian groups (e.g. Artiodactyla, Carnivores, Chiroptera, Primates, Rodents, etc.) available in online databases. In the European rabbit, IL17A is located in the forward strand of chromosome 12 and has a similar structure to other mammals with three coding exons. The IL17A cDNA sequence obtained in this work for *Lepus europaeus* showed a similar structure.

In humans, IL17A codes for a protein with 155 amino acids (aa) and has the ability to bind with high affinity to IL17RA and IL17RC (Ely et al., 2009; Liu et al., 2013; Sabat et al., 2013). The interaction between interleukins and their receptors is crucial for their function and signalling and any changes in the amino acid composition may induce alterations in the protein conformation. In humans and rodents these interactions sites are described (Ely et al., 2009; Sabat et al., 2013) and include Leu52, Ile54, Ser61, Ser70-Tyr72, Arg75, Arg84, Arg88-Val94, Trp96, Leu103, His114, His115, Asn117, Ser118, Gln122-Glu124, Leu128, Arg130, Phe139, Pro155-Met160 (Figure 3.4).

In leporids, the IL17A codes for a protein with 153 aa and we observed that the sites that likely interact with the receptors are quite conserved. Indeed, from the thirty three amino acids involved in the linkage between IL17A and IL17RA, eighteen are conserved: twelve are maintained between mammals and the other six, despite being different, do not alter the charge neither the polarity. For the remaining fifteen amino acids, only three are differently charged, seven have distinct polarity and five have both different charge and polarity (Table 3.6).

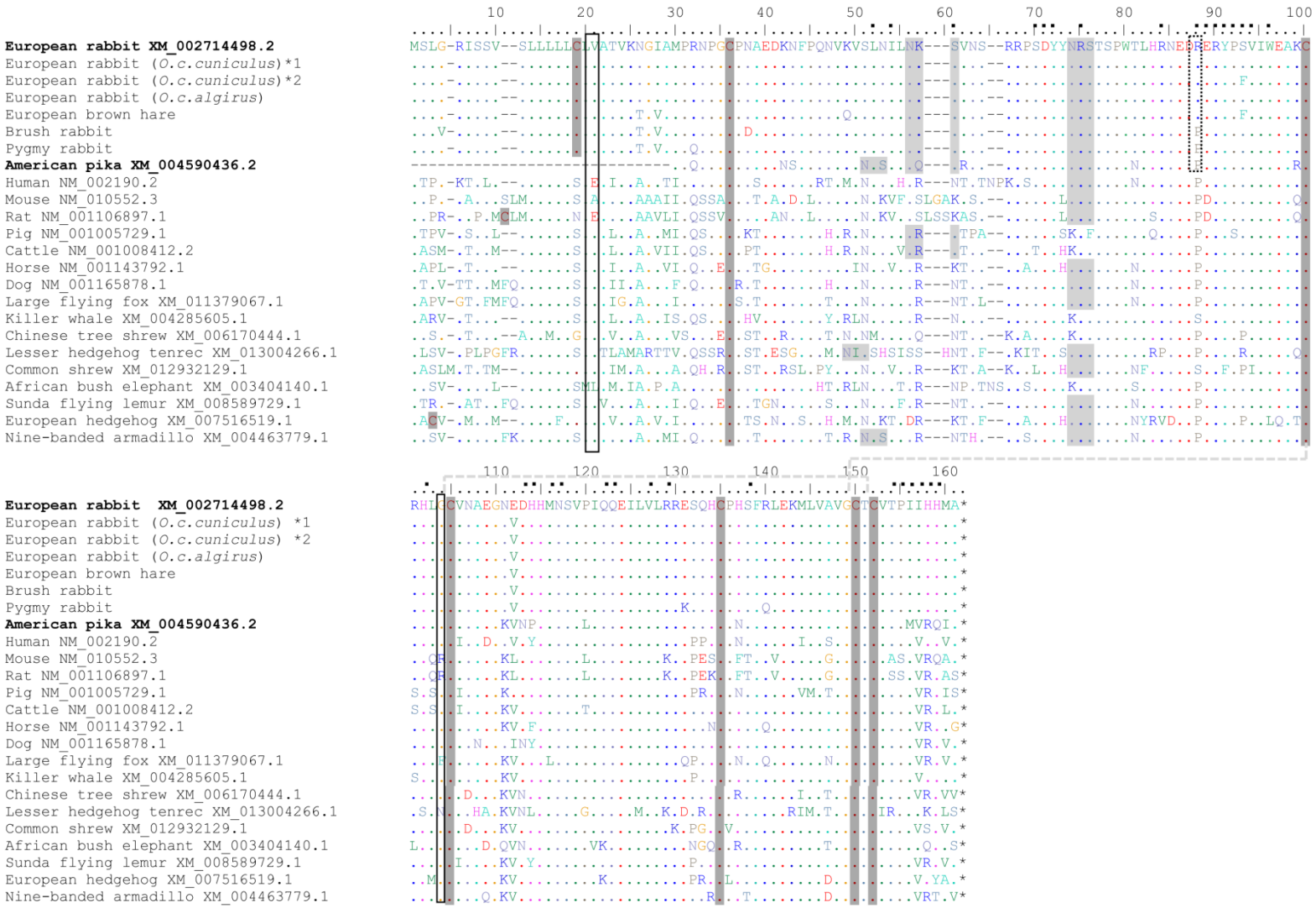


Figure 3. 4. Alignment of IL17A for several mammalian species. GenBank and Ensembl accession numbers are indicated in bold for the retrieved sequences. Positively selected amino acids are boxed (according to Neves et al., 2014) N-glycosylation sites are shaded in light grey and cysteine residues are shaded in dark grey. A black dashed box represent the Agr>Pro mutation between leporids. (*) represent stop codons; (-) represent indels; (▪) above the numbering represent the sites important for IL17A-IL17Ra interaction. *1 and *2 represent different alleles. Numbering is according to the European rabbit sequence (GenBank accession number XM_002714498.2) and the signal peptide and indels were included in the numbering. Disulfide bonds between side chain cysteines are represented by a light grey dashed line.

Between leporids these sites are highly conserved, but a mutation was observed that is located in the external coil of the IL17A in a site where this protein interacts with IL17RA (Figure 3.5).

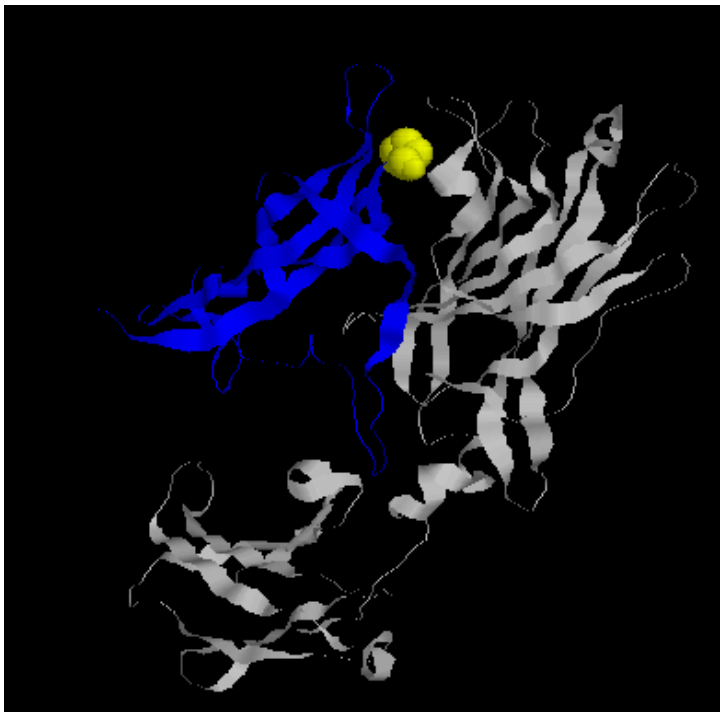


Figure 3. 5**.** 3D structures of the IL17A-IL17RA complex. IL17A appears in blue while IL17RA appears in grey. Marked in yellow is the 88Arg>Pro mutation described for leporids.

This mutation, 88Arg>Pro, occurs in the European rabbit and in the European brown hare, while in brush rabbit and the American pika the amino acid present is a Proline as in most mammals. Some studies showed that Arg>Pro mutations have crucial effects in the protein function (Fang et al., 2007; Ferrer-Costa et al., 2002; Yap et al., 2005). Indeed, the 332Arg>Pro mutation in human Trim5α restricts infection by HIV-1 (Human Immunodeficiency Virus-1) (Yap et al., 2005) while the 132Arg>Pro mutation in the helicase protein of coronavirus infectious bronchitis virus was lethal to infectivity *in vitro* (Fang et al., 2007). Additionally, this mutation alters the physiochemical properties of the amino acid by changing from a basic polar and positively charged Arginine to a nonpolar and neutral Proline.

Table 3. 6. Characterization of the IL17A amino acids differences in the sites important for binding to IL17RA.

Amino acid position	Amino acids				
	Leporids				
	European rabbit	European brown hare	Brush rabbit	Pygmy rabbit	Other mammals
52		L$^{\#}$			M$^{\#}$, S*
54		I$^{\#}$			V$^{\#}$, S*, T*
61		S*			N*, K^{*+}
70		S*			T*, L$^{\#}$
71			D^{*-}		
72			Y*		
75			R^{*+}		
84		R^{*+}			P$^{\#}$, V$^{\#}$
88	R^{*+}			P$^{\#}$	P$^{\#}$, S*
89		E^{*-}			D^{*-}
90			R^{*+}		
91		Y*			F$^{\#}$
92		P$^{\#}$			S*
93	S*, F$^{\#}$	F$^{\#}$		S*	F$^{\#}$, P$^{\#}$, R^{*+}
94			V$^{\#}$		
96		W$^{\#}$			L$^{\#}$
103		L$^{\#}$			Q*, S*, M$^{\#}$
114		H^{*+}			P$^{\#}$, Y*, F$^{\#}$, L$^{\#}$
115			H^{*+}		
117			N*		
118			S*		
122		Q*			K^{*+}
123			Q*		
124			E^{*-}		
128			L$^{\#}$		
130		R^{*+}			K^{*+}
139			F$^{\#}$		
155		P$^{\#}$			S*
156		I$^{\#}$			M$^{\#}$
157		I$^{\#}$			V$^{\#}$
158		H^{*+}			S*, R^{*+}, Q*, K^{*+}
159		H^{*+}			Q*, Y*, T*
160		M$^{\#}$			I$^{\#}$, V$^{\#}$, A, L$^{\#}$

The amino acid polarity (*hydrophilic; $^{\#}$hydrophobic) and charge ($^{+}$positive; $^{-}$negative) are properly annotated. The amino acid present in the human IL17A sequence is underlined. Numbering is according to the European rabbit IL17A sequence.

Disulfide bounds and N-glycosylation sites (Asn-X-Ser/Thr/Cys motifs where X can be any amino acid except proline) are important for the protein structure, stability and function (Helenius and Aebi, 2004; Lee et al., 2015; Rudd et al., 2001). Disulphide bounds occur between cysteines side chains and these linkages are also important for protein protection (Fass, 2012). In human and rodents, IL17A has a cysteine knot fold characterized by two sets of paired β-strands (1/2 and 3/4) interconnected by two disulfide bounds between strand 2

and 4 linked between four conserved cysteines (Cys100-Cys150 and Cys105-Cys152) (Hymowitz et al., 2001; Liu et al., 2013; Wright et al., 2008). In addition to these cysteines two other cysteines are common to all mammals, Cys36 and Cys135. For the European rabbit, the Psi-Pred predicted secondary structure and the DIANNA predicted disulfide bonds results are in agreement to those obtained and described for human and rodents (Hymowitz et al., 2001; Liu et al., 2013). An extra linkage is also predicted between Cys31 and Cys135. When compared to other mammals, there is an extra cysteine (Cys19) in leporids located in the signal peptide. The rat and the European hedgehog also have an extra cysteine located in different sites of the signal peptide (Cys11 and Cys3, respectively). Given that the signal peptide is cleaved in order to the protein to become active, this extra cysteine should not have an impact in the IL17A structure.

N-glycosylation is a crucial factor for the modulation of protein activity; therefore, alteration on these sites may interfere with recognition of targets, including receptors, and consequently affects the biological activity of the proteins and also their ability to diffuse through the organism (Chamorey et al., 2002; Shental-Bechor and Levy, 2008). Human IL17A is N-glycosylated at Asn68. Detection of putative N-glycosylation sites indicated that this N-glycosylation site is present in the majority of mammals, including rodents and lagomorphs. Other putative N-glycosylation sites were detected and include Asn56 in lagomorphs, pig and cattle, Asn51 in American pika and armadillo and Asn49 in the lesser hedgehog tenrec. The killer whale and the African bush elephant have no putative N-glycosylation sites. The implications of the absence/presence of N-glycosylation sites in IL17A are unknown, however some studies indicate that presence/removal of glycans in some proteins do not alter their folding or function, although a decrease in the protein dynamics is observed (Helenius and Aebi, 2004; Lee et al., 2015; Shental-Bechor and Levy, 2008).

Table 3. 7. IL17A nucleotide distances (the lowest values are in bold and the highest values are underlined).

	1	2	3	4	5	6	7	8	9
1 European_rabbit (*O.c.cuniculus*)	-								
2 European_rabbit (*O.c.algirus*)	0.002	-							
3 European brown hare	**0.011**	0.013	-						
4 Brush rabbit	**0.011**	0.013	0.013	-					
5 Pygmy rabbit	0.024	0.026	0.026	0.022	-				
6 American pika XM_004590436.2	0.112	0.115	0.120	0.112	0.112	-			
7 Human NM_002190.2	**0.169**	0.171	0.175	0.171	0.173	0.159	-		
8 Mouse NM_010552.3	0.251	0.251	0.249	0.249	0.245	0.240	0.236	-	
9 Rat NM_001106897.1	0.260	0.260	0.258	0.258	0.253	0.232	0.236	0.111	-

Comparison of the nucleotide sequences (Table 3.7) indicated that in leporids, the European rabbit and the European brown hare IL17A sequences are the least divergent (0.011) while the European rabbit and the pygmy rabbit IL17 sequences are the most divergent (0.026). Between the European rabbit and American pika, the genetic diversity obtained was 0.112-0.115. For the remaining mammals the highest divergence occurs for the lesser hedgehog tenrec (0.312) and the lowest for the flying lemur (0.145). The comparison of the nucleotide diversity of several interleukins in the European rabbit suggested that it could represent a better animal model for research (Neves et al., 2015). For IL17A, similar results were obtained, with the human sequence being more closely related to the European rabbit (0.169) than to mouse or rat IL17A sequences (0.236). This is further supported by a Maximum Likelihood tree inferred for IL17A mammalian sequences (Figure 3.6).

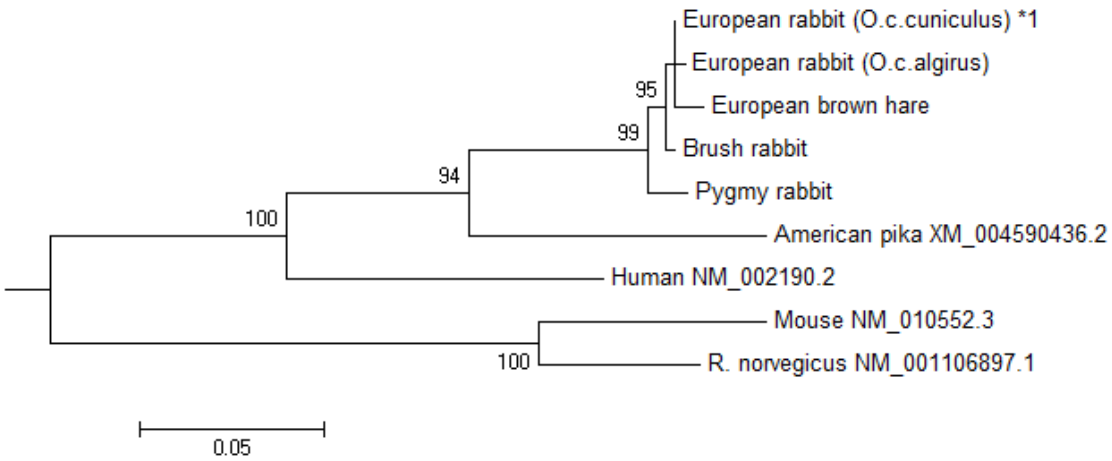


Figure 3. 6**.** Maximum Likelihood (ML) tree of the IL17A nucleotide sequences. Only bootstrap values ≥ 94% are shown. In order to facilitate visualization, only one sequence/allele of each species was used.

5. CONCLUSIONS

In the present study we sequenced and characterized IL17A for four leporids. Overall, the genomic organization, the location of the cysteine residues and the presence of N-glycosylation sites are highly conserved in leporids. Nevertheless, a single mutation was detected within the interaction site with IL17RA which may induce crucial changes in IL17A structure, function, stability, signaling and conformation. Further functional and structural studies should be performed to fully understand the impact of this specific mutation. The lowest divergence between the European rabbit and human IL17A sequences reinforces the hypothesis that the European rabbit might be a more suitable animal model for studies in the human innate immunity.

6. REFERENCES

Bettelli, E., Korn, T., Kuchroo, V.K., 2007. **Th17: the third member of the effector T cell trilogy**. Curr Opin Immunol 19, 652-657.

Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K., Jones, D.T., 2013. **Scalable web services for the PSIPRED Protein Analysis Workbench**. Nucleic Acids Res 41, W349-357.

Carvalho, C.L., Lopes de Carvalho, I., Ze-Ze, L., Nuncio, M.S., Duarte, E.L., 2014. **Tularaemia: a challenging zoonosis**. Comp Immunol Microbiol Infect Dis 37, 85-96.

Chamorey, A.L., Magne, N., Pivot, X., Milano, G., 2002. **Impact of glycosylation on the effect of cytokines. A special focus on oncology**. Eur Cytokine Netw 13, 154-160.

Edgar, R.C., 2004. **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. Nucleic Acids Res 32, 1792-1797.

Ely, L.K., Fischer, S., Garcia, K.C., 2009. **Structural basis of receptor sharing by interleukin 17 cytokines**. Nat Immunol 10, 1245-1251.

Fang, S., Chen, B., Tay, F.P., Ng, B.S., Liu, D.X., 2007. **An arginine-to-proline mutation in a domain with undefined functions within the helicase protein (Nsp13) is lethal to the coronavirus infectious bronchitis virus in cultured cells**. Virology 358, 136-147.

Fass, D., 2012. **Disulfide bonding in protein biophysics**. Annu Rev Biophys 41, 63-79.

Ferre, F., Clote, P., 2006. **DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification**. Nucleic Acids Res 34, W182-185.

Ferrer-Costa, C., Orozco, M., de la Cruz, X., 2002. **Characterization of disease-associated single amino acid polymorphisms in terms of sequence and structure properties**. J Mol Biol 315, 771-786.

Fischer, B., Chavatte-Palmer, P., Viebahn, C., Navarrete Santos, A., Duranthon, V., 2012. **Rabbit as a reproductive model for human health**. Reproduction 144, 1-10.

Gaffen, S.L., 2008. **An overview of IL-17 function and signaling**. Cytokine 43, 402-407.

Gerhardt, S., Abbott, W.M., Hargreaves, D., Pauptit, R.A., Davies, R.A., Needham, M.R., Langham, C., Barker, W., Aziz, A., Snow, M.J., Dawson, S., Welsh, F., Wilkinson, T., Vaugan, T., Beste, G., Bishop, S., Popovic, B., Rees, G., Sleeman, M., Tuske, S.J., Coales, S.J., Hamuro, Y., Russell, C., 2009. **Structure of IL-17A in complex with a potent, fully human neutralizing antibody**. J Mol Biol 394, 905-921.

Gupta, R., Jung, E., Brunak, S., 2004. **Prediction of N-glycosylation sites in human proteins**. in press.

Helenius, A., Aebi, M., 2004. **Roles of N-linked glycans in the endoplasmic reticulum**. Annual review of biochemistry 73, 1019-1049.

Hemdan, N.Y., Birkenmeier, G., Wichmann, G., Abu El-Saad, A.M., Krieger, T., Conrad, K., Sack, U., 2010. **Interleukin-17-producing T helper cells in autoimmunity**. Autoimmun Rev 9, 785-792.

Hestvik, G., Warns-Petit, E., Smith, L.A., Fox, N.J., Uhlhorn, H., Artois, M., Hannant, D., Hutchings, M.R., Mattsson, R., Yon, L., Gavier-Widen, D., 2015. **The status of tularemia in Europe in a one-health context: a review**. Epidemiology and infection 143, 2137-2160.

Horner, D.S., Lefkimmiatis, K., Reyes, A., Gissi, C., Saccone, C., Pesole, G., 2007. **Phylogenetic analyses of complete mitochondrial genome sequences suggest a basal divergence of the enigmatic rodent Anomalurus**. BMC Evol Biol 7, 16.

Huang, W., Na, L., Fidel, P.L., Schwarzenberger, P., 2004. **Requirement of interleukin-17A for systemic anti-Candida albicans host defense in mice**. The Journal of infectious diseases 190, 624-631.

Hymowitz, S.G., Filvaroff, E.H., Yin, J.P., Lee, J., Cai, L., Risser, P., Maruoka, M., Mao, W., Foster, J., Kelley, R.F., Pan, G., Gurney, A.L., de Vos, A.M., Starovasnik, M.A., 2001. **IL-17s adopt a cystine knot fold: structure and activity of a novel cytokine, IL-17F, and implications for receptor binding**. The EMBO journal 20, 5332-5341.

Ishigame, H., Kakuta, S., Nagai, T., Kadoki, M., Nambu, A., Komiyama, Y., Fujikado, N., Tanahashi, Y., Akitsu, A., Kotaki, H., Sudo, K., Nakae, S., Sasakawa, C., Iwakura, Y., 2009. **Differential roles of interleukin-17A and -17F in host defense against mucoepithelial bacterial infection and allergic responses**. Immunity 30, 108-119.

Iwakura, Y., Ishigame, H., Saijo, S., Nakae, S., 2011. **Functional specialization of interleukin-17 family members**. Immunity 34, 149-162.

Jones, D.T., 1999. **Protein secondary structure prediction based on position-specific scoring matrices**. J Mol Biol 292, 195-202.

Khader, S.A., Gopal, R., 2010. **IL-17 in protective immunity to intracellular pathogens**. Virulence 1, 423-427.

Kim, D.Y., Reilly, T.J., Schommer, S.K., Spagnoli, S.T., 2010. **Rabbit tularemia and hepatic coccidiosis in wild rabbit**. Emerg Infect Dis 16, 2016-2017.

Kubelkova, K., A., M., 2015. **Putting the Jigsaw Together - A Brief Insight Into the Tularemia** Open Life Sciences 10, 195-216.

Lee, H.S., Qi, Y., Im, W., 2015. **Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study**. Scientific reports 5, 8926.

Lepitzki, D.A., Woolf, A., Cooper, M., 1990. **Serological prevalence of tularemia in cottontail rabbits of southern Illinois**. Journal of wildlife diseases 26, 279-282.

Librado, P., Rozas, J., 2009. **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data**. Bioinformatics 25, 1451-1452.

Liu, S., Song, X., Chrunyk, B.A., Shanker, S., Hoth, L.R., Marr, E.S., Griffor, M.C., 2013. **Crystal structures of interleukin 17A and its complex with IL-17 receptor A**. Nature communications 4, 1888.

Mailles, A., Vaillant, V., 2014. **10 years of surveillance of human tularaemia in France**. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin 19, 20956.

Matthee, C.A., van Vuuren, B.J., Bell, D., Robinson, T.J., 2004. **A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene**. Syst Biol 53, 433-447.

Murphy, W.J., Eizirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A., O'Brien, S.J., 2001. **Molecular phylogenetics and the origins of placental mammals**. Nature 409, 614-618.

Neves, F., Abrantes, J., Almeida, T., de Matos, A.L., Costa, P.P., Esteves, P.J., 2015. **Genetic characterization of interleukins (IL-1alpha, IL-1beta, IL-2, IL-4, IL-8, IL-10, IL-12A, IL-12B, IL-15 and IL-18) with relevant biological roles in lagomorphs**. Innate immunity 21, 787-801.

Neves, F., Abrantes, J., Steinke, J.W., Esteves, P.J., 2014. **Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals**. Innate immunity 20, 184-191.

Ouyang, W., Kolls, J.K., Zheng, Y., 2008. **The biological functions of T helper 17 cell effector cytokines in inflammation**. Immunity 28, 454-467.

Paranavitana, C., Zelazowska, E., DaSilva, L., Pittman, P.R., Nikolich, M., 2010. **Th17 cytokines in recall responses against Francisella tularensis in humans**. J Interferon Cytokine Res 30, 471-476.

Pinheiro, A., Neves, F., Lemos de Matos, A., Abrantes, J., van der Loo, W., Mage, R., Esteves, P.J., 2016. **An overview of the lagomorph immune system and its genetic diversity**. Immunogenetics 68, 83-107.

Rijks, J.M., Kik, M., Koene, M.G., Engelsma, M.Y., van Tulden, P., Montizaan, M.G., Oomen, T., Spierenburg, M.A., Ijzer, J., van der Giessen, J.W., Grone, A., Roest, H.J., 2013. **Tularaemia in a brown hare (Lepus europaeus) in 2013: first case in the Netherlands in 60 years**. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin 18.

Rouvier, E., Luciani, M.F., Mattei, M.G., Denizot, F., Golstein, P., 1993. **CTLA-8, cloned from an activated T cell, bearing AU-rich messenger RNA instability sequences, and homologous to a herpesvirus saimiri gene**. J Immunol 150, 5445-5456.

Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A., Dwek, R.A., 2001. **Glycosylation and the immune system**. Science 291, 2370-2376.

Sabat, R., Witte, E., Witte, K., Wolk, K., 2013. **IL-22 and IL-17: An Overview**, in: Quesniaux, V., Ryffel, B., Padova, F.D. (Eds.), IL-17, IL-22 and Their Producing Cells: Role in Inflammation and Autoimmunity. Springer.

Schnupf, P., Sansonetti, P.J., 2012. **Quantitative RT-PCR profiling of the rabbit immune response: assessment of acute Shigella flexneri infection**. PLoS One 7, e36446.

Shen, F., Gaffen, S.L., 2008. **Structure-function relationships in the IL-17 receptor: implications for signal transduction and therapy**. Cytokine 41, 92-104.

Shental-Bechor, D., Levy, Y., 2008. **Effect of glycosylation on protein folding: a close look at thermodynamic stabilization**. Proc Natl Acad Sci U S A 105, 8256-8261.

Skyberg, J.A., Rollins, M.F., Samuel, J.W., Sutherland, M.D., Belisle, J.T., Pascual, D.W., 2013. **Interleukin-17 protects against the Francisella tularensis live vaccine strain but not against a virulent F. tularensis type A strain**. Infection and immunity 81, 3099-3105.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. **MEGA6: Molecular Evolutionary Genetics Analysis version 6.0**. Mol Biol Evol 30, 2725-2729.

Tesmer, L.A., Lundy, S.K., Sarkar, S., Fox, D.A., 2008. **Th17 cells in human disease**. Immunol Rev 223, 87-113.

Valentino, M.D., Maben, Z.J., Hensley, L.L., Woolard, M.D., Kawula, T.H., Frelinger, J.A., Frelinger, J.G., 2010. **Identification of T-cell epitopes in Francisella tularensis using an ordered protein array of serological targets**. Immunology 132, 348-360.

van de Veerdonk, F.L., Gresnigt, M.S., Kullberg, B.J., van der Meer, J.W., Joosten, L.A., Netea, M.G., 2009. **Th17 responses and host defense against microorganisms: an overview**. BMB reports 42, 776-787.

Vaure, C., Liu, Y., 2014. **A comparative review of toll-like receptor 4 expression and functionality in different animal species**. Front Immunol 5, 316.

Vuillaumier, S., Kaltenboeck, B., Lecointre, G., Lehn, P., Denamur, E., 1997. **Phylogenetic analysis of cystic fibrosis transmembrane conductance regulator gene in mammalian species argues for the development of a rabbit model for cystic fibrosis**. Mol Biol Evol 14, 372-380.

Wobeser, G., Campbell, G.D., Dallaire, A., McBurney, S., 2009. **Tularemia, plague, yersiniosis, and Tyzzer's disease in wild rodents and lagomorphs in Canada: a review**. Can Vet J 50, 1251-1256.

Wright, J.F., Bennett, F., Li, B., Brooks, J., Luxenberg, D.P., Whitters, M.J., Tomkinson, K.N., Fitz, L.J., Wolfman, N.M., Collins, M., Dunussi-Joannopoulos, K., Chatterjee-Kishore, M.,

Carreno, B.M., 2008. **The human IL-17F/IL-17A heterodimeric cytokine signals through the IL-17RA/IL-17RC receptor complex**. J Immunol 181, 2799-2805.

Yao, Z., Painter, S.L., Fanslow, W.C., Ulrich, D., Macduff, B.M., Spriggs, M.K., Armitage, R.J., 1995. **Human IL-17: a novel cytokine derived from T cells**. J Immunol 155, 5483-5486.

Yap, M.W., Nisole, S., Stoye, J.P., 2005. **A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction**. Current biology : CB 15, 73-78.

Zhu, S., Qian, Y., 2012. **IL-17/IL-17 receptor system in autoimmune disease: mechanisms and therapeutic potential**. Clin Sci (Lond) 122, 487-511.

GENETIC CHARACTERIZATION OF INTERLEUKINS (IL1A, IL1B, IL2, IL4, IL8, IL10, IL12A, IL12B, IL15 AND IL18) WITH RELEVANT BIOLOGICAL ROLES IN LAGOMORPHS

Fabiana Neves, Joana Abrantes, Tereza Almeida, Ana Lemos de Matos, Paulo P Costa, Pedro J Esteves

1. ABSTRACT

Interleukins (ILs), as essential innate immune modulators, are involved in an array of biological processes. In the European rabbit (*Oryctolagus cuniculus*) IL1α, IL1β, IL2, IL4, IL8, IL10, IL12A, IL12B, IL15 and IL18 have been implicated in inflammatory processes and in the immune response against rabbit hemorrhagic disease virus and myxoma virus infections. In this study we characterized these ILs in six Lagomorpha species (European rabbit, pygmy rabbit, two cottontail rabbit species, European brown hare and American pika). Overall, these interleukins are conserved between lagomorphs, including in their exon/intron structure. Most differences were observed between leporids and American pika. Indeed, when comparing both, some relevant differences were observed in American pika such as the location of the stop codon in IL1α and IL2, the existence of a different transcript in IL8 and the number of cysteine residues in IL1β. Changes at N-glycosylation motifs were also detected in IL1, IL10, IL12B and IL15. IL1α is the protein that presents the highest evolutionary distances, in contrast to IL12A where the distances between lagomorphs are the lowest. For all these ILs, sequences of human and European rabbit are more closely related than between human and mouse or European rabbit and mouse.

Keywords: Immune system, Interleukins, lagomorphs

2. INTRODUCTION

Interleukins (ILs) are polypeptides of low molecular weight involved in several biological activities, including immunity, inflammation, inflammatory diseases, hematopoiesis, oncogenesis, and fertility, among others (Afzal et al., 2012; Heinrich et al., 2003; Ishihara and Hirano, 2002). In vertebrates, many of these proteins participate in host defense with complementary and conflicting roles in induction, regulation and functioning of the immune system by regulating growth, differentiation, effector functions and survival of cells (Brocker et al., 2010; Kaiser et al., 2004). ILs are crucial for the immune response. They are produced as an integral part of the innate immune response, and have the ability to influence the result and nature of adaptive immune response (Brocker et al., 2010; Kaiser et al., 2004; O'Connell and McInerney, 2005; Zelus et al., 2000; Zhang and Nei, 2000). Despite these functions, ILs are also considered important therapeutic targets (Bessis and Boissier, 2001; He et al., 2014; Rossi et al., 2015), thereby any changes in their sequence or structure may lead to alteration in their normal functioning. In mammals, ILs encoding genes are among the ones with faster evolution (Brocker et al., 2010; O'Connell and McInerney, 2005), and 28 of the 46 known ILs present signatures of positive selection (Neves et al., 2014b).

Although being well characterized in most mammalian groups, little is known about ILs in lagomorphs, with the exception of the European rabbit (*Oryctolagus cuniculus*). The order Lagomorpha comprises two families, Ochotonidae and Leporidae that diverged ~35 million years ago (mya) (Matthee et al., 2004). While family Ochotonidae comprise only one genus (*Ochotona* or pikas), the family Leporidae comprises eleven genera of hares and rabbits widely distributed: pygmy rabbit (genus *Brachylagus*), riverine rabbit (genus *Bunolagus*), striped rabbit (genus *Nesolagus*), European rabbit (genus *Oryctolagus*), Amami rabbit (genus *Pentalagus*), Bunyoro rabbit (genus *Poelagus*), Red rock rabbit (genus *Pronolagus*), volcano rabbit (genus *Romerolagus*), cottontail rabbits (*Sylvilagus* spp.), hispid hare (genus *Caprolagus*) and true hares and jackrabbits (*Lepus* spp.) (Chapman and Flux, 2008; Ge et al., 2013; Matthee et al., 2004). The ones that are more closely related to the *Oryctolagus* genus are *Bunolagus, Caprolagus* and *Pentalagus*

with divergence times of ~7, ~8 and ~9 mya, respectively. In contrast, *Nesolagus*, *Poelagus* and *Pronolagus* are the less related to *Oryctolagus* with divergence times of ~15 mya (Matthee et al., 2004). The other two leporidae genera with some information for ILs, *Sylvilagus* and *Lepus*, diverged from *Oryctolagus* at ~12 mya (Esteves et al., 2005; Matthee et al., 2004; Pinheiro et al., 2013; van der Loo et al., 2009).

The European rabbit is one of the most used laboratory animal models for immunological research, including the study of atherosclerosis (Tian et al., 2012), intestinal immunity (Jimenez-Garcia et al., 2004), arthritis (Desando et al., 2013), cancer (Kang and Grossniklaus, 2011), Alzheimer's disease (Woodruff-Pak et al., 2007) and several viral infections (Marchandeau et al., 2014; Marques et al., 2012; Teixeira et al., 2012). In addition, the European rabbit is a key species in the Mediterranean ecosystem, where it is strongly affected by two viral diseases, the rabbit hemorrhagic disease (RHD) and myxomatosis (Abrantes et al., 2013; Dalton et al., 2014; Delibes-Mateos et al., 2008; Garcia-Bocanegra et al., 2010; Lopes et al., 2014; Muller et al., 2009). RHD is caused by a single-stranded RNA virus, the rabbit hemorrhagic disease virus (RHDV), while myxomatosis is caused by the myxoma virus (MYXV), a double-stranded DNA virus.

Rabbit resistance to both viral diseases is highly dependent on the immune response of the host in order to develop an effective adaptive immune response to control these viruses (Kerr and McFadden, 2002; Marques et al., 2012). ILs are important in the European rabbit immune response against RHDV and MYXV infections (Marchandeau et al., 2014; Marques et al., 2012), but also in inflammatory processes (Perkins et al., 2000). For RHDV, IL1, IL2, IL6, IL8 and IL10 are among the most important ILs. Indeed, young European rabbits infected with RHDV showed an increase of IL1 during an early stage of infection until 18 hours post-infection, with a decrease to normal values at 24 hours of infection while IL8 is particularly increased at 24 hours after infection, probably due to an increase of leucocyte migration to the site of infection (Marques et al., 2012). In contrast, in infected adult rabbits, IL10 is significantly increased in the course of the disease (Marques et al., 2012; Teixeira et al., 2012).

For MYXV, IL12, IL15 and IL18 have an important anti-viral activity. Indeed, recombinants of myxoma virus expressing human IL12, despite being similar to the wild type, do not induce myxomatosis in rabbits (Stanford et al., 2007). IL15 prevents lethal myxomatosis in the New Zealand rabbit breed through the stimulation of an immune response that leads to the elimination of the viral infection (Liu et al., 2009). Poxviruses developed the ability to block IL18 by interfering with proteins crucial for IL18 activation or function highlighting the importance of its anti-viral activity (Johnston and McFadden, 2003, 2004; Vande Walle and Lamkanfi, 2011). In contrast, expression of European rabbit IL4 by recombinant myxoma virus strains increases virus virulence and overcomes genetic resistance in wild rabbits (Kerr et al., 2004).

Considering the important biological role in the European rabbit immune response, we performed a genetic characterization of IL1, which includes two biologically similar antagonist proteins IL1α and IL1β , IL2, IL4, IL8, IL10, IL12A, IL12B, IL15 and IL18 in five Lagomorpha genera (*Oryctolagus*, *Brachylagus*, *Sylvilagus*, *Lepus* and *Ochotona*).

3. MATERIALS AND METHODS

Samples of European rabbit (*O. c. cuniculus and O. c. algirus),* pygmy rabbit (*Brachylagus idahoensis),* cottontail rabbits (brush rabbit - *Sylvilagus bachmani* and eastern cottontail - *Sylvilagus floridanus)*, European brown hare (*Lepus europaeus)* and American pika (*Ochotona princeps)* were provided by the CIBIO Lagomorpha tissue collection. Genomic DNA (gDNA) was extracted using the EasySpin Genomic DNA Minipreps Tissue Kit (Citomed) according to the manufacturer's instructions. Total RNA was extracted by using the RNeasy Mini Kit according to the manufacturer's instructions (Qiagen) from one specimen of: European rabbit, European brown hare, eastern cottontail and American pika. Complementary DNA (cDNA) was synthesized using oligo(dT) as primers and SuperScript III reverse transcriptase (Invitrogen). The European rabbit and American pika IL sequences were retrieved from public databases (accession numbers are given in bold in Figure 3.7). PCR amplification was performed with the Multiplex PCR Kit (Qiagen) by using several pairs of primers designed according to the retrieved sequences (supplementary material Table

3.10). Sequencing was performed on an ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems) and PCR products were sequenced in both directions. The sequences obtained were submitted to GenBank (accession numbers: KT216045-KT216070; KT273911-273919; KT279631-KT279692).

The program PHASE, built into the software DnaSP (Librado and Rozas, 2009), was used to reconstruct the haplotype phases of the obtained sequences that were aligned using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) available at http://www.ebi.ac.uk/ (Edgar, 2004).

Putative *N*-glycosylation sites were predicted using the NetNGlyc 1.0 server available at http://www.cbs.dtu.dk/services/NetNGlyc/ (Blom et al., 2004).

Predicted splicing sites were determined by using the NetGene2 server available at http://www.cbs.dtu.dk/services/NetGene2/ (Brunak et al., 1991; Hebsgaard et al., 1996).

The number of amino acid differences per site between sequences was estimated in MEGA6 (Tamura et al., 2013) with the options: bootstrap method (1000 replicates), p-distance as model and pairwise deletion for gaps/missing data treatment. A Maximum Likelihood (ML) approach was used to estimate, for each gene, the phylogenetic relationships between the amino acid sequences. The ML trees were estimated in MEGA6 by using the best-fit nucleotide substitution model predicted by the same software and bootstrap 1000 replicates. In order to simplify the outputs and avoid duplicated sequences, only one sequence from each species was used.

The secondary structure of each IL was predicted using PsiPred (http://bioinf.cs.ucl.ac.uk/psipred/) (Buchan et al., 2013; Jones, 1999) and DiAminoacid Neural Network Application (DiANNA) (http://clavius.bc.edu/~clotelab/DiANNA/) (Ferre and Clote, 2006). These methods predict protein cysteines that create permanent structural disulfide bonds. PsiPred uses Position Specific Iterated – BLAST (PSI-BLAST) searches of the non-redundant protein sequence database to obtain evolutionary information used to predict the secondary structure of the query protein. DiANNA is a neural network trained to recognize cysteines in an oxidized state (sulfur covalently bonded) telling them apart from those in a reduced state.

4. RESULTS

The interleukins IL1α, IL1β, IL2, IL4, IL8, IL10, IL12A, IL12B, IL15 and IL18 were amplified from gDNA for the two European rabbit subspecies, European brown hare, pygmy rabbit, brush rabbit and American pika. These ILs were also successfully amplified from cDNA of European rabbit (*O. c. cuniculus*), European brown hare, eastern cottontail and American pika. The allelic forms identified in DnaSP are shown in Figure 3.7. The results obtained are summarized in Table 3.8.

Table 3. 8. Summary of the alterations observed between lagomorph species for the ILs studied.

	Insertion/Deletion	**Cysteine residues**	**N-Glycosylation (N-X-T)**
IL1α	American pika: deletion aa 52; insertion aa 179, aa 272-275	-	Cottontails: N-X-T missing (Asn64) America pika: two different N-X-T (Asn55 and Asn92)
IL1β	American pika: deletion aa 39-43; deletion aa 102 and 108. Pygmy rabbit and cottontail rabbits: deletion aa254	American pika: extra Cys aa 135.	European rabbit and European brown hare: extra N-X-T (Asn254)
IL2	Pygmy rabbit: insertion aa 7 and 8. American pika: insertion aa 139-141 and aa 174; deletion of aa 146.	-	-
IL4	-	-	-
IL8	American pika: insertion aa 50-53;	-	-
IL10	-	-	European brown hare: extra N-X-T (Asn67)
IL12A	-	Brush rabbit: extra cys aa 54	-
IL12B	European rabbit, Brush rabbit and European brown hare: insertion aa 174-177; Brush rabbit, European brown hare and American pika: insertion aa 189.	-	Brush rabbit: extra N-X-T (Asn147 and Asn 280);
IL15	-	American pika: extra Cys aa 140.	Pygmy rabbit and European brown hare: N-X-T missing (Asn152). American pika: N-X-T missing (Asn168).
IL18	European rabbit: deletion aa 14;	-	-

IL1

IL1 is composed by two proteins IL1α and IL1β. According to the public databases NCBI, Ensembl and Uniprot, in the European rabbit these proteins are located in tandem in the reverse strand of chromosome 2, while in the American pika they are located in the forward strand.

IL1α

In lagomorphs, IL1α is organized into 7 exons, with 6 of them being coding exons. The coding sequence comprises 801 base pairs (bp) that translate into a protein with 267 amino acids (aa). However, for American pika, and like the human IL1α, the coding sequence has 813 bp translating into a protein with 271 aa (Figure 3.7.a). Some differences existed between lagomorphs: American pika presents an amino acid deletion at position 52 of the European rabbit and an insertion of five amino acids at positions 179 and from 272 to 275. All cysteine residues are conserved between lagomorphs (Cys14 and Cys47). Three potential N-glycosylation sites were detected for all leporids (Asn 64, Asn103 and Asn144), but Asn64 is absent in the cottontail rabbits studied. For American pika, there are two other potential N-glycosylation sites (Asn55 and Asn92).

L1β

In lagomorphs, IL1β comprises 6 coding exons, translating into proteins with different amino acid lengths: 268 aa in the European rabbit and 261 aa in American pika (Figure 3.7.b). In the American pika sequence we observed three deletions: five aa between positions 39 and 43 and which include the deletion of two cysteine residues that are conserved in all leporids; the two other deletions are located at amino acids 102 and 108. Between leporids there are nine conserved cysteines at sites 35, 42, 43, 54, 81, 105, 117, 126 and 189. In the American pika sequence, only four cysteines are conserved (Cys81, Cys117, Cys126 and Cys189), and an extra cysteine is found at position 135. All lagomorphs have a putative N-glycosylation site at Asn58; European rabbit and European brown hare have an additional putative N-glycosylation site at Asn254 that had been lost in both the brush rabbit and in the eastern cottontail rabbit due to a deletion of this amino acid.

Figure 3.7.a. IL1α

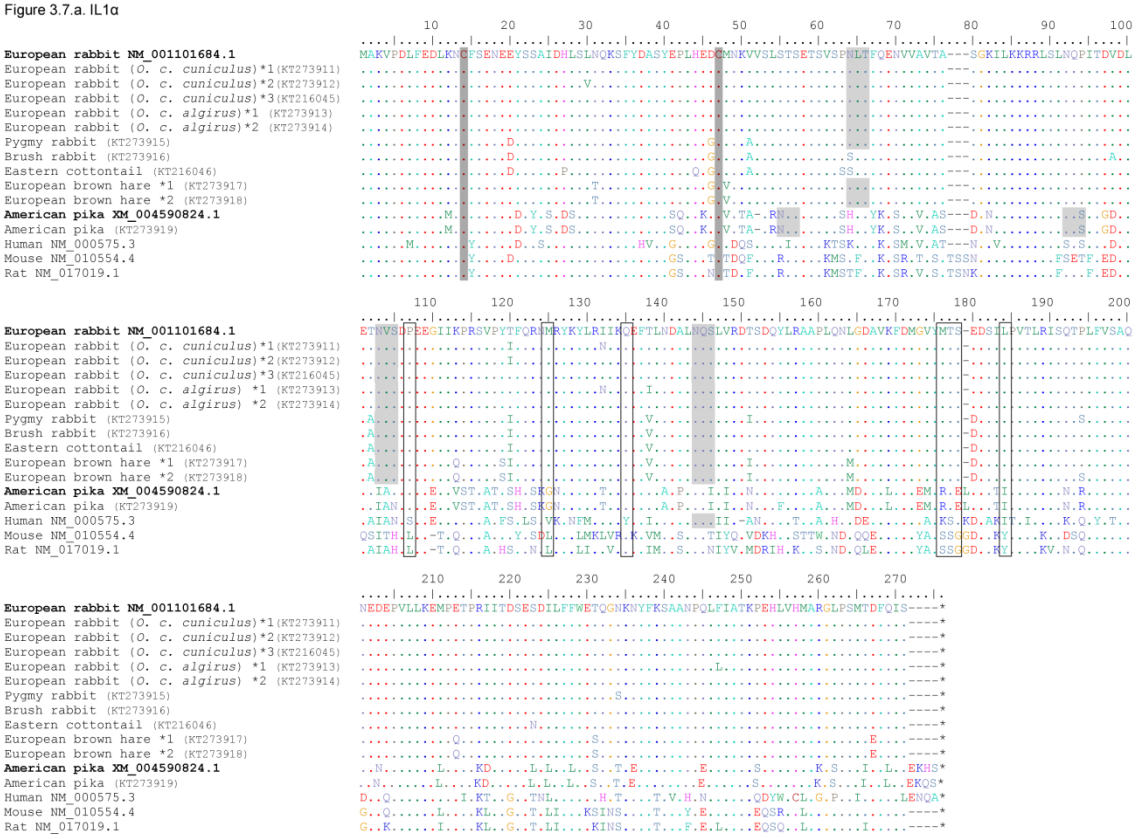


Figure 3.7.b. IL1β

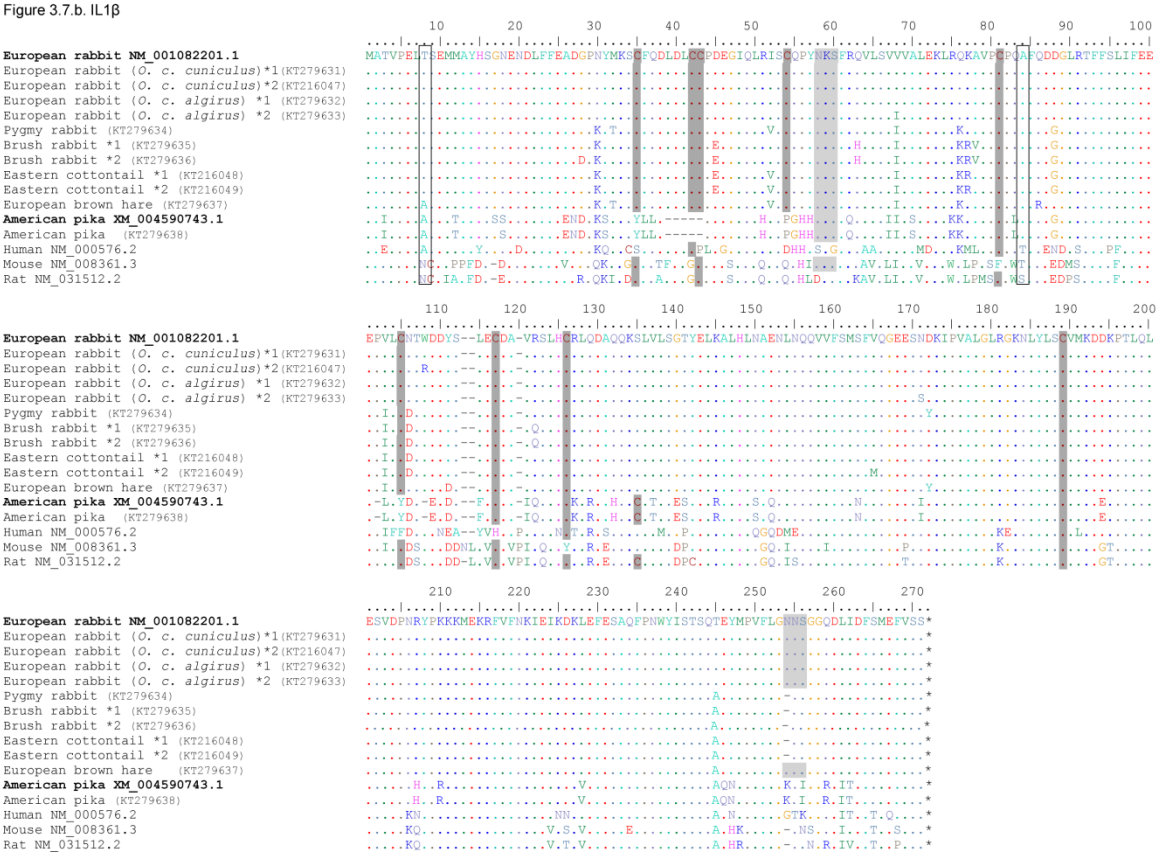


Figure 3.7.c. IL2

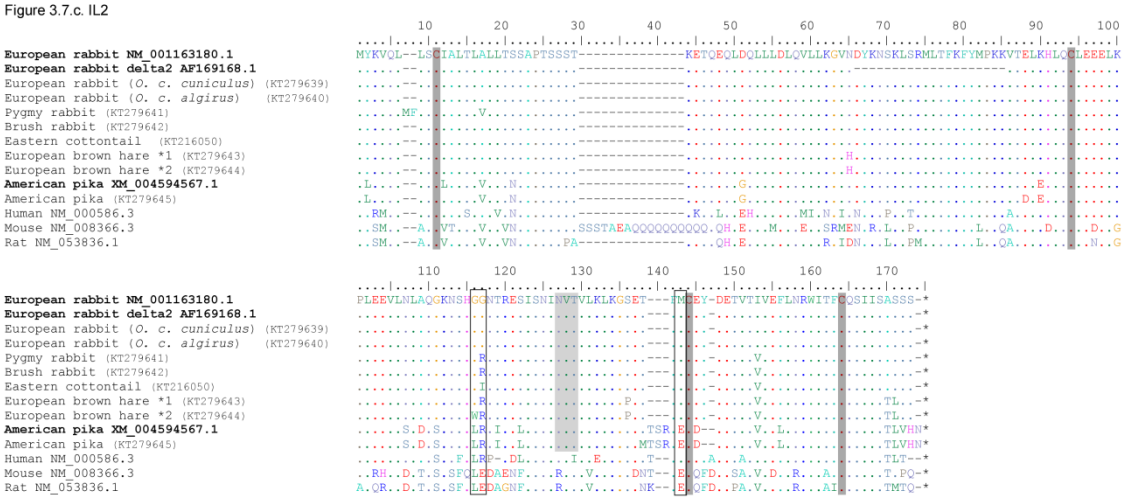


Figure 3.7.d. IL4

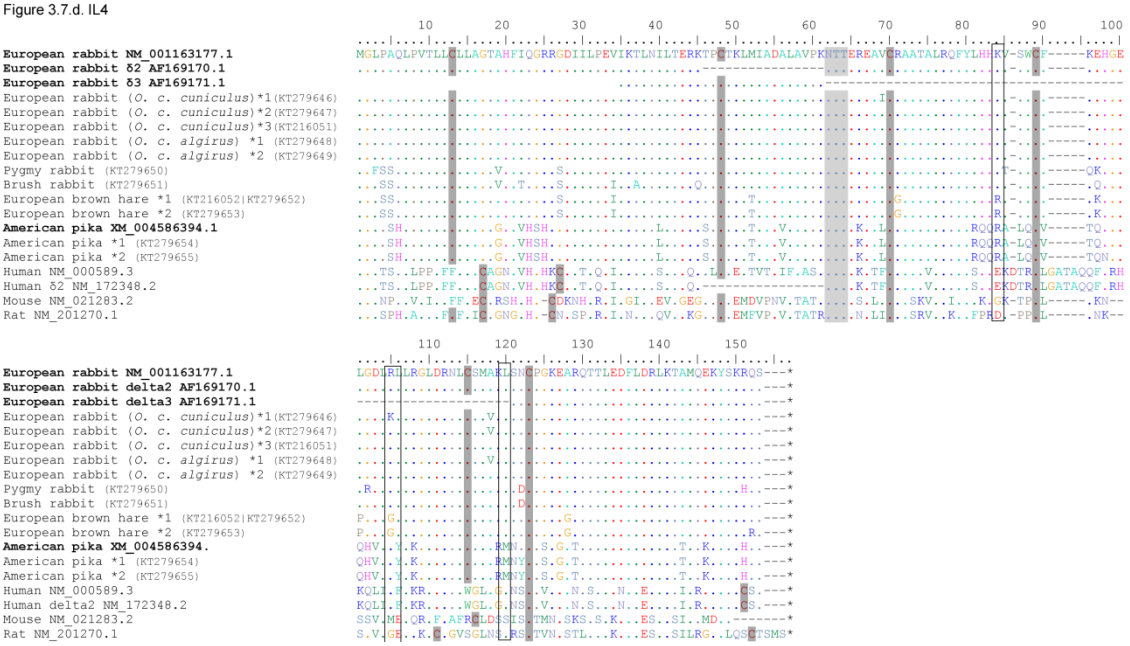


Figure 3.7.e. IL8

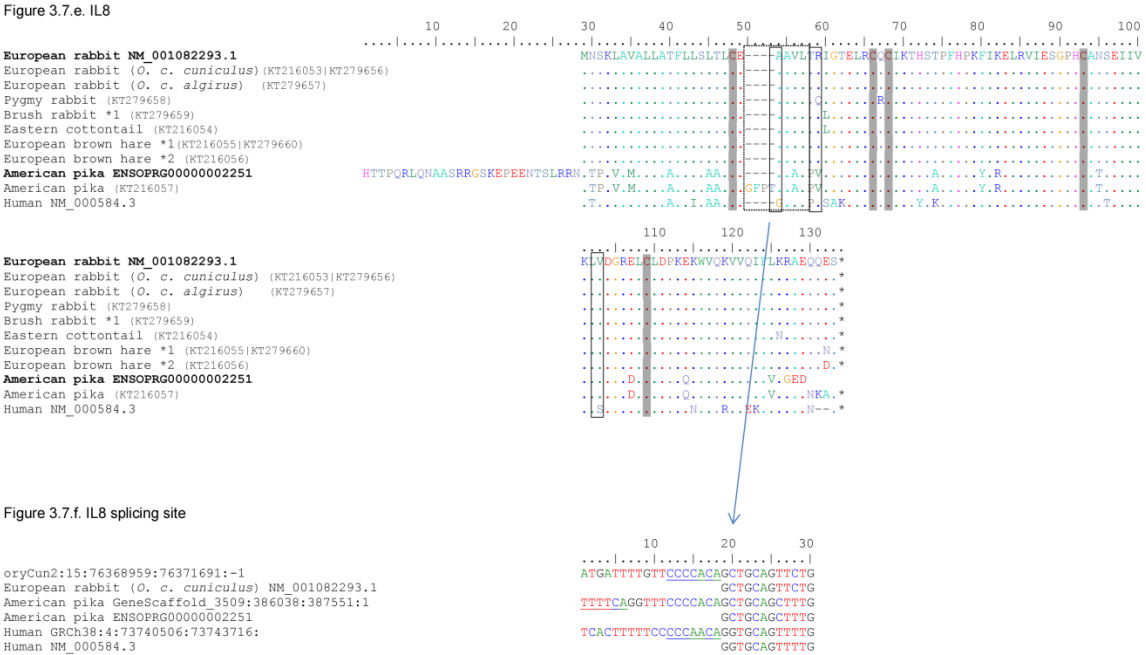


Figure 3.7.f. IL8 splicing site

Figure 3.7.g. IL10

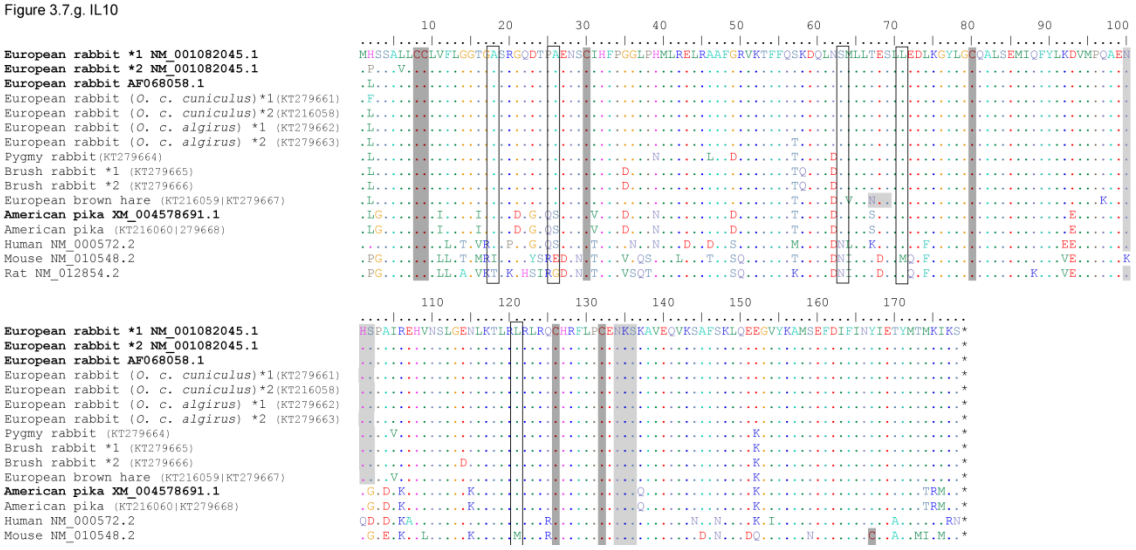


Figure 3.7.h. IL12A

Figure 3.7.i. IL12B

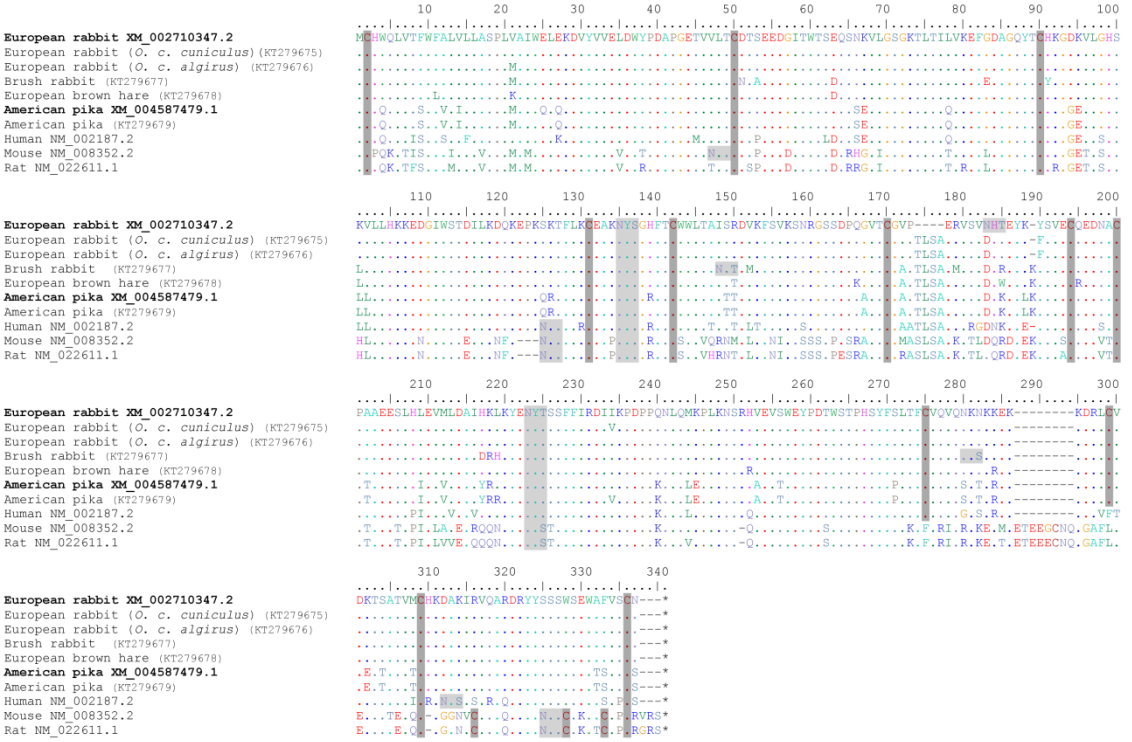
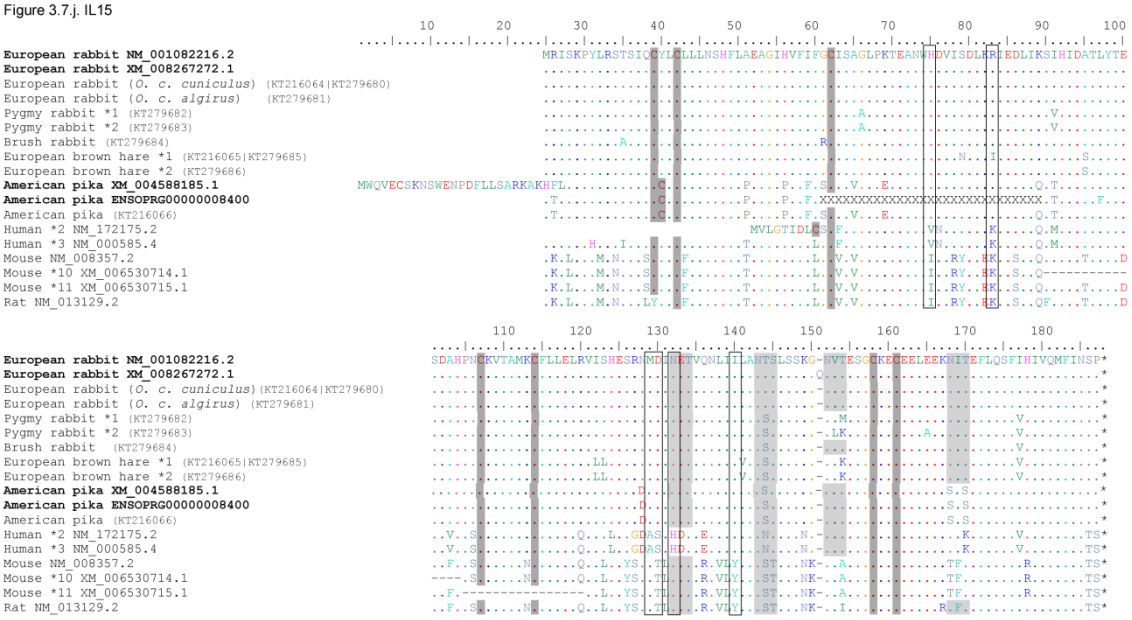


Figure 3.7.j. IL15



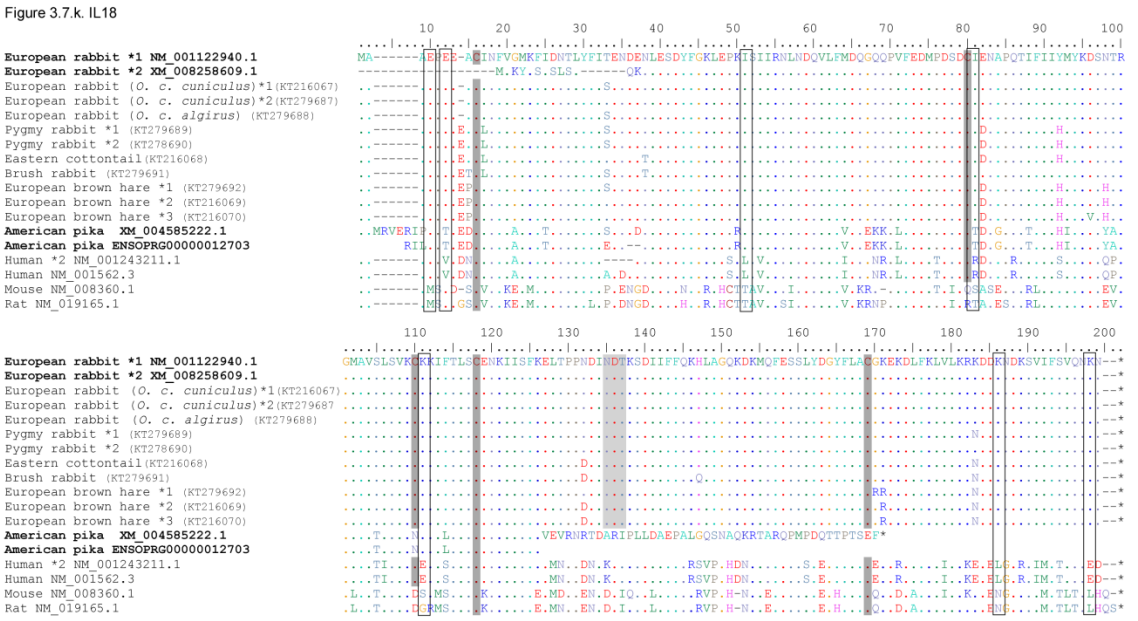


Figure 3. 7. Alignment of the studied interleukins for the different lagomorphs a) IL1α; b) IL1β; c) IL2; d) IL4; e) IL8; f) detail of the splicing region of exon 2 of IL8, with the splicing regions underlined; g) IL10; h) IL12A; i) IL12B; j) IL15; k) IL18. GenBank and Ensembl accession numbers are indicated in bold for the retrieved sequences. Positively selected amino acids are boxed according to Neves et al., 2014. N-glycosylation sites are shaded in light grey and cysteine residues are shaded in dark grey. (*) represent stop codons; (-) represent deletions; *1 and *2 represent alleles. The numbering is according to the European rabbit sequences. The signal peptide and indels, indicated as (-), were included in the numbering.

IL2

Regarding IL2 sequence, this protein is highly conserved between European rabbit, human, cat and horse. In the European rabbit, IL2 is located in the reverse strand of chromosome 15, with 4 coding exons that translate into a protein with 153 aa. In American pika IL2 has also 4 exons that translate into a 156 aa protein (Figure 3.7.c). Differences between lagomorphs include the insertion of two aa in pygmy rabbit (Met7 and Phe8) and four aa in American pika (Thr139, Ser140, Arg141 and Asn174). In American pika we also observed a deletion of one aa that corresponds to position 146 of European rabbit. All cysteine residues are conserved between lagomorphs (Cys11, Cys94, Cys144 and Cys164). The search of putative N-glycosylation sites identified one site (Asn127) common to all lagomorphs.

IL4

In the European rabbit, IL4 is located in chromosome 3 (forward strand). In the European rabbit and American pika, IL4 is organized into 4 coding exons

that translate into a protein of 147 aa (Figure 3.7.d). No alterations, including insertions/deletions, were observed between IL4 of lagomorphs. Also, cysteine residues were conserved (Cys13, Cys48, Cys70, Cys89, Cys115 and Cys123) and the same potential N-glycosylation site was detected (Asn62).

IL8

IL8 from European rabbit is located in the reverse strand of chromosome 15 and it is encoded by 4 exons that translate into a 101 aa protein. For American pika, the coding sequence available in Ensembl.org (ENSOPRT00000002239) is only known by projection and it is incomplete by not starting with an initiation codon and by missing the stop codon. Nevertheless, we successfully amplified IL8 for American pika from both gDNA and cDNA (Figure 3.7.e and 3.7.f). Our results showed that American pika has an insertion of four aa (Gly50, Phe51, Pro52 and Thr53). The cysteine residues are conserved among lagomorph species (Cys48, Cys61, Cys63, Cys93 and Cys109) and there are no putative N-glycosylation sites.

IL10

IL10 from European rabbit is located in the forward strand of chromosome 2 and, in both the European rabbit and American pika, it is composed by 5 coding exons that translate into a protein with 178 aa (Figure 3.7.g). The cysteine residues are conserved between species (Cys8, Cys9 Cys30, Cys80, Cys126 and Cys132). The search for N-glycosylation sites revealed two potential sites in leporids (Asn100 and Asn134). Ans67 was also predicted for European brown hare and American pika has only one of the sites predicted for N-glycosylation (Asn134).

IL12

IL12 is a heterodimeric protein composed by two proteins, IL12A and IL12B. For European rabbit, and according to the public databases NCBI and Ensembl, the encoding genes are located in the forward strand of chromosome 14 and in the reverse strand of chromosome 3, respectively.

IL12A

IL12A from European rabbit is composed by 7 coding exons that translate into a protein with 219 aa. For American pika, there are 2 different coding sequences available in NCBI (XM_012930663.1) and Ensembl (ENSOPRT00000006274). When compared with the European rabbit sequence, the American pika sequence available in NCBI has a start codon 109 amino acids upstream, while the sequence available in Ensembl is incomplete, since it lacks an initiation codon and amino acid information in the middle of the sequence. We were not able to amplify IL12A from gDNA or cDNA of American pika (Figure 3.7.h). All cysteine residues are conserved between leporids (Cys2, Cys71, Cys98, Cys119, Cys130, Cys144, Cys157 and Cys230). In addition, brush rabbit presents an extra cysteine at position 54. Four potential N-glycosylation sites were detected for all leporids (Asn73, Asn127, Asn141 and Asn214).

IL12B

IL12B from European rabbit and American pika is composed by 6 coding exons that translate into a protein with 324 and 329 aa, respectively. Despite all the attempts, we were unable to amplify IL12B from cDNA of any of the lagomorphs studied. Successful amplification from gDNA for all leporids showed that in the middle of exon 4 there is an insertion of four aa (Thr174 to Ala177) that is absent in the predicted European rabbit sequence available in NCBI (XM_002710347.2). However, this insertion is present in human, mouse and rat cDNA (Figure 3.7.i). All cysteine residues (Cys2, Cys50, Cys90, Cys131, Cys142, Cys170, Cys194 and Cys200, Cys275, Cys291, Cys309 and Cys336) and the two potential N-glycosylation sites (Asn135, Asn223) are conserved between lagomorphs. In addition, brush rabbit evidenced two other potential N-glycosylation sites (Asn147 and Asn280).

IL15

In the European rabbit, IL15 is located in chromosome 15 (reverse strand) and it is organized into 6 coding exons that translate into a protein with 162 aa. There are two sequences described for the European rabbit in NCBI, with the accession numbers NM_001082216.2 and XM_008267272.1, being the

latter a variant with an insertion of one amino acid (Gln127) (Figure 3.7.j). For American pika the information available in NCBI and Ensembl is different for the 5' coding region. Amplification was only successful using the primers constructed according to the American pika IL15 sequence available in Ensembl (ENSOPRT00000008393). All cysteine residues are conserved in lagomorphs (Cys39, Cys42, Cys62, Cys107, Cys114, Cys158 and Cys161), although American pika has an extra cysteine (Cys40). Four potential N-glycosylation sites were detected in lagomorphs, Asn132, Asn143, Asn152 and Asn168, but Asn152 is absent in pygmy rabbit and in European brown hare, and Asn168 is absent in American pika.

IL18

The European rabbit's IL18 is located in the forward strand of chromosome 1 and is organized into 5 coding exons that translate into a protein with 192 aa. For American pika there are two different predicted sequences available, one from NCBI (XM_004585222.1) and the other from Ensembl (ENSOPRG00000012703). When compared with the other lagomorphs, the American pika sequence from NCBI had an insertion of six amino acids in the first exon, while exon 5 presented low amino acid identity and a smaller size. The sequence from Ensembl is incomplete since it lacks an initiation codon and exon 5 (Figure 3.7.k). For this reason, and despite all attempts, we were unable to amplify IL18 from gDNA or cDNA of American pika. The European rabbit has a deletion at amino acid 14 whereas all the other leporids have a glutamine residue. All cysteine residues are conserved between leporids (Cys16, Cys80, Cys110, Cys118 and Cys169). Our search revealed that Asn135 is the only putative N-glycosylation site.

Regarding the number of amino acid differences, brush rabbit and pygmy rabbit are among the species more related for IL1, 2, 4 and 18, while cottontail rabbits and European rabbit evidenced a closer relationship for IL4, 8, 10, 12 and 15. On the other hand, American pika was the most distantly related when compared to other lagomorphs (supplementary material Table 3.11). IL1α presented the higher amino acid distances between lagomorphs (0.038 – 0.308) while IL12A showed the lowest (0.018 – 0.096). When comparing the amino acids distances between human, mouse and European rabbit ILs sequences

(Table 3.9), we observed that for all the ILs studied the amino acid distances are significantly lower between human-European rabbit than between human-mouse or European rabbit-mouse. These results were further confirmed with the phylogenetic analysis (Figure 3.8). Overall, we also observed that for the studied interleukins, lagomorphs cluster according to the accepted evolutionary topology. However, for IL1, IL10, IL12A, IL15 and IL18, these proteins present low bootstrap values to support these clusters.

Table 3. 9. Amino acid distances between rabbit, mouse and human for the different interleukins studied

		Rabbit	Mouse
IL1α	Human	**0.350**	0.383
	Rabbit	-	0.383
IL1β	Human	**0.254**	0.315
	Rabbit	-	0.271
IL2	Human	**0.199**	0.373
	Rabbit	-	0.399
IL4	Human	**0.463**	0.564
	Rabbit	-	0.593
IL8	Human	0.192	-
	Rabbit	-	-
IL10	Human	**0.202**	0.270
	Rabbit	-	0.281
IL12A	Human	**0.183**	0.400
	Rabbit	-	0.391
IL12B	Human	**0.177**	0.310
	Rabbit	-	0.341
IL15	Human	**0.167**	0.272
	Rabbit	-	0.259
IL18	Human	**0.266**	0.358
	Rabbit	-	0.389

(the lowest values for each interleukin are in bold).

5. DISCUSSION

Interleukins are important proteins for the innate and adaptive immune responses. In the European rabbit IL1α, IL1β, IL2, IL4, IL8, IL10, IL12A, IL12B, IL15 and IL18 have been implicated in the immune response against two highly fatal viral diseases, rabbit hemorrhagic disease and myxomatosis, and in inflammatory processes. Recently, we described a mutation in the European rabbit and Amami rabbit (*Pentalagus furnessi*) IL6 stop codon that leads to a larger protein with a considerable increase of the number of cysteines (Neves et al., 2014a). For the ILs included in this study, no significant alterations were

observed in any of the stop codons. Interestingly, American pika IL1α and IL2 sequences presented an insertion of four and one aa, respectively, in the site where the others lagomorphs have a stop codon.

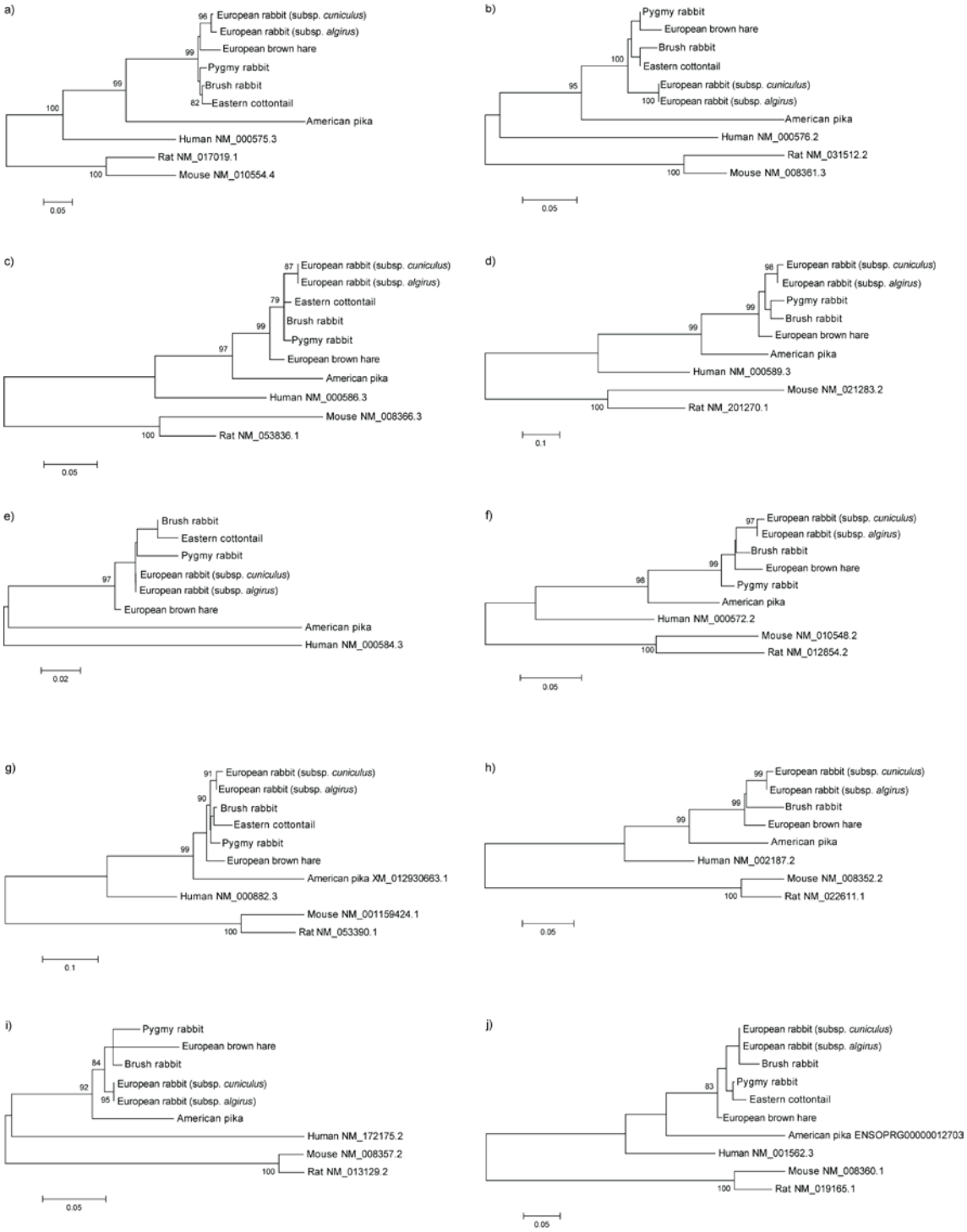


Figure 3. 8. Maximum Likelihood (ML) trees of the interleukins studied a) IL1α; b) IL1β; c) IL2; d) IL4; e) IL8; f) IL10; g) IL12A; h) IL12B; i) IL15; j) IL18. Only bootstrap values ≥ 0.75 are shown. In order to facilitate visualization, only one sequence/allele of each species was used. GenBank e Ensembl accession numbers are indicated for human, rodents and, in some cases, American pika sequences.

Some differences were observed between the European rabbit and the American pika sequences reported here and those already available in public databases (Supplementary material table 3.12). The major differences are in American pika IL8 and in the European rabbit IL12B, where both sequences have an insertion of 4 aa. However, the sequences available in public databases are predicted which might explain the differences observed. The remaining punctual differences are due to single polymorphisms arising from species diversity.

Alternative splicing is an important step for the production of mature mRNA that leads to protein diversity and may occur by: exon skipping (38%); alternative 5' or 3' spliced sites (26%); intron retention (3%); mutually exclusive exons, alternative promoters or multiple polyadenylation sites (33%) (Keren et al., 2010; Sahoo and Im, 2010). Exon skipping occurs by cleavage of specific exon-intron motifs located in the 5' and 3' termini regions of the intron. These motifs include, by decreasing order of relevance, GT-AG, CG-AG, AC-AT or AT-AC in the 5' and 3' regions, respectively (Holste and Ohler, 2008; Shimada et al., 2010; Wu and Krainer, 1999). Studies in the European rabbit reported alternative splicing in several ILs that leads to novel and functional proteins. Indeed, in IL2 and IL7 alternative splicing occurs with exclusion of exon 2 (Perkins et al., 2000; Siewe et al., 2010); for IL4 the presence of two variants (δ2, δ3) results from splicing out of exons 2 and 3, respectively (Perkins et al., 2000); for IL10 a spliced variant was described with the spliced region occurring between exon 5 and the 3' UTR region (Perkins et al., 2000). Beside these variants, abnormal transcripts were also characterized for European rabbit IL4 and IL10 (IL4 int2A, IL4 int3A, IL4 int3B and IL10/C) (Perkins et al., 2000). Although, more recently Mage and Mage suggested that there is no frameshit in exon 2 of IL4 (Mage and Mage, 2012). When comparing American pika IL8 sequence with the corresponding European rabbit sequence, we observed an insertion of four aa in the 5' region of exon 2 which is derived from the 3' region of the preceding intron. This suggests that in American pika the splicing occurred in a CA motif located 12bp upstream of the motif found in the European rabbit (Figure 3.7.f). These results were further confirmed by analyzing the human and American pika sequences using NetGene2 (Brunak et

al., 1991; Hebsgaard et al., 1996). Indeed, for human IL8 the splicing is predicted to occur by NetGene2 at nucleotide position 18 (Figure 3.7.f) (CCCCCAACA|GGTGCAGTTTT; with a 0.77 confidence value), while for American pika the confidence value is of 0.17, with the most likely splicing motif (confidence value of 0.65) being located at nucleotide position 6 (TAATTTTCA|GGTTTCCCCA). Thus, IL8 from American pika has a different transcript.

Disulfide bonds and glycosylation are important for protein protection against denaturation and proteolytic degradation (Bulaj, 2005). Disulfide bonds are formed between the thiol groups of cysteine residues and ensure important roles in folding, stability, function and in the regulation of protein activity, being well conserved between species (Bulaj, 2005; Fass, 2012; Li et al., 2011; Wong et al., 2011). In the studied ILs, and with the exception of IL1β, the cysteine residues are well conserved in lagomorphs. For American pika IL1β, and when comparing with other lagomorphs, the number and location of five out of nine cysteines residues is different. The analysis of protein secondary structure using PsiPred (Buchan et al., 2013; Jones, 1999) and DiANNA (Ferre and Clote, 2006) indicated that the loss of cysteine residues reduces the predicted bonds from four in the European rabbit (Cys35-Cys115; Cys42-Cys186; Cys43-Cys123 and Cys54-Cys81) to two in American pika (Cys108-Cys125 and Cys116-Cys179). These differences lead to the formation of an extra helix (Gly84 to Phe89) in American pika (Figure 3.9).

The glycosylation process is also important for several functions that include protein folding and interaction with cell surface receptors (Chamorey et al., 2002; Schwarz and Aebi, 2011). Thus, glycosylation can lead to protein diversity and modulation of protein properties (Chamorey et al., 2002; Helenius and Aebi, 2004; Rudd et al., 2001; Shental-Bechor and Levy, 2008). Putative N-glycosylation sites prediction showed some differences between lagomorphs, in particular for IL1α, IL1β, IL10, IL12B and IL15, however the consequences of these alterations should be considered in future studies.

Animal models are important tools for the study of human diseases. Historically, the European rabbit was the first animal model for immunological studies. Indeed, some of the foundations of molecular immunology were laid with the use of the European rabbit reviewed in (Pinheiro et al., 2011).

Nevertheless, over the last decade the European rabbit as a research animal model has been replaced by the mouse model due to its smaller size, lower cost, ease of breeding, etc (Webb, 2014). However, the choice of mouse as the best model to study human diseases has been controversial due to the high variability of results observed between mouse and human especially when the focus are inflammatory diseases (Burkhardt and Zlotnik, 2013; Mullane and Williams, 2014; Seok et al., 2013; Shay et al., 2015; Takao and Miyakawa, 2015; Webb, 2014).

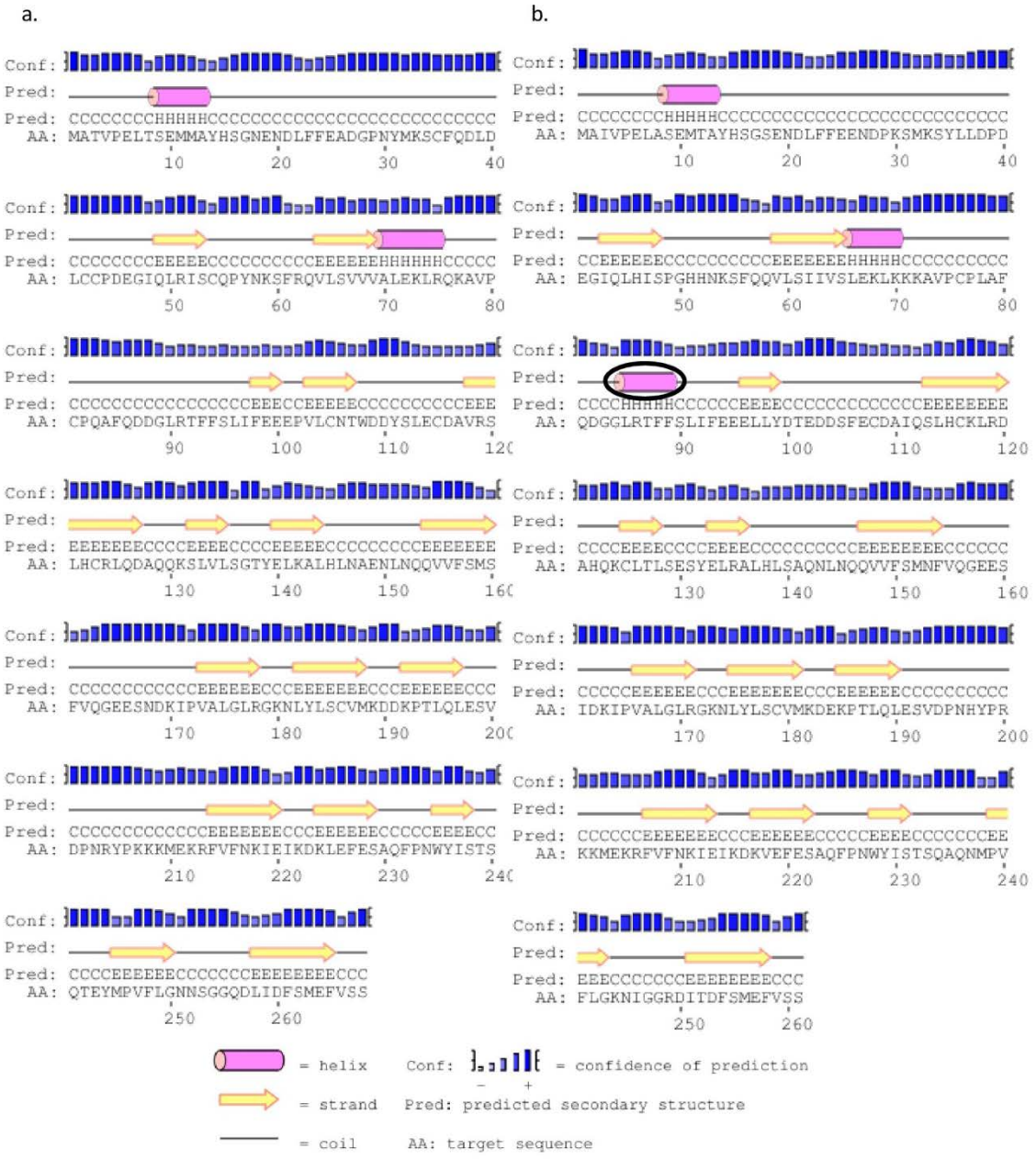


Figure 3. 9. IL1β PsiPred sequence analysis results for the a) European rabbit and b) American pika (the extra helix is marked with a black circle).

A previous study in ILs showed that European rabbit-human sequences are more related than mouse-human or European rabbit-mouse sequences (Perkins et al., 2000). Overall, the results of the phylogenetic analysis are in agreement with the evolutionary topology described for lagomorphs (Matthee et al., 2004), where, the American pika is the most divergent species. Nevertheless, the ML trees depicted in figure 3.8 correspond to gene trees that do not necessarily reflect the true species tree (Nakhleh et al., 2009; Nichols, 2001). Overall, the human IL sequences appear more closely related with those from European rabbit than with mouse and rat. This pattern had been already observed for other molecular markers and might suggest that the European rabbit might be a more appropriate animal model for the study of immunity in humans (Fischer et al., 2012; Koch et al., 2013; Perkins et al., 2000; Vaure and Liu, 2014; Vuillaumier et al., 1997; Wang et al., 2014).

6. CONCLUSIONS

In this study we sequenced and characterized ten interleukins for six lagomorph species and in the two European rabbit subspecies. As expected, most differences were observed between leporids and American pika. While these differences may lead to alterations in the biological roles of these proteins, the overall genomic organization, the location of the cysteine residues and the presence of N-glycosylation sites is well conserved. In addition, and according to divergence between the European rabbit and human for the ILs studied, the European rabbit might be a more suitable animal model for studies in the human innate immunity.

7. REFERENCES

Abrantes, J., Lopes, A.M., Dalton, K.P., Melo, P., Correia, J.J., Ramada, M., Alves, P.C., Parra, F., Esteves, P.J., 2013. **New variant of rabbit hemorrhagic disease virus, Portugal, 2012-2013**. Emerg Infect Dis 19, 1900-1902.

Afzal, N., Tahir, R., Jahan, S., 2012. **Cytokines: an ever expanding area** Biological and Biomedical Reports 2, 37-43.

Bessis, N., Boissier, M.C., 2001. **Novel pro-inflammatory interleukins: potential therapeutic targets in rheumatoid arthritis**. Joint Bone Spine 68, 477-481.

Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., Brunak, S., 2004. **Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence**. Proteomics 4, 1633-1649.

Brocker, C., Thompson, D., Matsumoto, A., Nebert, D.W., Vasiliou, V., 2010. **Evolutionary divergence and functions of the human interleukin (IL) gene family**. Hum Genomics 5, 30-55.

Brunak, S., Engelbrecht, J., Knudsen, S., 1991. **Prediction of human mRNA donor and acceptor sites from the DNA sequence**. J Mol Biol 220, 49-65.

Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K., Jones, D.T., 2013. **Scalable web services for the PSIPRED Protein Analysis Workbench**. Nucleic Acids Res 41, W349-357.

Bulaj, G., 2005. **Formation of disulfide bonds in proteins and peptides**. Biotechnol Adv 23, 87-92.

Burkhardt, A.M., Zlotnik, A., 2013. **Translating translational research: mouse models of human disease**. Cell Mol Immunol 10, 373-374.

Chamorey, A.L., Magne, N., Pivot, X., Milano, G., 2002. **Impact of glycosylation on the effect of cytokines. A special focus on oncology**. Eur Cytokine Netw 13, 154-160.

Chapman, J.A., Flux, J.E.C., 2008. **Introduction to the Lagomorpha**, in: Alves, P.C., Ferrand, N., Hackländer, K. (Eds.), Lagomorph Biology: Evolution, Ecology and Conservation. Springer, pp. 1-9.

Dalton, K.P., Nicieza, I., Abrantes, J., Esteves, P.J., Parra, F., 2014. **Spread of new variant RHDV in domestic rabbits on the Iberian Peninsula**. Vet Microbiol 169, 67-73.

Delibes-Mateos, M., Delibes, M., Ferreras, P., Villafuerte, R., 2008. **Key role of European rabbits in the conservation of the Western Mediterranean basin hotspot**. Conserv Biol 22, 1106-1117.

Desando, G., Cavallo, C., Sartoni, F., Martini, L., Parrilli, A., Veronesi, F., Fini, M., Giardino, R., Facchini, A., Grigolo, B., 2013. **Intra-articular delivery of adipose derived stromal cells attenuates osteoarthritis progression in an experimental rabbit model**. Arthritis Res Ther 15, R22.

Edgar, R.C., 2004. **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. Nucleic Acids Res 32, 1792-1797.

Esteves, P.J., Lanning, D., Ferrand, N., Knight, K.L., Zhai, S.K., van der Loo, W., 2005. **The evolution of the immunoglobulin heavy chain variable region (IgVH) in Leporids: an unusual case of transspecies polymorphism**. Immunogenetics 57, 874-882.

Fass, D., 2012. **Disulfide bonding in protein biophysics**. Annu Rev Biophys 41, 63-79.

Ferre, F., Clote, P., 2006. **DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification**. Nucleic Acids Res 34, W182-185.

Fischer, B., Chavatte-Palmer, P., Viebahn, C., Navarrete Santos, A., Duranthon, V., 2012. **Rabbit as a reproductive model for human health**. Reproduction 144, 1-10.

Garcia-Bocanegra, I., Astorga, R.J., Napp, S., Casal, J., Huerta, B., Borge, C., Arenas, A., 2010. **Myxomatosis in wild rabbit: design of control programs in Mediterranean ecosystems**. Prev Vet Med 93, 42-50.

Ge, D., Wen, Z., Xia, L., Zhang, Z., Erbajeva, M., Huang, C., Yang, Q., 2013. **Evolutionary history of lagomorphs in response to global environmental change**. PLoS One 8, e59668.

He, Y., Huang, C., Zhang, L., Lv, X.W., Li, J., 2014. **Interleukin-21, a potential diagnostic and therapeutic target for systemic lupus erythematosus**. Rheumatology international 34, 1027-1028.

Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P., Brunak, S., 1996. **Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information**. Nucleic Acids Res 24, 3439-3452.

Heinrich, P.C., Behrmann, I., Haan, S., Hermanns, H.M., Muller-Newen, G., Schaper, F., 2003. **Principles of interleukin (IL)-6-type cytokine signalling and its regulation**. The Biochemical journal 374, 1-20.

Helenius, A., Aebi, M., 2004. **Roles of N-linked glycans in the endoplasmic reticulum**. Annual review of biochemistry 73, 1019-1049.

Holste, D., Ohler, U., 2008. **Strategies for identifying RNA splicing regulatory motifs and predicting alternative splicing events**. PLoS Comput Biol 4, e21.

Ishihara, K., Hirano, T., 2002. **Molecular basis of the cell specificity of cytokine action**. Biochim Biophys Acta 1592, 281-296.

Jimenez-Garcia, A., Balongo-Garcia, R., Alconero, F.F., Araji, O.A., Martinez, G.J., Haba, M.G., Morales, L.C., Bevia, J.M., Martinez, J.C., 2004. **Intestinal wall damage in simple ileus in rabbits: immune-modulator role of somatostatin**. Hepato-gastroenterology 51, 1030-1036.

Johnston, J.B., McFadden, G., 2003. **Poxvirus immunomodulatory strategies: current perspectives**. J Virol 77, 6093-6100.

Johnston, J.B., McFadden, G., 2004. **Technical knockout: understanding poxvirus pathogenesis by selectively deleting viral immunomodulatory genes**. Cellular microbiology 6, 695-705.

Jones, D.T., 1999. **Protein secondary structure prediction based on position-specific scoring matrices**. J Mol Biol 292, 195-202.

Kaiser, P., Rothwell, L., Avery, S., Balu, S., 2004. **Evolution of the interleukins**. Developmental and comparative immunology 28, 375-394.

Kang, S.J., Grossniklaus, H.E., 2011. **Rabbit model of retinoblastoma**. Journal of biomedicine & biotechnology 2011, 394730.

Keren, H., Lev-Maor, G., Ast, G., 2010. **Alternative splicing and evolution: diversification, exon definition and function**. Nat Rev Genet 11, 345-355.

Kerr, P., McFadden, G., 2002. **Immune responses to myxoma virus**. Viral immunology 15, 229-246.

Kerr, P.J., Perkins, H.D., Inglis, B., Stagg, R., McLaughlin, E., Collins, S.V., Van Leeuwen, B.H., 2004. **Expression of rabbit IL-4 by recombinant myxoma viruses enhances virulence and overcomes genetic resistance to myxomatosis**. Virology 324, 117-128.

Koch, E., Hue-Beauvais, C., Galio, L., Solomon, G., Gertler, A., Revillon, F., Lhotellier, V., Aujean, E., Devinoy, E., Charlier, M., 2013. **Leptin gene in rabbit: cloning and expression in mammary epithelial cells during pregnancy and lactation**. Physiological genomics 45, 645-652.

Li, X.Q., Zhang, T., Donnelly, D., 2011. **Selective loss of cysteine residues and disulphide bonds in a potato proteinase inhibitor II family**. PLoS One 6, e18615.

Librado, P., Rozas, J., 2009. **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data**. Bioinformatics 25, 1451-1452.

Liu, J., Wennier, S., Reinhard, M., Roy, E., MacNeill, A., McFadden, G., 2009. **Myxoma virus expressing interleukin-15 fails to cause lethal myxomatosis in European rabbits**. J Virol 83, 5933-5938.

Lopes, A.M., Correia, J., Abrantes, J., Melo, P., Ramada, M., Magalhaes, M.J., Alves, P.C., Esteves, P.J., 2014. **Is the new variant RHDV replacing genogroup 1 in Portuguese wild rabbit populations?** Viruses 7, 27-36.

Mage, R.G., Mage, G.G., 2012. **Sequence of Rabbit (Oryctolagus cuniculus) DNA from the OryCun2.0 Donor does not Confirm a Frameshift in Exon 2 of IL4**. Immunology and Immunogenetics Insights 2012, 1-5.

Marchandeau, S., Pontier, D., Guitton, J.S., Letty, J., Fouchet, D., Aubineau, J., Berger, F., Leonard, Y., Roobrouck, A., Gelfi, J., Peralta, B., Bertagnoli, S., 2014. **Early infections by myxoma virus of young rabbits (Oryctolagus cuniculus) protected by maternal antibodies activate their immune system and enhance herd immunity in wild populations**. Vet Res 45, 26.

Marques, R.M., Costa, E.S.A., Aguas, A.P., Teixeira, L., Ferreira, P.G., 2012. **Early inflammatory response of young rabbits attending natural resistance to calicivirus (RHDV) infection**. Veterinary immunology and immunopathology 150, 181-188.

Matthee, C.A., van Vuuren, B.J., Bell, D., Robinson, T.J., 2004. **A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene**. Syst Biol 53, 433-447.

Mullane, K., Williams, M., 2014. **Animal models of asthma: reprise or reboot?** Biochemical pharmacology 87, 131-139.

Muller, A., Freitas, J., Silva, E., Le Gall-Recule, G., Zwingelstein, F., Abrantes, J., Esteves, P.J., Alves, P.C., van der Loo, W., Kolodziejek, J., Nowotny, N., Thompson, G., 2009. **Evolution of rabbit haemorrhagic disease virus (RHDV) in the European rabbit (Oryctolagus cuniculus) from the Iberian Peninsula**. Vet Microbiol 135, 368-373.

Nakhleh, L., Ruths, D., Innan, H., 2009. **Gene trees, species trees, and species networks**, in: Guerra, R., Goldstein, D. (Eds.), Meta-analysis and Combining Information in Genetics and Genomics. CRC Press, Boca Raton, FL, pp. 275-293.

Neves, F., Abrantes, J., Pinheiro, A., Almeida, T., Costa, P.P., Esteves, P.J., 2014a. **Convergent evolution of IL-6 in two leporids (Oryctolagus and Pentalagus) originated an extended protein**. Immunogenetics 66, 589-595.

Neves, F., Abrantes, J., Steinke, J.W., Esteves, P.J., 2014b. **Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals**. Innate immunity 20, 184-191.

Nichols, R., 2001. **Gene trees and species trees are not the same**. Trends Ecol Evol 16, 358-364.

O'Connell, M.J., McInerney, J.O., 2005. **Gamma chain receptor interleukins: evidence for positive selection driving the evolution of cell-to-cell communicators in the mammalian immune system**. Journal of molecular evolution 61, 608-619.

Perkins, H.D., van Leeuwen, B.H., Hardy, C.M., Kerr, P.J., 2000. **The complete cDNA sequences of IL-2, IL-4, IL-6 AND IL-10 from the European rabbit (Oryctolagus cuniculus)**. Cytokine 12, 555-565.

Pinheiro, A., de Mera, I.G., Alves, P.C., Gortazar, C., de la Fuente, J., Esteves, P.J., 2013. **Sequencing of modern Lepus VDJ genes shows that the usage of VHn genes has been retained in both Oryctolagus and Lepus that diverged 12 million years ago**. Immunogenetics 65, 777-784.

Pinheiro, A., Lanning, D., Alves, P.C., Mage, R.G., Knight, K.L., van der Loo, W., Esteves, P.J., 2011. **Molecular bases of genetic diversity and evolution of the immunoglobulin heavy chain variable region (IGHV) gene locus in leporids**. Immunogenetics 63, 397-408.

Rossi, J., Lu, Z., Jourdan, M., Klein, B., 2015. **Interleukin-6 as a Therapeutic Target**. Clinical Cancer Research.

Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A., Dwek, R.A., 2001. **Glycosylation and the immune system**. Science 291, 2370-2376.

Sahoo, A., Im, S.H., 2010. **Interleukin and interleukin receptor diversity: role of alternative splicing**. International reviews of immunology 29, 77-109.

Schwarz, F., Aebi, M., 2011. **Mechanisms and principles of N-linked protein glycosylation**. Current opinion in structural biology 21, 576-582.

Seok, J., Warren, H.S., Cuenca, A.G., Mindrinos, M.N., Baker, H.V., Xu, W., Richards, D.R., McDonald-Smith, G.P., Gao, H., Hennessy, L., Finnerty, C.C., Lopez, C.M., Honari, S., Moore, E.E., Minei, J.P., Cuschieri, J., Bankey, P.E., Johnson, J.L., Sperry, J., Nathens, A.B., Billiar, T.R., West, M.A., Jeschke, M.G., Klein, M.B., Gamelli, R.L., Gibran, N.S., Brownstein, B.H., Miller-Graziano, C., Calvano, S.E., Mason, P.H., Cobb, J.P., Rahme, L.G., Lowry, S.F., Maier, R.V., Moldawer, L.L., Herndon, D.N., Davis, R.W., Xiao, W., Tompkins, R.G., 2013. **Genomic responses in mouse models poorly mimic human inflammatory diseases**. Proc Natl Acad Sci U S A 110, 3507-3512.

Shay, T., Lederer, J.A., Benoist, C., 2015. **Genomic responses to inflammation in mouse models mimic humans: We concur, apples to oranges comparisons won't do**. Proc Natl Acad Sci U S A 112, E346.

Shental-Bechor, D., Levy, Y., 2008. **Effect of glycosylation on protein folding: a close look at thermodynamic stabilization**. Proc Natl Acad Sci U S A 105, 8256-8261.

Shimada, M.K., Hayakawa, Y., Takeda, J., Gojobori, T., Imanishi, T., 2010. **A comprehensive survey of human polymorphisms at conserved splice dinucleotides and its evolutionary relationship with alternative splicing**. BMC Evol Biol 10, 122.

Siewe, B.T., Kalis, S.L., Esteves, P.J., Zhou, T., Knight, K.L., 2010. **A novel functional rabbit IL-7 isoform**. Dev Comp Immunol 34, 828-836.

Stanford, M.M., Werden, S.J., McFadden, G., 2007. **Myxoma virus in the European rabbit: interactions between the virus and its susceptible host**. Vet Res 38, 299-318.

Takao, K., Miyakawa, T., 2015. **Genomic responses in mouse models greatly mimic human inflammatory diseases**. Proc Natl Acad Sci U S A 112, 1167-1172.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. **MEGA6: Molecular Evolutionary Genetics Analysis version 6.0**. Mol Biol Evol 30, 2725-2729.

Teixeira, L., Marques, R.M., Aguas, A.P., Ferreira, P.G., 2012. **Regulatory T cells are decreased in acute RHDV lethal infection of adult rabbits**. Veterinary immunology and immunopathology 148, 343-347.

Tian, J., Hu, S., Sun, Y., Ban, X., Yu, H., Dong, N., Wu, J., Yu, B., 2012. **A novel model of atherosclerosis in rabbits using injury to arterial walls induced by ferric chloride as evaluated by optical coherence tomography as well as intravascular ultrasound and histology**. Journal of biomedicine & biotechnology 2012, 121867.

van der Loo, W., Abrantes, J., Esteves, P.J., 2009. **Sharing of endogenous lentiviral gene fragments among leporid lineages separated for more than 12 million years**. J Virol 83, 2386-2388.

Vande Walle, L., Lamkanfi, M., 2011. **Inflammasomes: caspase-1-activating platforms with critical roles in host defense**. Frontiers in microbiology 2, 3.

Vaure, C., Liu, Y., 2014. **A comparative review of toll-like receptor 4 expression and functionality in different animal species**. Front Immunol 5, 316.

Vuillaumier, S., Kaltenboeck, B., Lecointre, G., Lehn, P., Denamur, E., 1997. **Phylogenetic analysis of cystic fibrosis transmembrane conductance regulator gene in mammalian species argues for the development of a rabbit model for cystic fibrosis**. Mol Biol Evol 14, 372-380.

Wang, Y., Fan, N., Song, J., Zhong, J., Guo, X., Tian, W., Zhang, Q., Cui, F., Li, L., Newsome, P.N., Frampton, J., Esteban, M.A., Lai, L., 2014. **Generation of knockout rabbits using transcription activator-like effector nucleases**. Cell Regen (Lond) 3, 3.

Webb, D.R., 2014. **Animal models of human disease: inflammation**. Biochemical pharmacology 87, 121-130.

Wong, J.W., Ho, S.Y., Hogg, P.J., 2011. **Disulfide bond acquisition through eukaryotic protein evolution**. Mol Biol Evol 28, 327-334.

Woodruff-Pak, D.S., Agelan, A., Del Valle, L., 2007. **A rabbit model of Alzheimer's disease: valid at neuropathological, cognitive, and therapeutic levels**. Journal of Alzheimer's disease : JAD 11, 371-383.

Wu, Q., Krainer, A.R., 1999. **AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes**. Molecular and cellular biology 19, 3225-3236.

Zelus, D., Robinson-Rechavi, M., Delacre, M., Auriault, C., Laudet, V., 2000. **Fast evolution of interleukin-2 in mammals and positive selection in ruminants**. Journal of molecular evolution 51, 234-244.

Zhang, J., Nei, M., 2000. **Positive selection in the evolution of mammalian interleukin-2 genes**. Molecular biology and evolution 17, 1413-1416.

8. SUPPLEMENTARY MATERIAL

Table 3. 10. Primers used for amplification of each interleukin in lagomorphs

		Reference sequence	Primers sequence (5'- 3')	Primer name	Exons amplified	Species amplified
IL1A	Leporids	OryCun2.0:2:97559306:97569785:1	CAATCTAGTAGAGGAGGCAG	IL1AOrcuF1.1	Exon 2	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit
			GCACAGCAATGGCTTGAATC	IL1AOrcuR1		
			CAAACTCCTGCTGAAAGRGTC	IL1AOrcuF2b	Exon 3 + exon 4	
			GAAGAAGGTAAGGGGTAAG	IL1AOrcuR2b		
			CTGTGCTTGTAGGAGAGTG	IL1AOrcuF3	Exon 5 + exon 6	
			CATCTGAATTCTCACCCAGC	IL1AOrcuR3a		
			GGTCATTGGCAATCTTCCTG	IL1AOrcuF4	Exon 7	European rabbit, Brush rabbit, European brown hare
			CTGCTTGAAGCTGAGCTTC	IL1AOrcuR4		
			CACAGTCATGGTCCTRAAG	IL1AOrcuF4a		Pygmy rabbit
			GAAGCCAAGTCTAATGGSAC	IL1AOrcuR4a		
		NM_001101684.1	ATGGCCAAAGTCCCTGAYC	IL1AOrcuFcDNA	cDNA	European rabbit, Eastern cottontail
			CTGAKTTTCAGATCTCATAA	IL1AOrcuRcDNA		
	American pika	GeneScaffold_4797:475551:484012:-1	CTTCACAGGCATCTGGTAC	IL1AOcprF1a	Exon1 + exon 2	American pika
			GGCGAGCAGTAGCTTGAATC	IL1AOcprR1		
			GGAGATCTTCCTAAGTCTGC	IL1AOcprF2	Exon 3	
			GGTGTCAAGATAACAAGAGGC	IL1AOcprR2		
			GTATCGTCTGTGACAGTACC	IL1AOcprF3	Exon 4 + exon 5	
			CAGAGAAGCCAACTGGTATC	IL1AOcprR3		
			CTGAGCCATACATCTAACAG	IL1AOcprF4a	Exon 6 + exon 7	
			GTGCTGAGCCTGTGTAATTC	IL1AOcprR4b		
IL1B	Leporids		ATGGCAACAGTACCTGAGCTC	IL1BOrcuF1	Exon1 + exon 2 + exon 3	European rabbit, Brush rabbit, Pygmy rabbit,
			GTGCTTGGAGACTAGATGTG	IL1BOrcuRex1		
			CTGCAYTTCATGCTCAGCTG	IL1BOrcuRex2		
			CTCAGAATCCATGGCTCAACAG	IL1BOrcuR1		
			CTAACCTWACTGTGTGCTC	IL1BOrcuF1e		European brown hare
			CCTTAATCTTTGAAGAAGG	IL1BOrcuR1e		
			TGACCACCCGTGCTTAGATGT	IL1BOrcuF2	Exon 4 + exon 5 + exon 6	European rabbit, Brush rabbit, European brown hare
			GACATCCTGGTTTCCCTTC	IL1BOrcuRint2		
			CATCACTCACTTCCCACACAG	IL1BOrcuR2		
			CACCTGTGCTCTCTGTTTC	IL1BorcuF2a		Pygmy rabbit,
			GWCCACCTTTCACTGCTGC	IL1BOrcuF2d	Exon 6	European brown hare
			CTTCAACCTGCTGAACCACA	IL1BOrcuF3	Exon 7	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			GGTGCTGAATGTGTTCACC	IL1BOrcuR3c		
			ATGGCAACAGTACCTGAGC	IL1BOrcuFcDNA	cDNA	Eastern cottontail
			GGAATTCGTGTCTTCCTRA	IL1BOrcuRcDNA		
	American pika		CTCATTATGCGACGATGCC	IL1BOcprF1a	Exon1 + exon 2 + exon 3	American pika
			GTTCTGGACCACCTTCTTC	IL1BOcprRint1		
			CTGTGAGCAGATAGTACAGC	IL1BOcprR1		
			GACCATCCATGCTTAGGTGT	IL1BOcprF2	Exon 4 + exon 5	
			TGGCCATCACAGCCATCTGTT	IL1BOcprR2		
			CAGGTACTCTACTAAGACACGC	IL1BOcprF3	Exon 6	
			CACGTTGTGATGCCTCACTG	IL1BOcprR3b		
IL2	Leporids	OryCun2.0:15:98312429:98317577:1	CACGCTCTGTAATCACTACTC	IL2OrcuF1	Exon1 + exon 2	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			GAGTACTATATAGCACCTC	IL2OrcuR1a		
			CTGGTCATAGCTACTGGAG	IL2OrcuF2	Exon 3	European rabbit, Brush rabbit,
			CATGCATCCAGAGCCAAG	IL2OrcuR2b		
			CAGGTCACAGAATTGAAAC	IL2OrcuF2c		European brown hare, Pygmy rabbit
			GYTCTCTCAGAAGTAAAM	IL2OrcuR2c		
			GGATGCTCATCCTCATTG	IL2OrcuF3	Exon 4	Brush rabbit, Pygmy rabbit,
			GGGCTCTAAAATGGCTTCAC	IL2OrcuR3		
			CTGGCTCTATCCTGGTTTG	IL2OrcuF3b		European rabbit, European brown hare,
			GGATGTATTGTTTGCTACC	IL2OrcuR3a		
		ENSOCUT00000010098	CACGCTCTGTAATCACTACTC	IL2OrcuF1	cDNA	Eastern cottontail
			GGGCTCTAAAATGGCTTCAC	IL2OrcuR3		
	American pika	scaffold_6999:1:91496:1	GCTTCTCTTGCTTAAGAGC	IL2OcprF1a	Exon1 + exon 2	American pika
			GAGCAGAACACCTGAAGTG	IL2OcprR1		
			GTATAAGGGAAATGGTGGCTG	IL2OcprF2a	Exon 3	
			CCTCAGTTCCACAGTTCAAC	IL2OcprR2		
			GGAGAAGCTCAGCCTTAAC	IL2OcprF3	Exon 4	
			CCGTTCAAGACAAATCAGAC	IL2OcprR3a		
IL4	Leporids	oryCun2:3:15702738:15715033:1	GGTAGCGTCTCCTGATAAAC	IL4OrcuF1	Exon1 + exon 2	European rabbit,
			GAGACTTCTCCCMAKAC	IL4OrcuF1c		
			GAGAATCTGCGTCTTCACG	IL4OrcuR1		
			CATTGGTAGCGTCTCCTG	IL4OrcuF1d		European brown hare, Brush rabbit, Pygmy rabbit,
			CCAGGCTATCTGTAAGGTG	IL4OrcuR1b		
			GTGGCCGTCTGGTATTTCTC	IL4OrcuF2a	Exon 3	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			GTGCTCTTCCAGACACTAGG	IL4OrcuR2		
			GCTTCCCAATGCAGCTAAC	IL4OrcuF3a	Exon 4	
			CTTCTCAACGCTGTTCCTC	IL4OrcuR3a		

IL8		NM_001163177.1	ATGGGGCTCCCTGCCCAG	IL4OrcuFcDNA	cDNA	European rabbit, European brown hare
			CTCWAAGCGTCRGAG	IL4OrcuRcDNA		
	American pika	GeneScaffold_3588:636330:644015:1	CTTCATCAGAATGCATCGG	IL4OrcuF1a	Exon1 + exon 2	American pika
			CTGTCCACAAGGGTTAGAAC	IL4OcprR1		
			GAGTACTTGGCTACAGCTC	IL4OcprF2	Exon 3	
			CTCCTGTTCAGGAGATGTAG	IL4OcprR2		
			CAAGTCCATAGGCATGAAAG	IL4OcprF3	Exon 4	
			CAAAGCCAAGGTATGCAAC	IL4OcprR3a		
	Leporids	oryCun2:15:76368959:76371691:-1	GAAGAAACCACTTCGCCTG	IL8OrcuF1a	Exon 1	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			CATCTGTGTGCCACTTGTG	IL8OrcuRint1a		European rabbit, European brown hare, Pygmy rabbit,
			GTCATCTGTATGCAACTC	IL8OrcuRint1b		Brush rabbit,
			GACAGTTCTCGGTACTGCA	IL8OrcuF2a	Exon 2 + exon 3 + exon 4	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			CTGCATGGATCTGTCGTAG	IL8OrcuR2a		
		NM_001082293.1	GAAGAAACCACTTCGCCTG	IL8OrcuF1a	cDNA	European rabbit, European brown hare, Eastern cottontail
			CTGCATGGATCTGTCGTAG	IL8OrcuR2a		
	American pika	scaffold_1225:1:361409:1	GCAGCAAGGAGCCAGAAG	IL8OcprF1	Exon1 + exon 2	American pika
			CAGTATACTGCAAAGTTGGG	IL8OcprR1		
			CTGTGCTTCAAGGCCCATG	IL8OcprF2	Exon 3 + exon 4	
			GGATCTGTCATAGAGTTGC	IL8OcprR3b		
		ENSOPRT00000002239	ATGACTCCCAAGGTGGCTATG	IL8OcprFcDNA	cDNA	
			GCTGAAAATCAAGCTTCATGA	IL8OcprRcDNA		
IL10	Leporids	OryCun2:16:65583171:65589395:1	GCAAAAGCAAACCACAAGGC	IL10OrcuF1b	Exon1 + exon 2	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			GTCCAGTTCTRTGTTGACTG	IL10OrcuRint1a		
			GATCTTAACTCAGCACTCAGC	IL10OrcuR1		
			GGAGGGTTACAAAGGACTAG	IL10OrcuF2a	Exon 3 + exon 4	
			CTGTGAAGTGTCTTTGGGG	IL10OrcuR2a		
			CTCGGAGTGAAGATGCTTAG	IL10OrcuF3	Exon 5	European rabbit, Brush rabbit,
			GGAAAGCTGTTGTACCTCTC	IL10OrcuR3a		European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			CAAAGTGACTCTCACCTTCC	IL10OrcuF3a		European brown hare, Pygmy rabbit,
		AF068058.1	GCAAAAGCAAACCACAAGGC	IL10OrcuF1b	cDNA	European rabbit, European brown hare,
			GGAAAGCTGTTGTACCTCTC	IL10OrcuR3a		
	American pika		GTGGAGCCTACCCATCAAG	IL10OcprF1b	Exon1 + exon 2 + exon3	American pika
			GAGCATCAGAACAGCATGC	IL10OcprF2a		
			CATTCTGCAAGAGCTCCAAG	IL10OcprR1a		
			GTTGTGAACCTGCATAAGGC	IL10OcprF2	Exon 4 + exon 5	
			CTCTGAGGCAGCTGACCCAGC	IL10OcprR2b		
		XM_004578691.1	ATGCTCGGCTCTGCTCTAC	IL10OcprF1a	cDNA	
			CCACGAGGATGAAAAGCTAA	IL10OcprR2a		
IL12A	Leporids	OryCun2.0:14:54457840:54464306:1	CACGAGTCCACCCGCCTG	IL12AOrcuF1a	Exon1 + exon 2	European rabbit,
			CAGCAGCATCCGCCCTC	IL12AOrcuF1b		European brown hare, Brush rabbit, Pygmy rabbit,
			GTCTACACAGAGGAAACGAG	IL12AOrcuR1		European rabbit, European brown hare, Brush rabbit, Pygmy rabbit,
			GTTAGCAAGCTGAAGTTCTCC	IL12AOrcuF2	Exon 3 + exon 4 + exon 5 + exon 6	European rabbit,
			CTGCCTCATTCTGAAGATGG	IL12AOrcuR2		
			CTCCAGTCTCATGGGAAG	IL12AOrcuF2b		European brown hare, Brush rabbit, Pygmy rabbit,
			GAAAGCCAGGCAGAACTTC	IL12AOrcuR2b		
			GTCATRCTCTGTGTTCATTC	IL12AOrcuF3a	Exon 7	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit
			GTTGACCTGGTCCTACTG	IL12AOrcuR3a		
		ENSOCUT00000010149	ATGTGCTCCCTGCGCGG	IL12AOrcuFcDNA		European rabbit, Eastern cottontail
			GCTATCTGAGTTCTTCCTGA	IL12AOrcuRcDNA		
IL12B	Leporids	OryCun2.0:3:40708623:40721869:1	GCTGCATAATTGCTGACATG	IL12BOrcuF1a	Exon 1	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit
			CTCAGAGTAAGGAGCCAACA	IL12BOrcuR1		
			CAGAGTGCATGGAGCTTAC	IL12BOrcuF2	Exon 2 + exon 3	
			CTGAGAATAGTCGCTGCCATG	IL12BOrcuR2		
			CAGGCCTTATCTGTCTCAC	IL12BOrcuF3a	Exon 4 + exon 5	European brown hare,
			CACACCAGGAAGATGATGAAG	IL12BOrcuR3a		
			GCTCCCGTTAGTGTATAAC	IL12BOrcuF3b		European rabbit, Brush rabbit, Pygmy rabbit
			GTAAGGTACGGTCTGGGTTC	IL12BOrcuR3b		
			CTGTCTTGTCACTGCATTGC	IL12BOrcuF4	Exon 6	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit
			CCCAGAGTCACAAGCTAAC	IL12BOrcuR4b		
	American pika	scaffold_4486:1:129844:1	CAACATTATTGACATGCCC	IL12BOcprF1a	Exon 1 + exon 2	American pika
			CTGACACCTGCAAAGAACTTCC	IL12BOcprR1		
			GTGCAATGTAGTACTCTGCG	IL12BOcprF2	Exon 3 + exon 4	
			CTTAAGCAACTTGCTCCAGG	IL12BOcprR2		
			CCATGTGCTCAGAGAAGTG	IL12BOcprF3	Exon 5 + exon 6	
			GGGAGCATAGGGCCATAG	IL12BOcprR3a		
IL15	Leporids	oryCun2:15:22774049:22789507:-1	GTCATAGCCAGCTCTTCTTC	IL15OrcuF1	Exon 1 + exon 2	European rabbit, European brown hare, Brush rabbit, Pygmy rabbit
			GAGCTGTATTGGAAGCAGAG	IL15OrcuR1		
			GCACTRGTGYGCTCATTCCA	IL15OrcuF2b	Exon 3	
			CTCCTTGTGTTTTCTGTCAC	IL15OrcuR2b		
			CCTGTACAAGAAGAGCAGG	IL15OrcuF3	Exon 4	
			CTTAGGCTGAGTTAGGCATG	IL15OrcuR3		
			CCTCTCTCTGTGTAACTCTG	IL15OrcuF4	Exon 5	
			CCATCCAGAATTACTCATGCGC	IL15OrcuR4		

			GAAGTGCATTAGGCATGCC	IL15OrcuF5	Exon 6	
			GYGGATCCTGTGCAGTG	IL15OrcuR5a		
		NM_001082216.2	ATGAGAATTTCGAAACCG	IL15OrcuFcDNA	cDNA	European rabbit, European brown hare
			GTTCATCAATTCTCCTTGA	IL15OrcuRcDNA		
	Ameri can pika	ENSOPRT00000008393	ATGACAATTTCGAAACC	IL15OcprFcDNA	cDNA	American pika
			GTTCATCAACTCTCCTTGA	IL15OcprR4		
IL18	Leporids	oryCun2:1:104225304:104234571:1	GAGGTGGCACAAGTTAGTAG	IL18OrcuF1a	Exon 1 + exon 2	*European rabbit, European brown hare, Brush rabbit, Pygmy rabbit*
			GACACATCCATATCACCACG	IL18OrcuR1		
			GTAGGCAACACAGCAGTAAG	IL18OrcuF2	Exon 3	*European rabbit, Brush rabbit, Pygmy rabbit*
			yCTGAAATACATGRTAGGTC	IL18OrcuF2a		*European brown hare*
			CTTAGATGAAGTCACAGGAG	IL18OrcuR2		*European rabbit, European brown hare, Brush rabbit, Pygmy rabbit*
			GCCTGATCTTCTGCTGATTG	IL18OrcuF3	Exon 4	*European rabbit, Brush rabbit,*
			CTAGAGCACGAACAGCAG	IL18OrcuR3		
			CTCAGAGTTCAAGGTTGGC	IL18OrcuF3a		*European brown hare, Pygmy rabbit*
			GCTGTCACAGGGAAGATAG	IL18OrcuR3a		
			CTGCACATGTTCAAATTGCC	IL18OrcuF4a	Exon 5	*European rabbit, European brown hare, Brush rabbit, Pygmy rabbit*
			CAGTCTATGAGCCTTATCTAGTC	IL18OrcuR4a		
		NM_001122940.1	ATGGCTGCTGAACCAGAAG	Il18FcDNA	cDNA	*European rabbit, European brown hare, Eastern cottontail*
			CAGTGTCCAAAACAAGAATTAG	IL18OrcuR4		

Table 3. 11. Estimates of evolutionary divergence between lagomorphs' sequences: a) IL1α; b) IL1β; c) IL2; d) IL4: e) IL8; f) IL10; g) IL12A; h) IL12B; i) IL15; j) IL18.

a) IL1α	1	2	3	4	5	6	7	8	9	10	11	12	13
1 European rabbit NM 001101684.1													
2 European rabbit (O. c. cuniculus) *1	0.007												
3 European rabbit (O. c. cuniculus) *2	0.007	0.007											
4 European rabbit (O. c. cuniculus) *3	0.000	0.007	0.007										
5 European rabbit (O. c. algirus) *1	0.011	0.011	0.019	0.011									
6 European rabbit (O. c. algirus) *2	0.000	0.007	0.007	0.000	0.011								
7 Pygmy rabbit	0.037	0.037	0.037	0.037	0.045	0.037							
8 Brush rabbit	0.034	0.034	0.034	0.034	0.041	0.034	0.019						
9 Eastern cottontail	0.045	0.045	0.045	0.045	0.052	0.045	0.030	0.019					
10 European brown hare *1	0.052	0.052	0.052	0.052	0.060	0.052	0.045	0.049	0.060				
11 European brown hare *2	0.056	0.056	0.056	0.056	0.064	0.056	0.041	0.052	0.064	0.004			
12 American pika XM 004590824.1	0.282	0.282	0.286	0.282	0.289	0.282	0.297	0.297	0.297	0.301	0.301		
13 American pika	0.289	0.289	0.293	0.289	0.297	0.289	0.305	0.305	0.305	0.308	0.308	0.011	-

b) IL1β	1	2	3	4	5	6	7	8	9	10	11	12	13
1 European rabbit NM 001082201.1													
2 European rabbit (O. c. cuniculus) *1	0.000												
3 European rabbit (O. c. cuniculus) *2	0.004	0.004											
4 European rabbit (O. c. algirus) *1	0.000	0.000	0.004										
5 European rabbit (O. c. algirus) *2	0.007	0.007	0.011	0.007									
6 Pygmy rabbit	0.037	0.037	0.041	0.037	0.037								
7 Brush rabbit *1	0.045	0.045	0.049	0.045	0.045	0.030							
8 Brush rabbit *2	0.049	0.049	0.052	0.049	0.049	0.034	0.004						
9 Eastern cottontail *1	0.037	0.037	0.041	0.037	0.037	0.015	0.015	0.019					
10 Eastern cottontail *2	0.041	0.041	0.045	0.041	0.041	0.019	0.019	0.022	0.004				
11 European brown hare	0.041	0.041	0.045	0.041	0.041	0.019	0.049	0.052	0.034	0.037			
12 American pika XM 004590743.1	0.222	0.222	0.222	0.222	0.218	0.208	0.208	0.204	0.208	0.212	0.215		
13 American pika	0.218	0.218	0.218	0.218	0.215	0.204	0.204	0.200	0.204	0.208	0.211	0.004	-

c) IL2	1	2	3	4	5	6	7	8	9	10	11
1 European rabbit NM 001163180.1											
2 European rabbit δ2 AF169168.1	0.000										
3 European rabbit (O. c. cuniculus)	0.000	0.000									
4 European rabbit (O. c. algirus)	0.000	0.000	0.000								
5 Pygmy rabbit	0.020	0.023	0.020	0.020							
6 Brush rabbit	0.013	0.015	0.013	0.013	0.007						
7 Eastern cottontail	0.013	0.015	0.013	0.013	0.013	0.007					
8 European brown hare. *1	0.039	0.045	0.039	0.039	0.033	0.026	0.033				
9 European brown hare *2	0.046	0.053	0.046	0.046	0.039	0.033	0.039	0.007			
10 American pika XM 004594567.1	0.138	0.159	0.138	0.138	0.118	0.125	0.132	0.125	0.125		
11 American pika	0.145	0.167	0.145	0.145	0.125	0.132	0.138	0.132	0.132	0.019	-

d) IL4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 European rabbit NM 001163177.1															
2 European rabbit δ2 AF169170.1	0.000														
3 European rabbit δ3 AF169171.1	0.000	0.000													
4 European rabbit (O. c. cuniculus) *1	0.020	0.023	0.000												
5 European rabbit (O. c. cuniculus) *2	0.007	0.008	0.000	0.014											
6 European rabbit (O. c. cuniculus) *3	0.000	0.000	0.000	0.020	0.007										
7 European rabbit (O. c. algirus) *1	0.007	0.008	0.000	0.014	0.000	0.007									
8 European rabbit (O. c. algirus) *2	0.000	0.000	0.000	0.020	0.007	0.000	0.007								
9 Pygmy rabbit	0.075	0.084	0.033	0.095	0.082	0.075	0.082	0.075							
10 Brush rabbit	0.068	0.076	0.050	0.088	0.075	0.068	0.075	0.068	0.068						
11 European brown hare *1	0.075	0.076	0.033	0.088	0.082	0.075	0.082	0.075	0.095	0.082					
12 European brown hare *2	0.082	0.084	0.050	0.095	0.088	0.082	0.088	0.082	0.102	0.088	0.007				
13 American pika XM 004586394.1	0.252	0.260	0.183	0.265	0.259	0.252	0.259	0.252	0.265	0.279	0.272	0.279			
14 American pika *1	0.259	0.267	0.200	0.272	0.265	0.259	0.265	0.259	0.265	0.279	0.279	0.286	0.007		
15 American pika *2	0.265	0.275	0.200	0.279	0.272	0.265	0.272	0.265	0.272	0.286	0.286	0.293	0.014	0.007	-

e) IL8	1	2	3	4	5	6	7	8	9	10
1 European rabbit NM 001082293.1										
2 European rabbit (O. c. cuniculus)	0.000									
3 European rabbit (O. c. algirus)	0.000	0.000								
4 Pygmy rabbit	0.020	0.020	0.020							
5 Brush rabbit *1	0.010	0.010	0.010	0.030						
6 Eastern cottontail	0.020	0.020	0.020	0.040	0.010					
7 European brown hare *1	0.010	0.010	0.010	0.030	0.020	0.030				
8 European brown hare *2	0.010	0.010	0.010	0.030	0.020	0.030	0.010			
9 American pika ENSOPRG00000002251	0.206	0.206	0.206	0.216	0.216	0.227	0.206	0.206		
10 American pika	0.198	0.198	0.198	0.208	0.208	0.218	0.198	0.198	0.031	-

f) IL10	1	2	3	4	5	6	7	8	9	10	11	12	13
1 European rabbit *1NM 001082045.1													
2 European rabbit *2 NM 001082045.1	0.011												
3 European rabbit AF068058.1	0.006	0.011											
4 European rabbit (O. c. cuniculus) *1	0.006	0.011	0.006										
5 European rabbit (O. c. cuniculus) *2	0.006	0.011	0.000	0.006									
6 European rabbit (O. c. algirus) *1	0.006	0.011	0.000	0.006	0.000								
7 European rabbit (O. c. algirus) *2	0.011	0.017	0.006	0.011	0.006	0.006							
8 Pygmy rabbit	0.045	0.051	0.039	0.045	0.039	0.039	0.034						
9 Brush rabbit *1	0.034	0.039	0.028	0.034	0.028	0.028	0.022	0.034					
10 Brush rabbit *2	0.039	0.045	0.034	0.039	0.034	0.034	0.028	0.039	0.006				
11 European brown hare	0.045	0.051	0.039	0.045	0.039	0.039	0.034	0.034	0.034	0.039			
12 American pika XM 004578691.1	0.140	0.146	0.135	0.140	0.135	0.135	0.129	0.118	0.118	0.124	0.135		
13 American pika	0.140	0.146	0.135	0.140	0.135	0.135	0.129	0.118	0.118	0.124	0.135	0.000	-

g) IL12A	1	2	3	4	5	6	7	8	9	10	11
1 European rabbit XM 008266409.1											
2 European rabbit (O. c. cuniculus) *1	0.009										
3 European rabbit (O. c. cuniculus) *2	0.000	0.009									
4 European rabbit (O. c. algirus)	0.014	0.005	0.014								
5 Pygmy rabbit	0.027	0.027	0.027	0.032							
6 Brush rabbit *1	0.023	0.023	0.023	0.027	0.032						
7 Brush rabbit *2	0.027	0.027	0.027	0.032	0.037	0.005					
8 Eastern cottontail *1	0.023	0.023	0.023	0.027	0.032	0.018	0.023				
9 Eastern cottontail *2	0.018	0.018	0.018	0.023	0.027	0.014	0.018	0.005			
10 European brown hare	0.023	0.023	0.023	0.027	0.023	0.027	0.032	0.027	0.023		
11 American pika ENSOPRG00000006280	0.087	0.087	0.087	0.087	0.096	0.087	0.096	0.096	0.087	0.096	-

h) IL12B	1	2	3	4	5	6	7
1 European rabbit XM 002710347.2							
2 European rabbit (O. c. cuniculus)	0.009						
3 European rabbit (O. c. algirus)	0.009	0.006					
4 Brush rabbit	0.056	0.058	0.052				
5 European brown hare	0.037	0.040	0.037	0.055			
6 American pika XM 004587479.1	0.136	0.137	0.131	0.143	0.134		
7 American pika	0.136	0.131	0.131	0.140	0.134	0.012	-

i) IL15	1	2	3	4	5	6	7	8	9	10	11	12
1 European rabbit NM 001082216.2												
2 European rabbit XM 008267272.1	0.000											
3 European rabbit (O. c. cuniculus)	0.000	0.000										
4 European rabbit (O. c. algirus)	0.000	0.000	0.000									
5 Pygmy rabbit	0.031	0.031	0.031	0.031								
6 Pygmy rabbit	0.043	0.043	0.043	0.043	0.019							
7 Brush rabbit	0.025	0.025	0.025	0.025	0.031	0.043						
8 European brown hare *1	0.056	0.056	0.056	0.056	0.056	0.062	0.056					
9 European brown hare *2	0.043	0.043	0.043	0.043	0.043	0.049	0.043	0.012				
10 American pika XM004588185.1	0.099	0.099	0.099	0.099	0.111	0.123	0.105	0.142	0.130			
11 American pika ENSOPRG00000008400	0.083	0.083	0.083	0.083	0.090	0.105	0.090	0.120	0.120	0.030		
12 American pika	0.086	0.086	0.086	0.086	0.099	0.111	0.093	0.130	0.117	0.019	0.008	-

j) IL18	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 European rabbit *1 NM 001122940.1														
2 European rabbit *2 XM 008258609.1	0.051													
3 European rabbit (O. c. cuniculus) *1	0.005	0.051												
4 European rabbit (O. c. cuniculus) *2	0.000	0.051	0.005											
5 European rabbit (O. c. algirus)	0.005	0.051	0.000	0.005										
6 Pygmy rabbit *1	0.026	0.069	0.021	0.026	0.021									
7 Pygmy rabbit *2	0.021	0.063	0.016	0.021	0.016	0.005								
8 Eastern cottontail	0.031	0.080	0.036	0.031	0.036	0.016	0.021							
9 Brush rabbit	0.036	0.074	0.031	0.036	0.031	0.031	0.036	0.026						
10 European brown hare *1	0.042	0.086	0.036	0.042	0.036	0.026	0.031	0.041	0.052					
11 European brown hare *2	0.036	0.086	0.042	0.036	0.042	0.031	0.036	0.026	0.047	0.016				
12 European brown hare *3	0.042	0.091	0.047	0.042	0.047	0.036	0.041	0.031	0.052	0.021	0.005			
13 American pika XM 004585222.1	0.417	0.486	0.411	0.417	0.411	0.402	0.402	0.415	0.421	0.396	0.402	0.409		
14 American pika ENSOPRG00000012703	0.200	0.250	0.200	0.200	0.200	0.190	0.190	0.198	0.216	0.181	0.181	0.190	0.017	-

Table 3. 12. Summary of the amino acid alterations observed between the sequences reported in this work and those already available in public databases for the ILs studied.

	European rabbit	**American pika**
IL1α	30(Leu-Val); 128(Thr-Ile); 132(Ile-Asn); 138(Thr-Ile); 247(Phe-Leu)	105(Ser-Asn); 155(Tyr-Phe); 274(His-Gln)
IL1β	68(Val-Ile); 108(Trp-Arg); 171(Asn-Ser)	17(Ser-Gly)
IL2	-	12(Leu-Thr); 88(Glu-Asp); 138(Thr-Met)
IL4	69(Val-Ile); 105(Arg-Lys); 118(Arg-Val)	988(His-Asn); 122(Asn-Tyr)
IL8	-	Insertion 4 aa (50-53) No coherence in the last part of sequences.
IL10	-	-
IL12A	-	-
IL12B	Insertion 4aa (179-182); 21(Val-Met); 234(Ile-Val)	25(Gln-Glu); 139(Arg-His); 219(Leu-Arg); 234(Ile-Val)
IL15	-	-
IL18	33(Thr-Ser)	-

CHAPTER 4

Innate immunity – Genetic aspects of CC motif chemokines in lagomorphs,

Paper V. Neves F, Abrantes J, Lissovsky AA, Esteves PJ. (2015) Pseudogenization of CCL14 in the Ochotonidae (pika) family. Innate immunity 21(6):647-654

Paper VI. Neves F, Abrantes J, Esteves PJ. (2016) Evolution of CCL11: genetic characterization in lagomorphs and evidence of positive and purifying selection in mammals. Innate immunity 22(5):336-43. doi: 10.1177/1753425916647471

Paper VII. Neves F, Abrantes J, Lopes AM, Magalhães MJ, Esteves PJ. Evolution of CCL16 in Glires (Rodentia and Lagomorphs) shows an unusual random pseudogenization pattern. In preparation.

PSEUDOGENIZATION OF CCL14 IN THE OCHOTONIDAE (PIKA) FAMILY

Fabiana Neves, Joana Abrantes, Andrey A. Lissovsky, Pedro J Esteves

1. ABSTRACT

The interaction between chemokines and their receptors is crucial for inflammatory cells trafficking. CCL14 binds with high affinity to the CCR5. In leporids, CCR5 underwent gene conversion with CCR2. In some leporid species, CCL8 is pseudogenized while CCL3, CCL4 and CCL5 are all functional. Here, we study the evolution of CCL14 in mammals with emphasis in the order Lagomorpha. By employing Maximum Likelihood methods we detected 6 sites under positive selection. The Tajima´s Relative Rate Test showed that the *Ochotona princeps* sequence evolved faster than in leporids. Sequencing of 10 *Ochotona* species showed that *O. princeps*, *O. pallasi*, *O. alpina* and *O. turuchanensis* have a mutation at the start codon (Met>Thr) while *O. hoffmanni*, *O. mantchurica*, *O. dauurica* and *O. rufescens* present the mammalian conserved Met. *O. hyperborea* has the two alleles. In *O. pusilla,* CCL14 is a pseudogene due to a 7 base pair insertion. Our results suggest that, in mammals, CCL14 is evolving under positive selection. Like CCL3, CCL4 and CCL5, CCL14 is functional in all leporids but in the Ochotonidae family it underwent a pseudogenization process, suggesting that CCL14 has an important biological role in other mammals that apparently has been lost in the Ochotonidae family.

Keywords: Chemokine ligands, CCL14, Evolution, Ochotonidae, Pseudogenization

2. INTRODUCTION

Chemokines are proteins that have several functions in the immune system and are found only in vertebrates (Nomiyama et al., 2010). These proteins are encoded by a multigene family that emerged by gene duplication

events (Nomiyama et al., 2010). Gene duplication is considered a powerful force in the adaptive evolution of genes in which one gene gives rise to new copies (Ding et al., 2012; Levasseur and Pontarotti, 2011; Nomiyama et al., 2013). These new copies can evolve by pseudogenization, neofunctionalization or subfunctionalization (Kaessmann, 2010; Levasseur and Pontarotti, 2011; Zhang, 2003). Pseudogenization is the process in which a redundant functional copy passes through mutational decay and loses its ability to produce a functional protein (Clark, 1994; Zhang, 2003). In the remaining processes, both copies escape to mutation as either one copy emerges into new functions (neofunctionalization) or either divides the functions with the other copy (subfunctionalization) (Kaessmann, 2010; Levasseur and Pontarotti, 2011; Lynch and Force, 2000).

Chemokines can only exert their functions through the binding to specific receptors, having the ability to bind and activate different receptors; for example, CCL14 binds with high affinity to the chemokine receptors CCR1, CCR3 and CCR5 (Zlotnik and Yoshie, 2012). CCR5 is used as co-receptor by human immunodeficiency virus type-1 (HIV-1) for infection and, despite controversy, it has been also associated with myxoma virus infection in rabbits (Abrantes et al., 2008; Blain et al., 2007; Detheux et al., 2000; Lalani et al., 1999; Lederman and Sieg, 2007; Masters et al., 2001; Munch et al., 2002; van der Loo et al., 2012). In some mammalian species, CCR5 underwent gene conversion with CCR2 (Abrantes et al., 2011; Carmo et al., 2006; Esteves et al., 2007; Perelygin et al., 2008; Shields, 2000; Vazquez-Salat et al., 2007). Several studies have been conducted in the CCR5 ligands to determine the consequences of the CCR5-CCR2 gene conversion and its putative association with the myxoma virus infection. Indeed, the study of CCL8, a potential CCR5 ligand, revealed that in some leporid species, such as *Oryctolagus* and *Bunolagus*, this ligand is pseudogenized whilst in *Lepus* and *Sylvilagus* is intact (van der Loo et al., 2012). A study on CCL3, CCL4 and CCL5 showed that, unlike CCL8, these genes are all functional in leporids (de Matos et al., 2014). Although mouse CCL12 is the true ortholog to human CCL8 (Islam et al., 2011), in leporids, the gene identified as CCL8 is the true ortholog as supported by their phylogenetic position within the mammalian CCL8 cluster (van der Loo et al., 2012).

In leporids, other CCR5 ligands remain to be studied, including CCL14. CCL14, also known as Hemofiltrate CC-Chemokine 1 (HCC-1), is found in high concentrations in blood plasma and acts as an inflammatory chemokine after N-terminus cleavage (Blain et al., 2007; Richter et al., 2009; Schulz-Knappe et al., 1996). In humans, this protein is located in the Macrophage Inflammatory protein (MIP) region of C-C chemokine genes (Nomiyama et al., 2010; Shibata et al., 2013) and is expressed in numerous tissues (Blain et al., 2007; D'Ambrosio et al., 2003; Richter et al., 2009; Zlotnik and Yoshie, 2012). In humans, CCL14 encodes a protein with 93 amino acids (aa) with 3 predicted cleavage sites (Thr25, Glu31, and Ser33) that are necessary for the protein to become active (Tsou et al., 1998). There are two described variants: one encoded by 3 exons in which the active CCL14 consists in a 74 aa protein after the signal peptide cleavage, and another variant encoded by 4 exons, which has an insertion of 48 nucleotides between exons 1 and 2 encoding 16 aa (Blain et al., 2007; Detheux et al., 2000; Nomiyama et al., 2010; Richter et al., 2009). These forms have no effect against HIV-1 infection, however the cleaved form CCL14 (9-74) has the ability to block HIV-1 entry and replication (Munch et al., 2002). CCL14 (9-74) is an agonist of CCR1, CCR3 and CCR5 which promotes calcium influx and migration of T lymphocytes, monocytes and eosinophils; through the interaction with the second external loop of CCR5, it has also the ability to inhibit this receptor (Blain et al., 2007; Detheux et al., 2000; Munch et al., 2002; Richter et al., 2009). In C-C chemokines there are 2 main interactions important for receptor binding and activation, both located in the N-terminus close to the CC motif. In CCL14, the first interaction corresponds to the high affinity binding to the receptor and includes residues located after the CC motif, and the second interaction comprises the residues located before the CC motif and that are essential for the receptor activation (Blain et al., 2007; Lolis and Murphy, 2007; Pakianathan et al., 1997; Rajagopalan and Rajarathnam, 2006; Steen et al., 2014). The N-terminus is also important for CCL14 degradation (Savino et al., 2009). Degradation occurs following the binding of D6 to the N-terminus of CCL14 and then, through the interaction with the residue Pro53 of CCL14 (Nibbs et al., 1997; Savino et al., 2009).

The order Lagomorpha includes two families, Ochotonidae (pikas) and Leporidae (hares and rabbits) that diverged ~35 million years ago (Matthee et

al., 2004). The Ochotonidae family is composed of only one genus, *Ochotona*, which is divided into four subgenera, *Pika*, *Ochotona*, *Conothoa* and *Lagotona* (Lissovsky, 2014). The Leporidae family is divided in eleven genera (Matthee et al., 2004). Rodents and lagomorphs form the superorder Glires that is considered a sister group of the superorder Euarchonta that includes primates (Horner et al., 2007; Murphy et al., 2001). In rodents, CCL14 was shown to be absent in mouse and rat, while present in squirrel and guinea pig (Shibata et al., 2013); in lagomorphs, information is only available for the European rabbit (*Oryctolagus cuniculus*) where a functional gene exists.

By comparing the *O. princeps* CCL14 CDS available in GenBank/Ensembl with the remaining mammalian species (Figure 4.1) it was possible to observe that the initiation codon typically present in mammals is mutated into a Thr (ACG). Following this observation, and considering the apparent "random" absence of CCL14 in rodents and its role as a ligand for CCR5 which in some lagomorphs underwent gene conversion with CCR2, we aimed at determining the presence and function conservation of CCL14 in lagomorphs. In addition, we studied the evolution of CCL14 in mammals by identifying codons that are evolving under positive selection.

3. MATERIALS AND METHODS

Mammalian CCL14 sequences were retrieved from public databases (accession numbers are given in Figure 4.1). Sequences were aligned using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) available at http://www.ebi.ac.uk/ (Edgar, 2004). Similar results were obtained using Multiple Alignment using Fast Fourier Transform (MAFFT), however, MUSCLE alignment introduced less gaps, leading to a more conserved alignment.

In order to determine the impact of positive selection on CCL14 sequence evolution, we estimated ω, i.e, the ratio of nonsynonymous (dN) to synonymous (dS) differences in functional CCL14 orthologs (final dataset of 55 sequences). The codon-based Maximum Likelihood (ML) method (CODEML) implemented in PAML v4.4 was used (Yang, 2007). An unrooted neighbour-joining tree was constructed using MEGA6 (Tamura et al., 2013), with p-distance as substitution model and the pairwise deletion option for gaps/missing

data. The topology of the phylogenetic tree obtained is in accordance with the accepted mammalian phylogeny. Two pairs of site-based models were compared: M1 (nearly neutral) vs. M2 (selection) and M7 (neutral, beta) vs. M8 (selection, beta and ω), where M1 and M7 correspond to the null hypothesis and M2 and M8 to the alternative hypothesis by allowing positive selection. A likelihood ratio test (LRT) with 2 degrees of freedom determined if a selection model fitted better the data than a neutral model (Yang, 2002; Yang et al., 2000). Codons under positive selection were identified by using a Bayes Empirical Bayes (BEB) approach with probability > 95%. Codon-based ML methods available in the HYPHY package implemented in the DataMonkey webserver were also used (Delport et al., 2010; Poon et al., 2009): Single Likelihood Ancestor Counting (SLAC), Fixed-Effect Likelihood (FEL), Internal Branch Fel (iFEL), Mixed Effects Model of Evolution (MEME), Random Effect Likelihood (REL) and Fast Unconstrained Bayesian AppRoximation (FUBAR) (Murrell et al., 2013; Murrell et al., 2012; Poon et al., 2009). For the first 4 methods the p-value was set to ≤0.05; for FUBAR we used a p-value ≥0.95 and for REL we used a Bayes factor >95. The best fitting-model for nucleotide substitution was determined by the automatic model selection available in the webserver. As done previously for other immunity genes (Areal et al., 2011; Lemos de Matos et al., 2014; Neves et al., 2014b; Pinheiro et al., 2013; Pinheiro et al., 2014), only the codons detected by more than one method were considered as being under positive selection.

Samples from pikas (*Ochotona alpina, O. dauurica, O. hyperborea, O. hoffmanni, O. mantchurica, O. pallasi, O. princeps, O. pusilla, O. rufescens* and *O. turuchanensis*) were obtained from the Zoological Museum of Moscow University collection and samples of *O. cuniculus* (subspecies *O. c. cuniculus and O. c. algirus), Brachylagus idahoensis, Pentalagus furnessi, Sylvilagus bachmani*, *Sylvilagus floridanus*, *Lepus europaeus*, *Romerolagus diazi,* and *Ochotona princeps* were provided by the CIBIO Lagomorph tissue collection. For RNA extraction, a liver sample of *O. princeps* was kindly provided by Jay Storz from School of Biological Sciences, University of Nebraska, USA. Genomic DNA was extracted using the EasySpin Genomic DNA Minipreps Tissue Kit (Citomed) according to manufacturer's instructions. Total RNA was extracted for one specimen of *Oryctolagus cuniculus cuniculus*, *O. c. algirus*,

Lepus europaeus, *S. floridanus* and *O. princeps* by using RNeasy Mini Kit according to the manufacturer's instructions (Qiagen). cDNA was synthesized using oligo(dT) as primers and SuperScript III reverse transcriptase (Invitrogen). CCL14 was PCR-amplified and sequenced; sequences were submitted to GenBank under the following accession numbers: KP407602-KP407627 and KP641178. The European rabbit (*Oryctolagus cuniculus*) and Pika (*Ochotona princeps*) CCL14 sequences available in GenBank were used as templates for primer design. For amplification of CCL14 from gDNA, amplification was partitioned and amplified with Multiplex PCR Kit (Qiagen). Primers and conditions are listed in Table 4.1. Sequencing was performed on an ABI PRISM 310 Genetic Analyser (PE Applied Biosystems) and PCR products were sequenced in both directions.

Table 4. 1. Primers and conditions used in CCL14 PCR amplification.

Species amplified	Primers sequence (5'- 3')	Primer name	Exons amplified	PCR conditions	Fragment length
Leporids	CAGCTCTCCTGCAGCAGC	CCL14OrcuF1	Exon1	Initial denaturation - 95ºC (15min) 40 cycles (95ºC (45s), 58ºC (20s), 72ºC(10s) Final extension - 60ºC (10min)	~182bp
	GGAGTTGGATAGCTGAGAG	CCL14OrcuR1			
	CAGAAGCAGCCAGCAATC	CCL14OrcuF2	Exon 2 + exon 3	Initial denaturation - 95ºC (15min) 40 cycles (95ºC (45s), 58ºC (20s), 72ºC(45s) Final extension - 60ºC (10min)	~732bp
	CAAAGGCACAGCTCAGAG	CCL14OrcuR2			
Ochotona	CTCAGCACATCCTTACCCAG	CCL14OcprF1	Exon1	Initial denaturation - 95ºC (15min) 40 cycles (95ºC (45s), 60ºC (20s), 72ºC(10s) Final extension - 60ºC (10min)	~200bp
	CTTCTAGCCTCGTCTCCAG	CCL14OcprR1			
	CAGCAATCATGTCTCACCAG	CCL14OcprF2	Exon 2 + exon 3	Initial denaturation - 95ºC (15min) 40 cycles (95ºC (45s), 62ºC (20s), 72ºC(45s) Final extension - 60ºC (10min)	~612bp
	GAAGGCACAGTTCACAGACC	CCL14OcprR2			
	CCAGGGTAGCAGAAGCAGC	CCL14OcprF2a		Initial denaturation - 95ºC (15min) 40 cycles (95ºC (45s), 60ºC (20s), 72ºC(45s) Final extension - 60ºC (10min)	~692bp
	CTATGCCAGGAATTGCTCTG	CCL14OcprR2a			
	GTTCCTCTAACACCTCTCTGG	CCL14OcprF2b		Initial denaturation - 95ºC (15min) 40 cycles (95ºC (45s), 62ºC (20s), 72ºC(1min) Final extension - 60ºC (10min)	~801bp
	CTTGTGCACCTCACTGAG	CCL14OcprR2b			

The program PHASE, built into the software DnaSP (Librado and Rozas, 2009), was used to reconstruct the haplotype phases of the obtained sequences. These sequences were aligned and translated using BioEdit (Hall, 1999).

The Tajima's Relative Rate Test was conducted in order to check the evolutionary rate of CCL14 between lagomorphs using MEGA6 (Tajima, 1993; Tamura et al., 2013).

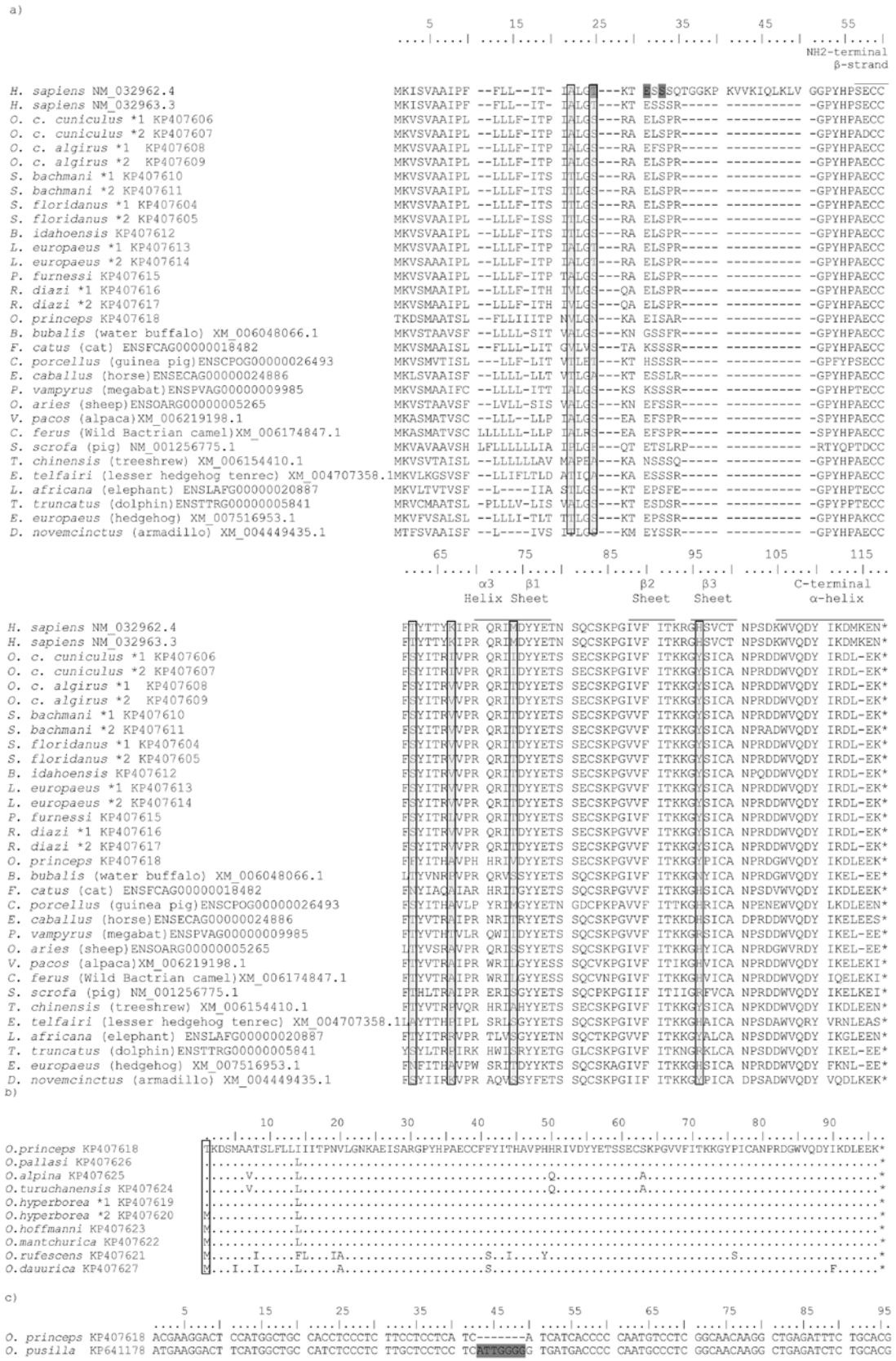


Figure 4. 1. a) Alignment of CCL14 for the different mammalian species (GenBank, Ensembl and Uniprot accession numbers are indicated for the retrieved sequences). Positively selected amino acids are boxed and the 3 possible cleavage sites are shaded in dark grey. (*) represent stop codons; (-) represent deletions; *1 and *2 represent alleles. b) Alignment of CCL14 for the different *Ochotona* species studied. The characteristic mammalian initiation codon is boxed. c) Alignment of the *CCL14* for *O. princeps* and *O. pusilla* with the insertion of the 7 base pairs shaded in grey.

In our study, 6 amino acids were identified as being under positive selection: Ala22, Thr25, Thr62, Lys67, Met74 and His96 (amino acid residues were numbered from the first methionine residue in human CCL14. The signal peptide and indels, indicated as (-), were included in the numbering; Figure 4.1). Ala22 and Thr25 are located in the pre peptide which is cleaved in order for the protein to become active (Detheux et al., 2000); however, the modifications observed at residue 22 are not expected to alter the CCL14 protein. In contrast, Thr25 is located within the region where CCL14 cleavage occurs. Since cleavage is required for protein activation, alterations in this site might negatively affect the CCL14 protein. Thr62 and Lys67 are in the region reported as to be essential to the binding of CCL14 to the several receptors (Blain et al., 2007; Lolis and Murphy, 2007; Pakianathan et al., 1997; Rajagopalan and Rajarathnam, 2006; Steen et al., 2014); thus, changes in this region may affect both the ligand-receptor binding, but also CCL14 degradation (Nibbs et al., 1997; Savino et al., 2009). Met74 and His96 are located in the β1 and β3 sheets, respectively. Alterations in these sites may lead to alterations in the protein conformation.

The comparison of the *O. princeps* CCL14 CDS available in GenBank and Ensembl with the remaining mammalian species (Figure 4.1a) showed that in *O. princeps* the initiation codon, in contrast to the other mammalian sequences, is mutated into a Thr (ACG). Following this observation, we attempted the amplification of CCL14 from gDNA of *O. princeps*. Our results indicated that, as for the sequences publicly available, the amplified fragment also presents the ACG mutation. Next, we attempted to amplify CCL14 from cDNA of *O. princeps*. Despite having used different Taq polymerases and different PCR conditions (e.g. including the addition of DMSO in the reaction, increasing extension times) and different pairs of primers designed according to the *O. princeps* CDS, and having previously successfully amplified other CDS of other genes from the same sample (Neves et al., 2014a), we were unable to amplify the CCL14. This may suggest that CCL14 is a pseudogene in *O. princeps*. We amplified and sequenced the CCL14 gene for 8 leporid species: *O. cuniculus cuniculus* and *O. c. algirus*, *Brachylagus idahoensis, Pentalagus*

furnessi, Sylvilagus bachmani and *S. floridanus*, *Lepus europaeus* and *Romerolagus diazi*.

To support the putative pseudogenization of CCL14 in *O. princeps* we performed a Tajima's Relative Rate Test in order to determine the rates of evolution of CCL14 in lagomorphs. This test determines if two genes (paralogs or orthologs) are evolving at similar rates, i.e. it determines statistically if the two sequences follow the molecular clock hypothesis of similar constant rates of evolution (null hypothesis) or if one sequence evolved at an accelerated or decelerated rate (alternative hypothesis) (Tajima, 1993). By comparing the *O. c. cuniculus* CCL14 gene with the *O. princeps* orthologs, and using as an outgroup each of the different leporids studied, the p-values of the chi-square were always statistically significant. This rejects the null hypothesis and favors the hypothesis of an acceleration of the mutation rate in *Ochotona princeps.* Thus, our results (Table 4.2) support a significant bias to neutrality between leporids and *O. princeps*, compatible with a pseudogenization process in this species.

Table 4. 2. Results obtained in Tajima's Relative Rate Test.

Outgroup	Nr. divergent sites in the three sequences	Nr. differences in *O.* c. *cuniculus*	Nr. differences in *O. princeps*	Chi-square [1]	p-value[2]
Sylvilagus bachmani	1	5	37	24.38	0.00000
Sylvilagus floridanus	2	7	34	17.78	0.00002
Brachylagus idahoensis	2	4	37	26.56	0.00000
Lepus europaeus	1	6	36	21.43	0.00000
Pentalagus furnessi	3	2	38	32.40	0.00000
Romerolagus diazi	1	9	33	13.71	0.00021

[1] Chi-square is a statistical test used to determine the substitution rates between species
[2] A p-value < 0.05 is used to reject the null hypothesis of equal rates between lineages.

We attempted to amplify the CCL14 gene from gDNA of several pika species that encompass the different *Ochotona* subgenera: *O. alpina, O. dauurica, O. hyperborea, O. hoffmanni, O. mantchurica, O. pallasi, O. pusilla,*

O. rufescens and *O. turuchanensis.* Interestingly, we found different results within the Ochotonidae family (Figure 4.1. b and c). Indeed, *O. alpina*, *O. pallasi* and *O. turuchanensis,* share the ATG>ACG mutation with *O. princeps*; in contrast, in *O. dauurica, O. hoffmanni, O. mantchurica* and *O. rufescens*, as in all the other mammals examined, the putative initiation codon is conserved. In addition, in *O. hyperborea*, we could define 2 alleles, one with the putative initiation codon and the other with ACG. In *O. pusilla*, despite presenting the putative start codon, CCL14 seems to be a pseudogene due to an insertion of 7 bp (base pairs 43-49) that leads to a frameshift that disrupts the CDS (Figure 4.1c). Our results seem to indicate that in the Ochotonidae family subgenus *Pika* CCL14 is under a process of pseudogenization with inactivation of the gene in some, but not all species; this pseudogenization, however, does not reflect the taxonomic relationships in the *Ochotona* genus (Figure 4.2). Indeed, the subgenus *Pika* where the mutation in the start codon emerged encompasses species from different lineages that present both the ATG and the ACG (Lissovsky, 2014; Melo-Ferreira et al., 2015). *O. pusilla* is the only species studied where CCL14 is obviously a pseudogene (Figure 4.1c), and belongs to a different subgenus (*Lagotona*) (Fostowicz-Frelik and Frelik, 2010; Fostowicz-Frelik et al., 2010; Lanier and Olson, 2009; Lissovsky, 2014; Melo-Ferreira et al., 2015).

Taking in account the CCL14 alterations described in the *Ochotona* species, we also determined if those alterations could have had any influence in the receptors. For this we sequenced the CCR5 and the CCR1 genes (data not shown). However, these sequences did not reveal alterations that could explain the pseudogenization of CCL14.

In the human genome, pseudogenes are quite common, with high prevalence in multigene families such as chemokines (Ondrej and Jianzhi, 2010), but little is known for other mammals. For CCL14, it was previously shown that in Rodentia, CCL14 is absent in rat and mouse of the Muridae family (Shibata et al., 2013), but is a functional gene in the families Dipodidae (*Jaculus jaculus*), Spalacidae (*Nannospalax galili*), Sciuridae (*Spermophilus tridecemlineatus*), Bathyergidae (*Heterocephalus glaber*), Chinchillidae (*Chinchilla laniger*) and Caviidae (*Cavia porcellus*). In addition, in the *Trichechus manatus latirostris* (Florida manatee) that belongs to the superorder

Afrotheria, the CCL14 predicted sequence available in GenBank (accession number XM_004385435.1) seems to be a pseudogene due to an early stop codon while for the other species within this order for which the CCL14 sequences are available (*Loxodonta africana* and *Echinops telfairi*), the gene seems to encode a functional protein. This indicates that although CCL14 is present as a functional gene in the common ancestor of these species, it was later inactivated or deleted in some species, even within the same clade. This suggests that CCL14 might have functions that overlap with the function of other genes making its presence unnecessary and leading to its pseudogenization.

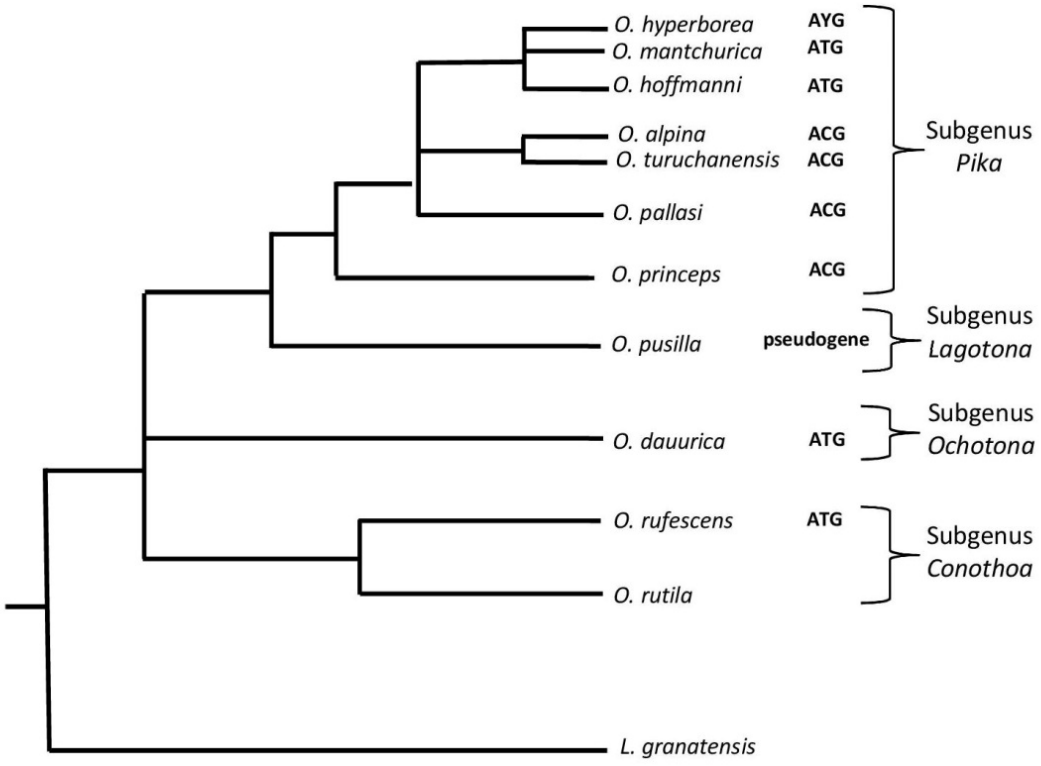


Figure 4. 2. Evolutionary topology showing the molecular phylogeny within the Ochotonidae family (adapted from Melo-Ferreira et al., 2015). In the AYG found in *O. hyperborea*, Y is a pyrimidine indicative that this species has 2 alleles, one with ATG and other with ACG.

5. CONCLUSIONS

In most of the mammals we observed that CCL14 encodes a functional gene that is evolving under positive selection, with positively selected codons located in regions important for ligand-receptor binding. In contrast, in two of the

Ochotonidae family subgenera (*Pika and Lagotona)* CCL14 is a pseudogene due to distinct disabling mutations and the detection of an acceleration of the mutation rates. While in some mammals CCL14 retained its biological significance in the Ochotonidae family (subgenera *Pika* and *Lagotona*) it seems to have been lost.

6. REFERENCES

Abrantes, J., Carmo, C.R., Matthee, C.A., Yamada, F., van der Loo, W., Esteves, P.J., 2011. **A shared unusual genetic change at the chemokine receptor type 5 between Oryctolagus, Bunolagus and Pentalagus**. Conserv Genet 12, 325-330.

Abrantes, J., Esteves, P.J., Carmo, C.R., Muller, A., Thompson, G., van der Loo, W., 2008. **Genetic characterization of the chemokine receptor CXCR4 gene in lagomorphs: comparison between the families Ochotonidae and Leporidae**. Int J Immunogenet 35, 111-117.

Areal, H., Abrantes, J., Esteves, P.J., 2011. **Signatures of positive selection in Toll-like receptor (TLR) genes in mammals**. BMC Evol Biol 11, 368.

Blain, K.Y., Kwiatkowski, W., Zhao, Q., La Fleur, D., Naik, C., Chun, T.W., Tsareva, T., Kanakaraj, P., Laird, M.W., Shah, R., George, L., Sanyal, I., Moore, P.A., Demeler, B., Choe, S., 2007. **Structural and functional characterization of CC chemokine CCL14**. Biochemistry 46, 10008-10015.

Carmo, C.R., Esteves, P.J., Ferrand, N., van der Loo, W., 2006. **Genetic variation at chemokine receptor CCR5 in leporids: alteration at the 2nd extracellular domain by gene conversion with CCR2 in Oryctolagus, but not in Sylvilagus and Lepus species**. Immunogenetics 58, 494-501.

Clark, A.G., 1994. **Invasion and maintenance of a gene duplication**. Proc Natl Acad Sci U S A 91, 2950-2954.

D'Ambrosio, D., Panina-Bordignon, P., Sinigaglia, F., 2003. **Chemokine receptors in inflammation: an overview**. J Immunol Methods 273, 3-13.

de Matos, A.L., Lanning, D.K., Esteves, P.J., 2014. **Genetic characterization of CCL3, CCL4 and CCL5 in leporid genera Oryctolagus, Sylvilagus and Lepus**. Int J Immunogenet 41, 154-158.

Delport, W., Poon, A.F., Frost, S.D., Kosakovsky Pond, S.L., 2010. **Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology**. Bioinformatics 26, 2455-2457.

Detheux, M., Standker, L., Vakili, J., Munch, J., Forssmann, U., Adermann, K., Pohlmann, S., Vassart, G., Kirchhoff, F., Parmentier, M., Forssmann, W.G., 2000. **Natural proteolytic processing of hemofiltrate CC chemokine 1 generates a potent CC chemokine receptor (CCR)1 and CCR5 agonist with anti-HIV properties**. J Exp Med 192, 1501-1508.

Ding, Y., Zhou, Q., Wang, W., 2012. **Origins of new genes and evolution of their novel functions**. Annu. Rev. Ecol. Evol. Syst. 43, 345-363.

Edgar, R.C., 2004. **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. Nucleic Acids Res 32, 1792-1797.

Esteves, P.J., Abrantes, J., van der Loo, W., 2007. **Extensive gene conversion between CCR2 and CCR5 in domestic cat (Felis catus)**. Int J Immunogenet 34, 321-324.

Fostowicz-Frelik, L., Frelik, G.J., 2010. **The earliest occurrence of the steppe pika (Ochotona pusilla) in Europe near the Pliocene/Pleistocene boundary**. Naturwissenschaften 97, 325-329.

Fostowicz-Frelik, L., Frelik, G.J., Gasparik, M., 2010. **Morphological phylogeny of pikas (Lagomorpha: Ochotona), with a description of a new species from the**

Pliocene/Pleistocene transition of Hungary. Proceedings of the Academy of Natural Sciences of Philadelphia 159, 97-118.

Hall, T.A., 1999. **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT**. Nucl. Acids. Symp. Ser. 41, 95-98.

Horner, D.S., Lefkimmiatis, K., Reyes, A., Gissi, C., Saccone, C., Pesole, G., 2007. **Phylogenetic analyses of complete mitochondrial genome sequences suggest a basal divergence of the enigmatic rodent Anomalurus**. BMC Evol Biol 7, 16.

Islam, S.A., Chang, D.S., Colvin, R.A., Byrne, M.H., McCully, M.L., Moser, B., Lira, S.A., Charo, I.F., Luster, A.D., 2011. **Mouse CCL8, a CCR8 agonist, promotes atopic dermatitis by recruiting IL-5+ T(H)2 cells**. Nat Immunol 12, 167-177.

Kaessmann, H., 2010. **Origins, evolution, and phenotypic impact of new genes**. Genome research 20, 1313-1326.

Lalani, A.S., Masters, J., Zeng, W., Barrett, J., Pannu, R., Everett, H., Arendt, C.W., McFadden, G., 1999. **Use of chemokine receptors by poxviruses**. Science 286, 1968-1971.

Lanier, H.C., Olson, L.E., 2009. **Inferring divergence times within pikas (Ochotona spp.) using mtDNA and relaxed molecular dating techniques**. Mol Phylogenet Evol 53, 1-12.

Lederman, M.M., Sieg, S.F., 2007. **CCR5 and its ligands: a new axis of evil?** Nat Immunol 8, 1283-1285.

Lemos de Matos, A., McFadden, G., Esteves, P.J., 2014. **Evolution of viral sensing RIG-I-like receptor genes in Leporidae genera Oryctolagus, Sylvilagus, and Lepus**. Immunogenetics 66, 43-52.

Levasseur, A., Pontarotti, P., 2011. **The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics**. Biology direct 6, 11.

Librado, P., Rozas, J., 2009. **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data**. Bioinformatics 25, 1451-1452.

Lissovsky, A.A., 2014. **Taxonomic revision of pikas Ochotona (Lagomorpha, Mammalia) at the species level**. Mammalia 78, 199-216.

Lolis, E., Murphy, J.W., 2007. **The Structural Biology of Chemokines**, in: Harrison, J.K., Lukacs, N.W. (Eds.), The Receptors: The Chemokine Receptors. Humana Press Inc, Totowa, New Jersey, pp. 9-30.

Lynch, M., Force, A., 2000. **The probability of duplicate gene preservation by subfunctionalization**. Genetics 154, 459-473.

Masters, J., Hinek, A.A., Uddin, S., Platanias, L.C., Zeng, W., McFadden, G., Fish, E.N., 2001. **Poxvirus infection rapidly activates tyrosine kinase signal transduction**. J Biol Chem 276, 48371-48375.

Matthee, C.A., van Vuuren, B.J., Bell, D., Robinson, T.J., 2004. **A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene**. Syst Biol 53, 433-447.

Melo-Ferreira, J., Lemos de Matos, A., Areal, H., Lissovsky, A.A., Carneiro, M., Esteves, P.J., 2015. **The phylogeny of pikas (Ochotona) inferred from a multilocus coalescent approach**. Mol Phylogenet Evol 84, 240-244.

Munch, J., Standker, L., Pohlmann, S., Baribaud, F., Papkalla, A., Rosorius, O., Stauber, R., Sass, G., Heveker, N., Adermann, K., Escher, S., Kluver, E., Doms, R.W., Forssmann, W.G., Kirchhoff, F., 2002. **Hemofiltrate CC chemokine 1[9-74] causes effective internalization of CCR5 and is a potent inhibitor of R5-tropic human immunodeficiency virus type 1 strains in primary T cells and macrophages**. Antimicrobial agents and chemotherapy 46, 982-990.

Murphy, W.J., Eizirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A., O'Brien, S.J., 2001. **Molecular phylogenetics and the origins of placental mammals**. Nature 409, 614-618.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., Scheffler, K., 2013. **FUBAR: a fast, unconstrained bayesian approximation for inferring selection**. Mol Biol Evol 30, 1196-1205.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Kosakovsky Pond, S.L., 2012. **Detecting individual sites subject to episodic diversifying selection**. PLoS genetics 8, e1002764.

Neves, F., Abrantes, J., Pinheiro, A., Almeida, T., Costa, P.P., Esteves, P.J., 2014a. **Convergent evolution of IL-6 in two leporids (Oryctolagus and Pentalagus) originated an extended protein**. Immunogenetics 66, 589-595.

Neves, F., Abrantes, J., Steinke, J.W., Esteves, P.J., 2014b. **Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals**. Innate immunity 20, 184-191.

Nibbs, R.J., Wylie, S.M., Pragnell, I.B., Graham, G.J., 1997. **Cloning and characterization of a novel murine beta chemokine receptor, D6. Comparison to three other related macrophage inflammatory protein-1alpha receptors, CCR-1, CCR-3, and CCR-5**. J Biol Chem 272, 12495-12504.

Nomiyama, H., Osada, N., Yoshie, O., 2010. **The evolution of mammalian chemokine genes**. Cytokine Growth Factor Rev 21, 253-262.

Nomiyama, H., Osada, N., Yoshie, O., 2013. **Systematic classification of vertebrate chemokines based on conserved synteny and evolutionary history**. Genes to cells : devoted to molecular & cellular mechanisms 18, 1-16.

Ondrej, P., Jianzhi, Z., 2010. **Pseudogenes and their evolution**, ENCYCLOPEDIA OF LIFE SCIENCES. John Wiley & Sons Ltd, Chichester.

Pakianathan, D.R., Kuta, E.G., Artis, D.R., Skelton, N.J., Hebert, C.A., 1997. **Distinct but overlapping epitopes for the interaction of a CC-chemokine with CCR1, CCR3 and CCR5**. Biochemistry 36, 9642-9648.

Perelygin, A.A., Zharkikh, A.A., Astakhova, N.M., Lear, T.L., Brinton, M.A., 2008. **Concerted evolution of vertebrate CCR2 and CCR5 genes and the origin of a recombinant equine CCR5/2 gene**. J Hered 99, 500-511.

Pinheiro, A., Woof, J.M., Abi-Rached, L., Parham, P., Esteves, P.J., 2013. **Computational analyses of an evolutionary arms race between mammalian immunity mediated by immunoglobulin A and its subversion by bacterial pathogens**. PLoS One 8, e73934.

Pinheiro, A., Woof, J.M., Almeida, T., Abrantes, J., Alves, P.C., Gortazar, C., Esteves, P.J., 2014. **Leporid immunoglobulin G shows evidence of strong selective pressure on the hinge and CH3 domains**. Open Biol 4, 140088.

Poon, A.F., Frost, S.D., Pond, S.L., 2009. **Detecting signatures of selection from DNA sequences using Datamonkey**. Methods Mol Biol 537, 163-183.

Rajagopalan, L., Rajarathnam, K., 2006. **Structural basis of chemokine receptor function--a model for binding affinity and ligand selectivity**. Bioscience reports 26, 325-339.

Richter, R., Casarosa, P., Standker, L., Munch, J., Springael, J.Y., Nijmeijer, S., Forssmann, W.G., Vischer, H.F., Vakili, J., Detheux, M., Parmentier, M., Leurs, R., Smit, M.J., 2009. **Significance of N-terminal proteolysis of CCL14a to activity on the chemokine receptors CCR1 and CCR5 and the human cytomegalovirus-encoded chemokine receptor US28**. J Immunol 183, 1229-1237.

Savino, B., Borroni, E.M., Torres, N.M., Proost, P., Struyf, S., Mortier, A., Mantovani, A., Locati, M., Bonecchi, R., 2009. **Recognition versus adaptive up-regulation and degradation of CC chemokines by the chemokine decoy receptor D6 are determined by their N-terminal sequence**. J Biol Chem 284, 26207-26215.

Schulz-Knappe, P., Magert, H.J., Dewald, B., Meyer, M., Cetin, Y., Kubbies, M., Tomeczkowski, J., Kirchhoff, K., Raida, M., Adermann, K., Kist, A., Reinecke, M., Sillard, R., Pardigol, A., Uguccioni, M., Baggiolini, M., Forssmann, W.G., 1996. **HCC-1, a novel chemokine from human plasma**. J Exp Med 183, 295-299.

Shibata, K., Nomiyama, H., Yoshie, O., Tanase, S., 2013. **Genome diversification mechanism of rodent and Lagomorpha chemokine genes**. Biomed Res Int 2013, 856265.

Shields, D.C., 2000. **Gene conversion among chemokine receptors**. Gene 246, 239-245.

Steen, A., Larsen, O., Thiele, S., Rosenkilde, M.M., 2014. **Biased and g protein-independent signaling of chemokine receptors**. Front Immunol 5, 277.

Tajima, F., 1993. **Simple methods for testing the molecular evolutionary clock hypothesis**. Genetics 135, 599-607.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. **MEGA6: Molecular Evolutionary Genetics Analysis version 6.0**. Mol Biol Evol 30, 2725-2729.

Tsou, C.L., Gladue, R.P., Carroll, L.A., Paradis, T., Boyd, J.G., Nelson, R.T., Neote, K., Charo, I.F., 1998. **Identification of C-C chemokine receptor 1 (CCR1) as the monocyte hemofiltrate C-C chemokine (HCC)-1 receptor**. J Exp Med 188, 603-608.

van der Loo, W., Afonso, S., de Matos, A.L., Abrantes, J., Esteves, P.J., 2012. **Pseudogenization of the MCP-2/CCL8 chemokine gene in European rabbit (genus Oryctolagus), but not in species of Cottontail rabbit (Sylvilagus) and Hare (Lepus)**. BMC Genet 13, 72.

Vazquez-Salat, N., Yuhki, N., Beck, T., O'Brien, S.J., Murphy, W.J., 2007. **Gene conversion between mammalian CCR2 and CCR5 chemokine receptor genes: a potential mechanism for receptor dimerization**. Genomics 90, 213-224.

Yang, Z., 2002. **Inference of selection from multiple species alignments**. Current opinion in genetics & development 12, 688-694.

Yang, Z., 2007. **PAML 4: phylogenetic analysis by maximum likelihood**. Mol Biol Evol 24, 1586-1591.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. **Codon-substitution models for heterogeneous selection pressure at amino acid sites**. Genetics 155, 431-449.

Zhang, J., 2003. **Evolution by gene duplication: an update**. TRENDS in Ecology and Evolution 18, 292-298.

Zlotnik, A., Yoshie, O., 2012. **The chemokine superfamily revisited**. Immunity 36, 705-716.

EVOLUTION OF CCL11: GENETIC CHARACTERIZATION IN LAGOMORPHS AND EVIDENCE OF POSITIVE AND PURIFYING SELECTION IN MAMMALS

Fabiana Neves, Joana Abrantes, Pedro J Esteves

1. ABSTRACT

The interactions between chemokines and their receptors are crucial for differentiation and activation of inflammatory cells. CCL11 binds to CCR3 and to CCR5 that in leporids underwent gene conversion with CCR2. Here, we genetically characterized CCL11 in lagomorphs (leporids and pikas). All lagomorphs have a potentially functional CCL11, and the pygmy rabbit has a mutation in the stop codon that leads to a longer protein. Other mammals also have mutations at the stop codon that result in proteins with different lengths. By employing Maximum Likelihood methods, we observed that in mammals CCL11 exhibits both signatures of purifying and positive selection. Signatures of purifying selection were detected in sites important for receptor binding and activation. Of the three sites detected as under positive selection, two were located close to the stop codon. Our results suggest that CCL11 is functional in all lagomorphs, and that the signatures of purifying and positive selection in mammalian CCL11 probably reflect the protein's biological roles.

Keywords: CCL11, purifying selection, positive selection, lagomorphs, mammals

2. INTRODUCTION

C-C chemokine ligand 11 (CCL11), also known as Eotaxin-1, is a chemoattractant for eosinophils with an important role in allergic and parasitic inflammation (Pease and Williams, 2001; Puxeddu et al., 2006; Stevenson et al., 2009). First isolated in a guinea pig model of asthma (Jose et al., 1994), this protein is located in the monocyte chemoattractant protein (MCP) region of the

CC cluster of several mammals, including human, mouse, rat, rabbit, horse and cow (Nomiyama et al., 2010; Shibata et al., 2013). In the European rabbit *(Oryctolagus cuniculus*) genome, CCL11 has been mapped to chromosome 19:23,739,202-23,742,061 on the forward strand with an Ensembl gene designation ENSOCU00000005935. CCL11 can exert its functions through the interaction between residues located in its extracellular loops and NH_2-terminus and the N-terminus of two receptors, CCR3 and CCR5 (Allen et al., 2007; D'Ambrosio et al., 2003; Millard et al., 2014).

An extensive gene conversion in the CCR5 transmembrane domain has been reported for several species (Abrantes et al., 2011; Carmo et al., 2006; Esteves et al., 2007; Perelygin et al., 2008; Shields, 2000; Vazquez-Salat et al., 2007). In contrast, in some leporids, CCR5 suffered a dramatic change at the second extracellular loop, resulting from a gene conversion event with the paralogous CCR2. This alteration was confirmed in the European rabbit, riverine rabbit (*Bunolagus monticularis*) and Amami rabbit (*Pentalagus furnessi*), but it was not observed in the Eastern cottontail (*Sylvilagus floridanus*) or in European and Iberian hares (*Lepus europaeus* and *L. granatensis*) (Abrantes et al., 2011; Carmo et al., 2006). In the other Lagomorpha family, Ochotonidae (pikas), this gene conversion is also absent (Pinheiro et al., 2016). The most likely evolutionary scenario that could explain this pattern is that the gene conversion event occurred in the ancestor of *Oryctolagus*, *Bunolagus* and *Pentalagus* genera at ~8 million years ago, probably conferring some selective advantage, and thus became fixed in the ancestral population (Pinheiro et al., 2016).

This CCR5 evolutionary pattern led to the study of the CCR5 ligands in lagomorphs. The study of CCL3, CCL4 and CCL5 revealed that these genes are all functional in leporids and showed evidence of strong purifying selection (de Matos et al., 2014). CCL14 is functional in the Leporidae family. In the Ochotonidae family, CCL14 is suffering a pseudogenization process with an intact gene in some species while in others it is a pseudogene. In contrast, among the leporids, CCL8 is pseudogenized in the European rabbit and riverine rabbit, but functional in *Sylvilagus* and *Lepus* (Neves et al., 2015; van der Loo et al., 2012).

In chemokine receptors the sites located intracellularly or in the transmembrane domains are involved in signal transduction and dimerization (Metzger and Thomas, 2010). Consequently, new nucleotide changes that lead to amino acid (aa) alterations tend to be quickly eliminated. In contrast, the aa residues localized in the extracellular domains are evolving under selective pressure resulting from ligand-binding and pathogen interactions. Thus, in these regions, the proportion of non-synonymous nucleotide substitutions is expected to be significantly higher than synonymous substitutions. This pattern has been observed in chemokine receptors such as CCR2 and CCR3 (Metzger and Thomas, 2010). The interactions between chemokine ligands and their receptors suggest that the chemokine ligands might also exhibit signatures of positive and purifying selection. Here, we characterized CCL11 in lagomorphs and investigated the selective pressures that have been driving the evolution of this molecule in mammals.

3. MATERIALS AND METHODS

Tissue samples were provided by the CIBIO Lagomorpha tissue collection with the exception of brush rabbit (*Sylvilagus bachmani*) tissue samples which were kindly provided by Jeff Wilcox and Dr. Michael Hamilton from Blue Oak Ranch Reserve, University of California, Berkeley. Genomic DNA (gDNA) was extracted from three tissue samples of pygmy rabbit (*Brachylagus idahoensis)* and from one sample from each European rabbit subspecies *(Oryctolagus cuniculus cuniculus and Oryctolagus cuniculus algirus),* and one sample of European brown hare (*Lepus europaeus)*, pygmy rabbit (*Brachylagus idahoensis),*brush rabbit (*Sylvilagus bachmani*), volcano rabbit (*Romerolagus diazi*) and American pika (*Ochotona princeps*) using the EasySpin Genomic DNA Minipreps Tissue Kit (Citomed, Torun, Poland) according to manufacturer's instructions. Total RNA was extracted using the RNeasy Mini Kit according to the manufacturer's instructions (Qiagen, Hilden, Germany) from samples of European rabbit (subspecies *cuniculus)* and European brown hare*.* cDNA was synthesized using oligo(dT) as primers and SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA, USA). The European rabbit and American pika sequences available in public databases

were used for primer design. PCR amplification was performed with the Multiplex PCR Kit (Qiagen) according to the manufacturer's protocol (Table 4.3). Sequencing was performed on an ABI PRISM 310 Genetic Analyser (PE Applied Biosystems) and PCR products were sequenced in both directions. Sequences were submitted to GenBank under the following accession numbers: KU987806-KU987816.

Table 4. 3. Primers and conditions used for PCR amplification and sequencing of CCL11 from lagomorphs' gDNA samples.

Species amplified	Primers sequence (5'- 3')	Primer name	Exons amplified	PCR conditions	Fragment length
Leporids	GATCAATCCAGAAGCCTCC	CCL11OrcuF1*	Exon1	95ºC (15min) 40 cycles: 95ºC (45s), 58ºC (20s), 72ºC (20s) 60ºC (10min)	~185bp
	CAGGTCAGCAAGTGTCTTC	CCL11OrcuR1			
European rabbit Brush rabbit	CTACCTCTGACACATCCTC	CCL11OrcuF2	Exon 2 + Exon 3	95ºC (15min) 40 cycles: 95ºC (45s), 58ºC (20s), 72ºC (50s) 60ºC (10min)	~805bp
	CCTTCTCAAGATGCGTTCTG	CCL11OrcuR2*			
European brown hare Pygmy rabbit Volcano rabbit	CTGTGTATCATTAGTAGCTC	CCL11OrcuF2a		95ºC (15min) 40 cycles: 95ºC (45s), 56ºC (20s), 72ºC (45s) 60ºC (10min)	~778bp
	GTGAGCTAAGCCTCACCTG	CCL11OrcuR2a			
Ochotona princeps	GCAGGCAGAACAATCAAATAC	CCL11OcprF1a*	Exon1 + Exon 2 + Exon 3	95ºC (15min) 40 cycles: 95ºC (45s), 64ºC (20s), 72ºC (1min40s) 60ºC (10min)	~1637bp
	GTTATTGCTGGTCGCTCAG	CCL11OcprR1a*			
	CCTCTTGATGAGCATGTTC	CCL11OcprFint1			

Sequences were aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE) available at http://www.ebi.ac.uk/ (Edgar, 2004). The program PHASE, built into the software DnaSP (Librado and Rozas, 2009), was used to reconstruct the haplotype phases of the obtained sequences. Haplotypes were translated using BioEdit (Hall, 1999).

In order to identify which codons of CCL11 are under selection (purifying and positive), we estimated ω, i.e, the ratio of nonsynonymous (dN) to synonymous (dS) substitutions in CCL11 orthologs (final dataset of 68 sequences) by employing codon-based Maximum Likelihood (ML) methods available in the HYPHY package implemented in the DataMonkey webserver (Delport et al., 2010; Poon et al., 2009): Single Likelihood Ancestor Counting (SLAC), Fixed-Effect Likelihood (FEL), Internal Branch Fel (iFEL), Random-Effect Likelihood (REL) and Fast Unconstrained Bayesian AppRoximation (FUBAR) (Murrell et al., 2013; Murrell et al., 2012; Poon et al., 2009). For the first 3 methods the p-value was set to ≤0.05; for FUBAR we used a p-value

≥0.95 and for REL we used a Bayes factor >95. The best fitting-model for nucleotide substitution was determined by the automatic model selection tool available in the webserver. We further used the codon-based ML method (CODEML) implemented in PAML v4.4 (Yang, 2007). An unrooted neighbor-joining tree was constructed using MEGA6 (Tamura et al., 2013), with p-distance as substitution model and the pairwise deletion option for gaps/missing data. The topology of the phylogenetic tree obtained follows the accepted mammalian phylogeny for the major groups. Two pairs of site-based models were compared: M1 (nearly neutral) vs. M2 (selection) and M7 (neutral, β) vs. M8 (selection, β and ω), where M1 and M7 correspond to the null hypothesis and M2 and M8 to the alternative hypothesis by allowing positive selection. A likelihood ratio test (LRT) with 2 degrees of freedom determined whether a selection model fit the data better than a neutral model (Yang, 2002; Yang et al., 2000). Codons under positive selection were identified by using a Bayes Empirical Bayes (BEB) approach with probability > 95%. As done previously for other immunity genes (Areal et al., 2011; Lemos de Matos et al., 2013, 2014; Neves et al., 2014; Pinheiro et al., 2013; Pinheiro et al., 2014), only the codons detected by more than one method were considered as being under selection.

The three-dimensional (3D) structure displaying the interaction of Human CCL11 with CCR3 was downloaded from the Protein Data Bank in Europe (PDBeurope) available in http://www.ebi.ac.uk/pdbe/entry/pdb/2MPM, and Discovery Studio 3.5 software (BIOVIA, San Diego, CA, USA) was used to map the sites under selection.

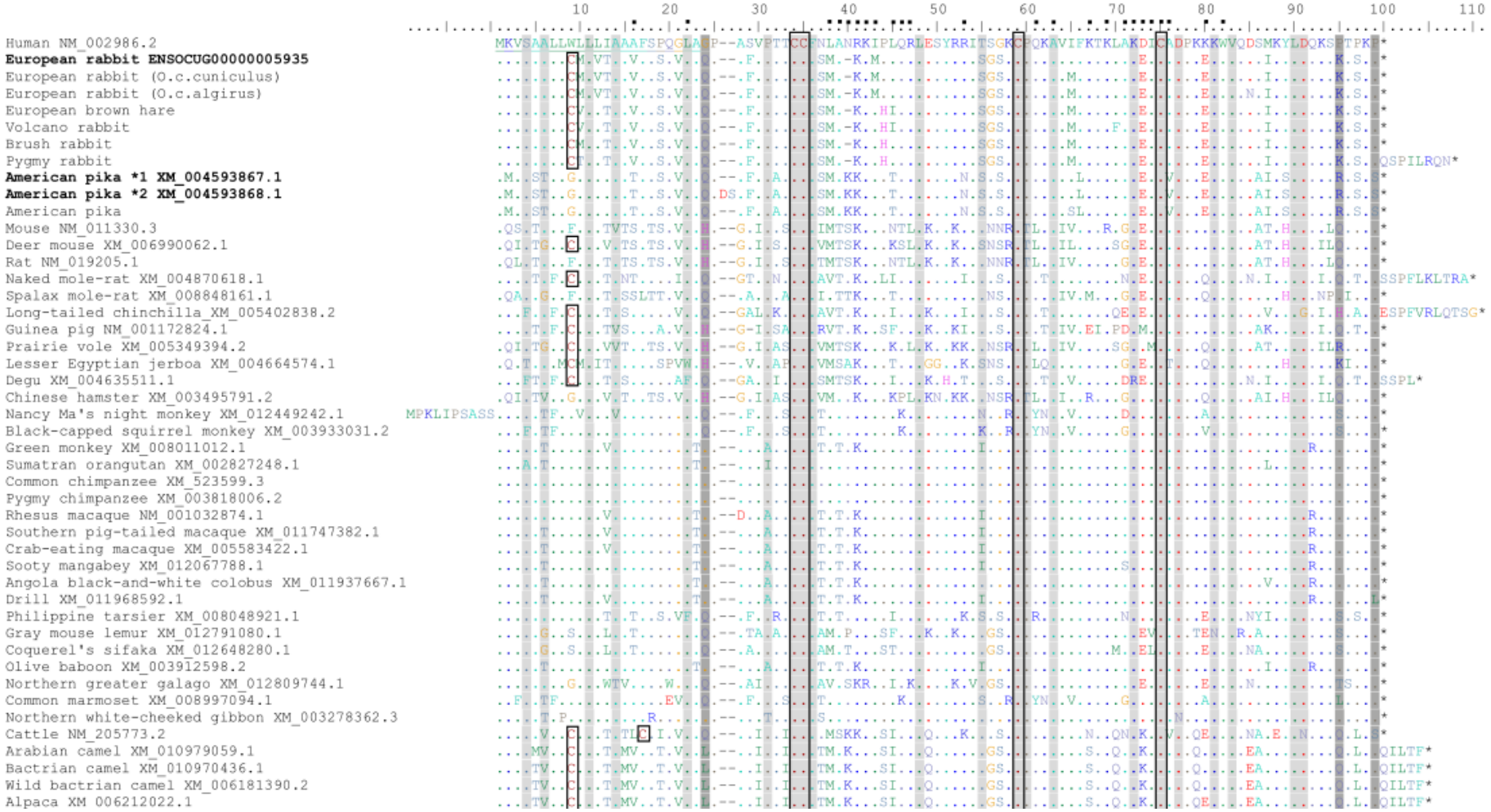
The secondary structures of human, European rabbit and pygmy rabbit CCL11 were predicted by using PsiPred (http://bioinf.cs.ucl.ac.uk/psipred/) . This software calculates the protein cysteines that create disulfide bonds by using Position Specific Iterated – BLAST (PSI-BLAST) to obtain evolutionary information that is used to predict the secondary structure of the query protein.

4. RESULTS AND DISCUSSION

We amplified and sequenced the CCL11 gene for six lagomorph species: both subspecies of European rabbit, European brown hare, brush rabbit, pygmy rabbit, volcano rabbit and American pika. These sequences were further

compared with CCL11 sequences available for other mammals and some differences were observed (Figure 4.3). Indeed, at position 40, where human has an Asn, all leporids have a deletion and the American pika has a Lys. There are also some amino acid changes that are only present in some lagomorph species: Val12 and Met43 in the European rabbit; Thr10 (pygmy rabbit); His44 (volcano, pygmy and brush rabbits and European brown hare); Phe70 (volcano rabbit) and Met2, Ser5, Asn53, Leu66, Ser89 (American pika). In addition, we found a Met65 in all CCL11 sequences from the leporids amplified in this work, while the sequence from the European rabbit available in public databases (ENSOCUG00000005935) has an isoleucine. Additionally, and despite the identification of two different transcripts in the American pika sequence available online (XM_004593867.1 and XM_004593868.1 – deletion/insertion of two aa Asp26 and Ser27, respectively), we were only able to detect the first transcript. The listed amino acid alterations lead to changes in charge and polarity (Supplementary material table 4.5) which can induce modifications in protein structure and conformation and even alter the protein functions (Majewski and Ott, 2003).

Strikingly, within leporids, the pygmy rabbit has a mutation in the stop codon leading to an extension of 8 amino acids (Gln100-Asn107) resulting in a protein with 104 amino acids. This mutation was detected in three different individuals confirming that this was not a PCR artifact. Other mammals also code for longer proteins due to mutations in the stop codon (Figure 4.3), leading to CCL11 proteins ranging from 100 to 108 amino acids. The protein length is important for the three-dimensional structure and different sizes may imply different folding patterns which may affect protein functions (Tiessen et al., 2012). The implications of these mutations in the protein structure are unknown, however PsiPred results do not predict any differences in the secondary structure of these species when compared with human.


10
20
30
40
50
60
70
80
90
100
110
Human NM_002986.2
European rabbit ENSOCUG00000005935
European rabbit (O.c.cuniculus)
European rabbit (O.c.algirus)
European brown hare
Volcano rabbit
Brush rabbit
Pygmy rabbit
American pika *1 XM_004593867.1
American pika *2 XM_004593868.1
American pika
Mouse NM_011330.3
Deer mouse XM_006990062.1
Rat NM_019205.1
Naked mole-rat XM_004870618.1
Spalax mole-rat XM_008848161.1
Long-tailed chinchilla_XM_005402838.2
Guinea pig NM_001172824.1
Prairie vole XM_005349394.2
Lesser Egyptian jerboa XM_004664574.1
Degu XM_004635511.1
Chinese hamster XM_003495791.2
Nancy Ma's night monkey XM_012449242.1
Black-capped squirrel monkey XM_003933031.2
Green monkey XM_008011012.1
Sumatran orangutan XM_002827248.1
Common chimpanzee XM_523599.3
Pygmy chimpanzee XM_003818006.2
Rhesus macaque NM_001032874.1
Southern pig-tailed macaque XM_011747382.1
Crab-eating macaque XM_005583422.1
Sooty mangabey XM_012067788.1
Angola black-and-white colobus XM_011937667.1
Drill XM_011968592.1
Philippine tarsier XM_008048921.1
Gray mouse lemur XM_012791080.1
Coquerel's sifaka XM_012648280.1
Olive baboon XM_003912598.2
Northern greater galago XM_012809744.1
Common marmoset XM_008997094.1
Northern white-cheeked gibbon XM_003278362.3
Cattle NM_205773.2
Arabian camel XM_010979059.1
Bactrian camel XM_010970436.1
Wild bactrian camel XM_006181390.2
Alpaca XM_006212022.1

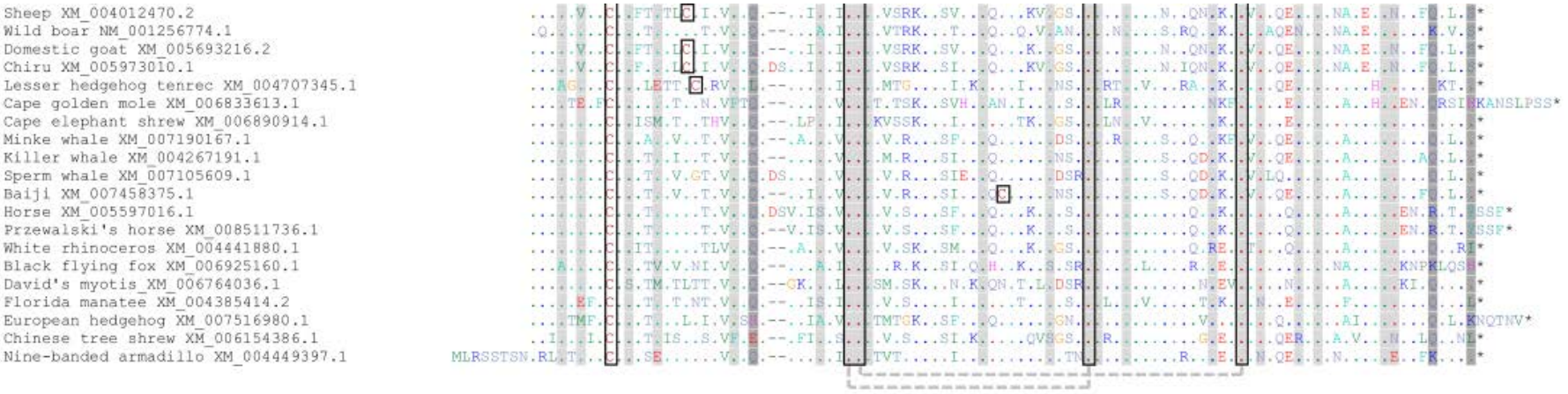


Figure 4. 3. Alignment of CCL11 for several mammalian species. GenBank and Ensembl accession numbers are indicated in bold for the European rabbit and American pika retrieved sequences. Purifying selected amino acids are shaded in light grey while positively selected amino acids are shaded in dark grey; cysteine residues are boxed and human signal peptide in underlined. (*) represent stop codons; (-) represent indels; (▪) above the numbering represent the sites important for the interaction of CCL11 with its receptors (Crump et al., 1998; Crump et al., 1999; Millard et al., 2014; Ye et al., 2000; Zhang and LiWang, 2014) *1 and *2 represent different alleles. Numbering is according to the Human CCL11 sequence and the signal peptide and indels were included in the numbering. Disulfide bonds between side chain cysteines are represented by a light grey dashed line. Human (*Homo sapiens*), European rabbit (*Oryctolagus cuniculus cuniculus and Oryctolagus cuniculus algirus),* European brown hare (*Lepus europaeus)*, Pygmy rabbit (*Brachylagus idahoensis),* Brush rabbit (*Sylvilagus bachmani*), Volcano rabbit (*Romerolagus diazi*) and American pika (*Ochotona princeps*); Mouse (*Mus musculus*), Deer mouse (*Peromyscus maniculatus bairdii*); Rat (*Rattus norvegicus*); Naked mole-rat (*Heterocephalus glaber*); Spalax mole-rat (*Nannospalax galili*); Long-tailed chinchilla (*Chinchilla lanigera*); Guinea pig (*Cavia porcellus*); Prairie vole (*Microtus ochrogaster*); Lesser Egyptian jerboa (*Jaculus jaculus*); Degu (*Octodon degus*); Chinese hamster (*Cricetulus griseus*); Nancy Ma's night monkey (*Aotus nancymaae*); Black-capped squirrel monkey (*Saimiri boliviensis boliviensis*); Green monkey (*Chlorocebus sabaeus*); Sumatran orangutan (*Pongo abelii*); Common chimpanzee (*Pan troglodytes*); Pygmy chimpanzee (*Pan paniscus*); Rhesus macaque (*Macaca mulatta*); Southern pig-tailed macaque (*Macaca nemestrina*); Crab-eating macaque (*Macaca fascicularis*); Sooty mangabey (*Cercocebus atys*); Angola black-and-white colobus (*Colobus angolensis palliates*); Drill (*Mandrillus leucophaeus*); Philippine tarsier (*Tarsius syrichta*); Gray mouse lemur (*Microcebus murinus*); Coquerel's sifaka (*Propithecus coquereli*); Olive baboon (*Papio anubis*); Northern greater galago (*Otolemur garnettii*); Common marmoset (*Callithrix jacchus*); Northern white-cheeked gibbon (*Nomascus leucogenys*); Cattle (*Bos taurus*); Arabian camel (*Camelus dromedarius*); Bactrian camel (*Camelus bactrianus*); Wild bactrian camel (*Camelus ferus*); Alpaca (*Vicugna pacos*); Sheep (*Ovis aries*); Wild boar (*Sus scrofa*); Domestic goat (*Capra hircus*); Chiru (*Pantholops hodgsonii*); Lesser hedgehog tenrec (*Echinops telfairi*); Cape golden mole (*Chrysochloris asiatica*); Cape elephant shrew (*Elephantulus edwardii*); Minke whale (*Balaenoptera acutorostrata scammoni*); Killer whale (*Orcinus orca*); Sperm whale (*Physeter catodon*); Baiji (*Lipotes vexillifer*); Horse (*Equus caballus*); Przewalski's horse (*Equus przewalskii*); White rhinoceros (*Ceratotherium simum simum*); Black flying fox (*Pteropus alecto*); David's myotis (*Myotis davidii*); Florida manatee (*Trichechus manatus latirostris*); European hedgehog (*Erinaceus europaeus*); Chinese tree shrew (*Tupaia chinensis*); Nine-banded armadillo (*Dasypus novemcinctus*).

Table 4. 4 .Phylogenetic Tests of Selection[a]

CCL11	No. aa residues	No. of species	lnL M7	lnL M8	2ΔlnL	PAML	SLAC[b]	FEL[c]	REL[d]	iFel[e]	MEME[f]	FUBAR[g]	Total no. of sites	% of sites
Purifying selection	**327**	61	-	-	-	-	4, 6, 22, 31, 35, 36, 48, 55, 59, 60, 63, 77, 83, 91	4, 6, 11, 14, 21, 22, 31, 35, 36, 48, 55, 59, 60, 63, 72, 77, 83, 90, 91, 94	1, 6, 17, 23, 27, 34, 35, 36, 42, 43, 48, 51, 59, 60, 63, 64, 66, 68, 72, 75, 76, 77, 78, 81, 82, 83, 85, 88, 90, 91	4, 8, 11, 14, 22, 35, 36, 48, 50, 55, 59, 60, 63, 69, 72, 74, 77, 81, 83, 91, 96	-	4, 6, 14, 22, 34, 35, 36, 48, 55, 59, 60, 63, 75, 77, 83, 91	21	6,4
Positive selection	**327**	61	-5406,996	-5397,951	18,090*	95	24, 95	24, 95	33, 38, 40, 95, 99	24, 56, 95	13, 20, 24, 64, 95, 99	95	3	0,92

Disulfide bonds between cysteine residues are important for protein structure and function, being involved in an array of biological processes (Fass, 2012; Helenius and Aebi, 2004; Rudd et al., 2001). For CCL11, the cysteine residues and the predicted disulfide bonds are highly conserved between all the mammals studied (Cys34-Cys59 and Cys35-Cys75). However, some extra cysteines were detected: cattle, sheep, domestic goat and chiru have Cys17 and the lesser hedgehog tenrec has Cys18. In addition, most mammals, with the exception of primates, American pika, mouse, rat, spalax mole-rat and Chinese hamster, have an extra cysteine at position 9. Their location in the signal peptide suggests no role in establishing extra disulfide bonds. In contrast, the Baiji sequence has an extra cysteine at position 50 that can potentially establish a disulfide bond.

Chemokines exert important roles in the immune response. In order to maintain their conformation and biological role, chemokines are expected to have signatures of purifying selection; however, if we consider that these proteins are targets for several types of molecules (such as pathogens, drugs, receptors) it should be expected that chemokines also have signatures of positive selection (Metzger and Thomas, 2010). The results obtained indicate that mammalian CCL11 has 21 codons negatively selected and 3 codons positively selected (Table 4.4). CCL11 interacts with both CCR3 and CCR5 receptors, with CCR3 being the higher affinity receptor (Manns et al., 2007). The most crucial region for receptor binding and activation is the N-terminus preceding the disulfide bonds. The first 20 amino acids (following the signal peptide) and the N-loop located after the disulfide bonds are also described to be important for receptor binding (Crump et al., 1998; Millard et al., 2014; Ye et al., 2000; Zhang and LiWang, 2014). Thirteen of the twenty-one sites identified as under purifying selection are located in close vicinity with these regions (Figure 4.3): Ala14, Leu22, Phe36, Leu48, Thr55, Cys59, Pro60, Ala63, Lys72, Cys75, Asp77, Lys81 and Val83. A similar pattern was observed for CCL3, CCL4 and CCL5 in leporids, where these CCR5 ligands have sites under purifying selection probably due to functional binding constraints (de Matos et al., 2014).

Secretory proteins such as CCL11 only become functional after crossing the membrane and arriving to the appropriate cellular compartment and

consequently being cleaved by signal peptidases (Singh et al., 2013; Thakur et al., 2013). The signal peptide plays important roles in targeting and membrane insertion and after being cleaved can also exert other functions such as protecting cells from being killed by other cells (Hegde, 2002; Li et al., 2009; Singh et al., 2013; Thakur et al., 2013). Before cleavage, the signal peptide may have also important functions in protein folding and maturation (Hegde, 2002). Therefore, mutations in the signal peptide may interfere with such functions. Indeed, from the 23 amino acids that compose the CCL11 signal peptide, five were found to be under purifying selection confirming the importance of their maintenance (Ser4, Ala6, Leu11, Ala14 and Leu22).

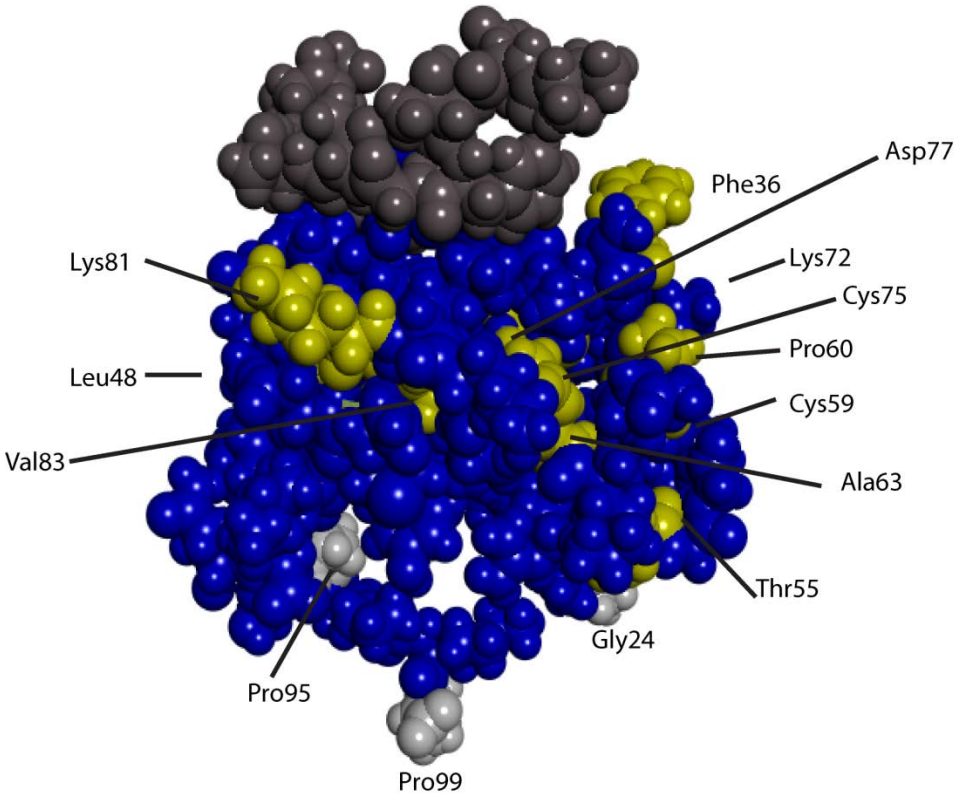


Figure 4. 4. 3D structure of the CCL11-CCR3 complex. CCL11 appears in blue while CCR3 appears in dark grey. The purifying selected amino acids identified in this study and located near important sites for ligand-receptor binding and interaction are marked in yellow; the positively selected sites are highlighted in light grey.

Signatures of positive selection in immune system genes tend to be associated with regions where the binding with other molecules occurs (proteins, receptors) and that can alter the proteins' activity and consequently their biological roles (Hage et al., 1999; Koyanagi et al., 2010; LaPorte et al.,

2008; Zhang et al., 2002). For CCL11, these regions had not been described, but the amino acids changes observed as positively selected show different polarity and charge that can cause changes in the protein. In addition, two of the positively selected sites (Pro95 and Pro99) are located in close vicinity to the stop codon that is mutated in several mammalian species, leading to proteins with different sizes.

Since the CCL11 predicted secondary structures of Human and the European rabbit obtained in PsiPred were similar (data not shown), we used the Human 3D structure available online (CCL11-CCR3 interaction) to locate the sites under selection (Figure 4.4). This confirmed their relevance for receptor binding. A previous study in lagomorphs identified one site under positive selection in chemokine ligand CCL5 (de Matos et al., 2014), but this was located in the signal peptide that is cleaved in protein maturation.

5. CONCLUSIONS

This work describes the detection of codons under positive and purifying selection in CCL11. Purifying selection may result from the proteins' functional constraints, while an increase in diversity, probably as such mutations are advantageous in the host response against several agents, suggests positive selection. Our results identified 21 codons under purifying selection in sites located in regions important for ligand-receptor binding and activation and in the signal peptide and 3 sites under positive selection near signal peptide and the stop codon. We observed that CCL11 is functional in lagomorphs, with pygmy rabbit having a longer protein due to a mutation in the stop codon. Further functional studies should evaluate the biological implications of this extension.

6. REFERENCES

Abrantes, J., Carmo, C.R., Matthee, C.A., Yamada, F., van der Loo, W., Esteves, P.J., 2011. **A shared unusual genetic change at the chemokine receptor type 5 between Oryctolagus, Bunolagus and Pentalagus**. Conserv Genet 12, 325-330.

Allen, S.J., Crown, S.E., Handel, T.M., 2007. **Chemokine: receptor structure, interactions, and antagonism**. Annu Rev Immunol 25, 787-820.

Areal, H., Abrantes, J., Esteves, P.J., 2011. **Signatures of positive selection in Toll-like receptor (TLR) genes in mammals**. BMC Evol Biol 11, 368.

Carmo, C.R., Esteves, P.J., Ferrand, N., van der Loo, W., 2006. **Genetic variation at chemokine receptor CCR5 in leporids: alteration at the 2nd extracellular domain by gene conversion with CCR2 in Oryctolagus, but not in Sylvilagus and Lepus species**. Immunogenetics 58, 494-501.

Crump, M.P., Rajarathnam, K., Kim, K.S., Clark-Lewis, I., Sykes, B.D., 1998. **Solution structure of eotaxin, a chemokine that selectively recruits eosinophils in allergic inflammation**. J Biol Chem 273, 22471-22479.

Crump, M.P., Spyracopoulos, L., Lavigne, P., Kim, K.S., Clark-lewis, I., Sykes, B.D., 1999. **Backbone dynamics of the human CC chemokine eotaxin: fast motions, slow motions, and implications for receptor binding**. Protein Sci 8, 2041-2054.

D'Ambrosio, D., Panina-Bordignon, P., Sinigaglia, F., 2003. **Chemokine receptors in inflammation: an overview**. J Immunol Methods 273, 3-13.

de Matos, A.L., Lanning, D.K., Esteves, P.J., 2014. **Genetic characterization of CCL3, CCL4 and CCL5 in leporid genera Oryctolagus, Sylvilagus and Lepus**. Int J Immunogenet 41, 154-158.

Delport, W., Poon, A.F., Frost, S.D., Kosakovsky Pond, S.L., 2010. **Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology**. Bioinformatics 26, 2455-2457.

Edgar, R.C., 2004. **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. Nucleic Acids Res 32, 1792-1797.

Esteves, P.J., Abrantes, J., van der Loo, W., 2007. **Extensive gene conversion between CCR2 and CCR5 in domestic cat (Felis catus)**. Int J Immunogenet 34, 321-324.

Fass, D., 2012. **Disulfide bonding in protein biophysics**. Annu Rev Biophys 41, 63-79.

Hage, T., Sebald, W., Reinemer, P., 1999. **Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface**. Cell 97, 271-281.

Hall, T.A., 1999. **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT**. Nucl. Acids. Symp. Ser. 41, 95-98.

Hegde, R.S., 2002. **Targeting and beyond: new roles for old signal sequences**. Molecular cell 10, 697-698.

Helenius, A., Aebi, M., 2004. **Roles of N-linked glycans in the endoplasmic reticulum**. Annual review of biochemistry 73, 1019-1049.

Jose, P.J., Griffiths-Johnson, D.A., Collins, P.D., Walsh, D.T., Moqbel, R., Totty, N.F., Truong, O., Hsuan, J.J., Williams, T.J., 1994. **Eotaxin: a potent eosinophil chemoattractant cytokine detected in a guinea pig model of allergic airways inflammation**. J Exp Med 179, 881-887.

Koyanagi, M., Kerns, J.A., Chung, L., Zhang, Y., Brown, S., Moldoveanu, T., Malik, H.S., Bix, M., 2010. **Diversifying selection and functional analysis of interleukin-4 suggests antagonism-driven evolution at receptor-binding interfaces**. BMC evolutionary biology 10, 223.

LaPorte, S.L., Juo, Z.S., Vaclavikova, J., Colf, L.A., Qi, X., Heller, N.M., Keegan, A.D., Garcia, K.C., 2008. **Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system**. Cell 132, 259-272.

Lemos de Matos, A., McFadden, G., Esteves, P.J., 2013. **Positive evolutionary selection on the RIG-I-like receptor genes in mammals**. PLoS One 8, e81864.

Lemos de Matos, A., McFadden, G., Esteves, P.J., 2014. **Evolution of viral sensing RIG-I-like receptor genes in Leporidae genera Oryctolagus, Sylvilagus, and Lepus**. Immunogenetics 66, 43-52.

Li, Y.D., Xie, Z.Y., Du, Y.L., Zhou, Z., Mao, X.M., Lv, L.X., Li, Y.Q., 2009. **The rapid evolution of signal peptides is mainly caused by relaxed selection on non-synonymous and synonymous sites**. Gene 436, 8-11.

Librado, P., Rozas, J., 2009. **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data**. Bioinformatics 25, 1451-1452.

Majewski, J., Ott, J., 2003. **Amino acid substitutions in the human genome: evolutionary implications of single nucleotide polymorphisms**. Gene 305, 167-173.

Manns, J., Rieder, S., Escher, S., Eilers, B., Forssmann, W.G., Elsner, J., Forssmann, U., 2007. **The allergy-associated chemokine receptors CCR3 and CCR5 can be inactivated by the modified chemokine NNY-CCL11**. Allergy 62, 17-24.

Metzger, K.J., Thomas, M.A., 2010. **Evidence of positive selection at codon sites localized in extracellular domains of mammalian CC motif chemokine receptor proteins**. BMC Evol Biol 10, 139.

Millard, C.J., Ludeman, J.P., Canals, M., Bridgford, J.L., Hinds, M.G., Clayton, D.J., Christopoulos, A., Payne, R.J., Stone, M.J., 2014. **Structural basis of receptor sulfotyrosine recognition by a CC chemokine: the N-terminal region of CCR3 bound to CCL11/eotaxin-1**. Structure 22, 1571-1581.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., Scheffler, K., 2013. **FUBAR: a fast, unconstrained bayesian approximation for inferring selection**. Mol Biol Evol 30, 1196-1205.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Kosakovsky Pond, S.L., 2012. **Detecting individual sites subject to episodic diversifying selection**. PLoS genetics 8, e1002764.

Neves, F., Abrantes, J., Lissovsky, A.A., Esteves, P.J., 2015. **Pseudogenization of CCL14 in the Ochotonidae (pika) family**. Innate immunity 21, 647-654.

Neves, F., Abrantes, J., Steinke, J.W., Esteves, P.J., 2014. **Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals**. Innate immunity 20, 184-191.

Nomiyama, H., Osada, N., Yoshie, O., 2010. **The evolution of mammalian chemokine genes**. Cytokine Growth Factor Rev 21, 253-262.

Pease, J.E., Williams, T.J., 2001. **Eotaxin and asthma**. Current opinion in pharmacology 1, 248-253.

Perelygin, A.A., Zharkikh, A.A., Astakhova, N.M., Lear, T.L., Brinton, M.A., 2008. **Concerted evolution of vertebrate CCR2 and CCR5 genes and the origin of a recombinant equine CCR5/2 gene**. J Hered 99, 500-511.

Pinheiro, A., Neves, F., Lemos de Matos, A., Abrantes, J., van der Loo, W., Mage, R., Esteves, P.J., 2016. **An overview of the lagomorph immune system and its genetic diversity**. Immunogenetics 68, 83-107.

Pinheiro, A., Woof, J.M., Abi-Rached, L., Parham, P., Esteves, P.J., 2013. **Computational analyses of an evolutionary arms race between mammalian immunity mediated by immunoglobulin A and its subversion by bacterial pathogens**. PLoS One 8, e73934.

Pinheiro, A., Woof, J.M., Almeida, T., Abrantes, J., Alves, P.C., Gortazar, C., Esteves, P.J., 2014. **Leporid immunoglobulin G shows evidence of strong selective pressure on the hinge and CH3 domains**. Open Biol 4, 140088.

Poon, A.F., Frost, S.D., Pond, S.L., 2009. **Detecting signatures of selection from DNA sequences using Datamonkey**. Methods Mol Biol 537, 163-183.

Puxeddu, I., Bader, R., Piliponsky, A.M., Reich, R., Levi-Schaffer, F., Berkman, N., 2006. **The CC chemokine eotaxin/CCL11 has a selective profibrogenic effect on human lung fibroblasts**. J Allergy Clin Immunol 117, 103-110.

Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A., Dwek, R.A., 2001. **Glycosylation and the immune system**. Science 291, 2370-2376.

Shibata, K., Nomiyama, H., Yoshie, O., Tanase, S., 2013. **Genome diversification mechanism of rodent and Lagomorpha chemokine genes**. Biomed Res Int 2013, 856265.

Shields, D.C., 2000. **Gene conversion among chemokine receptors**. Gene 246, 239-245.

Singh, P., Sharma, L., Kulothungan, S.R., Adkar, B.V., Prajapati, R.S., Ali, P.S., Krishnan, B., Varadarajan, R., 2013. **Effect of signal peptide on stability and folding of Escherichia coli thioredoxin**. PLoS One 8, e63442.

Stevenson, N.J., Addley, M.R., Ryan, E.J., Boyd, C.R., Carroll, H.P., Paunovic, V., Bursill, C.A., Miller, H.C., Channon, K.M., McClurg, A.E., Armstrong, M.A., Coulter, W.A., Greaves, D.R., Johnston, J.A., 2009. **CCL11 blocks IL-4 and GM-CSF signaling in hematopoietic cells and hinders dendritic cell differentiation via suppressor of cytokine signaling expression**. J Leukoc Biol 85, 289-297.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. **MEGA6: Molecular Evolutionary Genetics Analysis version 6.0**. Mol Biol Evol 30, 2725-2729.

Thakur, S., Normand, P., Daubin, V., Tisa, L.S., Sen, A., 2013. **Contrasted evolutionary constraints on secreted and non-secreted proteomes of selected Actinobacteria**. BMC genomics 14, 474.

Tiessen, A., Perez-Rodriguez, P., Delaye-Arredondo, L.J., 2012. **Mathematical modeling and comparison of protein size distribution in different plant, animal, fungal and microbial species reveals a negative correlation between protein size and protein number, thus providing insight into the evolution of proteomes**. BMC research notes 5, 85.

van der Loo, W., Afonso, S., de Matos, A.L., Abrantes, J., Esteves, P.J., 2012. **Pseudogenization of the MCP-2/CCL8 chemokine gene in European rabbit (genus Oryctolagus), but not in species of Cottontail rabbit (Sylvilagus) and Hare (Lepus)**. BMC Genet 13, 72.

Vazquez-Salat, N., Yuhki, N., Beck, T., O'Brien, S.J., Murphy, W.J., 2007. **Gene conversion between mammalian CCR2 and CCR5 chemokine receptor genes: a potential mechanism for receptor dimerization**. Genomics 90, 213-224.

Yang, Z., 2002. **Inference of selection from multiple species alignments**. Current opinion in genetics & development 12, 688-694.

Yang, Z., 2007. **PAML 4: phylogenetic analysis by maximum likelihood**. Mol Biol Evol 24, 1586-1591.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. **Codon-substitution models for heterogeneous selection pressure at amino acid sites**. Genetics 155, 431-449.

Ye, J., Kohli, L.L., Stone, M.J., 2000. **Characterization of binding between the chemokine eotaxin and peptides derived from the chemokine receptor CCR3**. J Biol Chem 275, 27250-27257.

Zhang, J.L., Simeonowa, I., Wang, Y., Sebald, W., 2002. **The high-affinity interaction of human IL-4 and the receptor alpha chain is constituted by two independent binding clusters**. J Mol Biol 315, 399-407.

Zhang, L., LiWang, P.J., 2014. **Chemokine-receptor interactions: solving the puzzle, piece by piece**. Structure 22, 1550-1552.

7. SUPLEMENTARY MATERIAL

Table 4. 5. Characterization of the CCL11 amino acids differences between lagomorphs and other mammals

Human amino acid position	Amino acids						
	Lagomorphs						
	Leporids						
	European rabbit	European brown hare	Brush rabbit	Volcano rabbit	Pygmy rabbit	American pika	Other mammals
2K^{*+}			K^{*+}			M$^{\#}$	K^{*+}, Q*, R^{*+}
5A$^{\#}$			A$^{\#}$			S*	A$^{\#}$, T*, M$^{\#}$, G$^{\#}$
6A$^{\#}$			A$^{\#}$			T*	A$^{\#}$, V$^{\#}$, M$^{\#}$, T*, E^{*-}, G$^{\#}$
9W$^{\#}$			C*			G$^{\#}$	F$^{\#}$, C*, W$^{\#}$, S*, G$^{\#}$
10L$^{\#}$	M$^{\#}$	V$^{\#}$	M$^{\#}$	V#	T*	L$^{\#}$	V$^{\#}$, M$^{\#}$, L$^{\#}$
12L$^{\#}$	V$^{\#}$			L$^{\#}$			I$^{\#}$, F$^{\#}$
13I$^{\#}$			T*			I$^{\#}$	T*, L$^{\#}$, S*, I$^{\#}$, V$^{\#}$, A$^{\#}$, W$^{\#}$
16A$^{\#}$			V$^{\#}$			T*	S*, A$^{\#}$, L$^{\#}$, V$^{\#}$, T*, I$^{\#}$
19P$^{\#}$			S*				T*, I$^{\#}$, S*, P$^{\#}$, N*, A$^{\#}$
21G$^{\#}$			V$^{\#}$				G$^{\#}$, I$^{\#}$, V$^{\#}$, A$^{\#}$
24G$^{\#}$			Q*				H^{*+}, Q*, L$^{\#}$, E^{*-}, G$^{\#}$
29S*			F$^{\#}$				F$^{\#}$, S*, L$^{\#}$, K^{*+}, A$^{\#}$, T*, V$^{\#}$
32T*			T*			A$^{\#}$	T*, R^{*+}, S*, A$^{\#}$, K^{*+}, N*, I$^{\#}$
37N*			S*				N*, V$^{\#}$, S*, I$^{\#}$, T*, A$^{\#}$, R^{*+}
38L$^{\#}$			M$^{\#}$				V$^{\#}$, L$^{\#}$, M$^{\#}$
40N*	-	-	-	-	-	K^{*+}	S*, K^{*+}, G$^{\#}$, R^{*+}, N*, P$^{\#}$, T*, A$^{\#}$
41R^{*+}			K^{*+}				R^{*+}, K^{*+}
43I$^{\#}$	M$^{\#}$			I$^{\#}$			I$^{\#}$
44P$^{\#}$	P$^{\#}$			H^{*+}		P$^{\#}$	S*, L$^{\#}$, P#
45L$^{\#}$	L$^{\#}$	I$^{\#}$	L$^{\#}$	I$^{\#}$	L$^{\#}$	T*	L$^{\#}$, T*, N*, F$^{\#}$, I$^{\#}$, V$^{\#}$, K^{*+}, M$^{\#}$
53R^{*+}			R^{*+}			N*	R^{*+}, K^{*+}, Q*, I$^{\#}$, T*
55T*			S*				T*, S*, N*, I$^{\#}$, K^{*+}
56S*			G$^{\#}$			S*	N*, S*, G$^{\#}$, D^{*-}, A$^{\#}$
57G$^{\#}$			S*				N*, S*, T*, G$^{\#}$
65I$^{\#}$	I$^{\#}$, M$^{\#}$			M$^{\#}$		I$^{\#}$, S*	I$^{\#}$, V$^{\#}$, L$^{\#}$
66F$^{\#}$			F$^{\#}$			L$^{\#}$	F$^{\#}$
70L$^{\#}$		L$^{\#}$		F$^{\#}$		L$^{\#}$	L$^{\#}$, Q*, R^{*+}, A$^{\#}$, S*, M$^{\#}$, P$^{\#}$
73D^{*-}			E^{*-}				K^{*+}, E^{*-}, D^{*-}, M$^{\#}$
76A$^{\#}$			A$^{\#}$			V$^{\#}$	A$^{\#}$, V$^{\#}$, T*
80K^{*+}			E^{*-}				K^{*+}, E^{*-}, Q*, A$^{\#}$, N*, V$^{\#}$
85D^{*-}	D^{*-}, N*			D^{*-}			D^{*-}, N*, E^{*-}, A$^{\#}$
87M$^{\#}$			I$^{\#}$				I$^{\#}$, T*, M$^{\#}$, V$^{\#}$, L$^{\#}$, K^{*+}
89Y*			Y*			S*	Y*, H^{*+}
95P$^{\#}$			K^{*+}			R^{*+}	P$^{\#}$, Q*, R^{*+}, K^{*+}, S*, T*, L$^{\#}$, H^{*+}
97P$^{\#}$			S*				P$^{\#}$, T*, L$^{\#}$, S*, Q*, V$^{\#}$, A$^{\#}$
99P$^{\#}$			P$^{\#}$			S*	P$^{\#}$, Y*, S*, H^{*+}, K^{*+}, L$^{\#}$, I$^{\#}$

EVOLUTION OF CCL16 IN GLIRES (RODENTIA AND LAGOMORPHA) SHOWS AN UNUSUAL RANDOM PSEUDOGENIZATION PATTERN

Fabiana Neves, Joana Abrantes, Ana M Lopes, Maria J Magalhães,Wessel van der Loo, Pedro J Esteves

1. ABSTRACT

The chemokine ligand CCL16 is a strong pro-inflammatory chemokine and a chemoattractant for monocytes and lymphocytes. It is present at high concentrations in normal plasma and elicits its effects on cells by interacting with cell surface chemokine receptors.

The CCL16 gene was identified as a pseudogene in the European rabbit and in some rodents such as mouse, rat and guinea pig, while in squirrel it seems to be functional. To elucidate the evolution of this gene in the superorder Glires, composed by rodents and lagomorphs, we sequenced eleven leporids and seven *Ochotona* species. Additionally, we retrieved the CCL16 sequences of twelve rodents available in public databases. Sequencing of CCL16 showed that similarly to the European rabbit, all leporid species showed pseudogenization of CCL16. In contrast, in *Ochotona* species CCL16 seems functional, except for Hoffmann's pika where one allele is a pseudogene. Pseudogenization of CCL16 in leporids is due to a mutation in the typical CC motif leading to a premature stop codon. In *Sylvilagus* spp. this motif is also mutated, but it codes for a Lysine residue. In addition, in the Mexican cottontail, CCL16 is a pseudogene due to a 95-base pair deletion that leads to a frameshift. Furthermore, we were not able to amplify this gene from Eastern cottontail (*S. floridanus*) cDNA, supporting that in this species CCL16 is a pseudogene. In rodents, the same pattern was observed with species presenting a functional CCL16 while in others it is a pseudogene. This pseudogenization is due to species-specific mutations. Our results suggest that despite being functional in the Glires ancestor, CCL16 underwent pseudogenization in some species. This process occurred stochastically or in

specific lineages at different moments in the evolution of Glires. In addition, in cottontails, although they do not present the same inactivating mutations as in the remaining leporids, CCL16 remains non-functional.

Keywords: Chemokine ligands, CCL16, Evolution, Glires, Pseudogenization

2. INTRODUCTION

Chemokines are small pleiotropic proteins of low-molecular weight with important roles in inflammation, homeostasis and immune response (Ono et al., 2003; Zlotnik and Yoshie, 2012). Chemokines are only found in vertebrates and are classified according to their conserved N-terminal Cysteines residues into CC, CXC, XC, CX3C and CX (only identified in zebrafish) (Nomiyama et al., 2008; Nomiyama et al., 2013). In C-C chemokines, both N-terminus Cysteines are juxtaposed. These proteins are able to exert their function trough the interaction between the residues located in both the extracellular loops and the NH_2-terminus and the chemokine receptors (Crump et al., 1999; Teran, 2000; Van Coillie et al., 1999).

Chemokines originated through gene duplication from an ancestral gene (Juan et al., 2009; Nomiyama et al., 2010) and can evolve by pseudogenization, neofunctionalization or subfunctionalization. In the pseudogenization process, disruption of the coding sequence leads to the loss of function. In the neofunctionalization and subfunctionalization, the genes are maintained in the either by acquiring new functions or maintaining functions with the other copy, respectively (Ding et al., 2012; Levasseur and Pontarotti, 2011; Zhang, 2003).

CCL16, also known as liver-expressed chemokine (LEC) or human C-C chemokine (HCC)-4, is located in the macrophage inflammatory protein (MIP) region of the CC cluster. CCL16 is a strong pro-inflammatory chemokine and a chemoattractant for monocytes and lymphocytes, enhancing their adhesive properties (Nomiyama et al., 2001; Youn et al., 1998). Commonly present at high concentrations in normal plasma, CCL16 elicits its effects on cells by interacting with cell surface chemokine receptors such as CCR1, CCR2, CCR5 and CCR8 (Zlotnik and Yoshie, 2012). In some mammalian species, including some leporids, CCR5 evolved under gene conversion with CCR2

(Abrantes et al., 2011; Carmo et al., 2006; Esteves et al., 2007; Perelygin et al., 2008; Shields, 2000; Vazquez-Salat et al., 2007). Thus, leporids are a good model to study the co-evolution of chemokine receptors and their ligands. Indeed, in the leporids European rabbit (*Oryctolagus cuniculus*), Amami rabbit (*Pentalagus furnessi*) and riverine rabbit (*Bunolagus monticularis*) CCR5 suffered a gene conversion with CCR2 in the second external cellular loop, while cottontail rabbits (*Sylvilagus* spp.) and hares (*Lepus* spp.) have a normal CCR5. In order to determine the consequences of this CCR5-CCR2 gene conversion, the CCR5 chemokine ligands have been studied in leporids. CCL3, CCL4, CCL5 and CCL11 genes are functional in leporids (de Matos et al., 2014; Neves et al., 2016). In rabbit, mouse and rat there is only one copy of CCL3 and CCL4, but in other rodents such as squirrel and guinea pig there are several copies that may be either functional or inactivated (Shibata et al., 2013; Zlotnik et al., 2006). The study of CCL8 in leporids showed that this gene is pseudogenized in the European rabbit, riverine rabbit and Amami rabbit, while intact in hares and Eastern cottontail (*S. floridanus*) (van der Loo et al., 2012; van der Loo et al., 2016). CCL14 is functional in the Leporidae family while in the Ochotonidae family it is suffering a pseudogenization process with some species presenting an intact CCL14 while in others it is disrupted (Neves et al., 2015b). Interestingly, mouse and rat lack the CCL14 gene (Shibata et al., 2013).

The superorder Glires includes two orders, Rodentia (rodents) and Lagomorpha (rabbits, hares and pikas) that diverged at approximately 82 million years ago (mya) (Hedges et al., 2015). Rodentia is the most diverse order from placental mammals with 2.277 species within 33 families (Blanga-Kanfi et al., 2009). Several phylogenies proposed for rodents. According to Blanga-Kanfi et al., (2009) there are two main groups of rodents that diverged ~ 73 mya. The first group is composed by the Octodontoidea (degu), Chinchilloideia (long tailed chinchilla), Cavioidea (guinea pig) and Bathyergidae (naked mole-rat and damaraland mole-rat) families that were separated ~ 43 mya. The second group includes the Heteromyidae (ord's kangaroo rat), Dipodidae (lesser Egyptian jerboa), Cricetidae (Chinese and golden hamsters) and Muridae (mouse and rat) families that diverged ~70 mya (Blanga-Kanfi et al., 2009). Lagomorpha includes only two families, Ochotonidae (pikas) and Leporidae (rabbits and

hares), that diverged at ~ 9.2 mya (Matthee et al., 2004). The Ochotonidae family is composed of only one genus, *Ochotona*, which is divided into four subgenera, *Pika*, *Ochotona*, *Conothoa* and *Lagotona* (Lissovsky, 2014). The Leporidae family comprises eleven genera *Poelagus*, *Pronolagus*, *Nesolagus*, *Oryctolagus*, *Caprolagus*, *Bunolagus*, *Pentalagus*, *Brachylagus*, *Sylvilagus*, *Lepus* and *Romerolagus*. From these, *Oryctolagus* is most closely related with *Bunolagus* and *Pentalagus*, while *Brachylagus* and *Sylvilagus* are closest to each other (Matthee et al., 2004).

CCL16 is a pseudogene in the European rabbit and in some rodents such as mouse, rat and guinea pig, while in squirrel it seems to be functional (Shibata et al., 2013). Here, we genetically characterized CCL16 in lagomorphs. We further included the sequences available for several rodent species aiming at determine the evolution of this gene in the superorder Glires. With the exception of cottontails, in the remaining leporids CCL16 is a pseudogene due to a mutation in the juxtaposed Cysteines that leads to a stop codon. Cottontail rabbits also present a mutation in this motif, but it leads to a Lysine. In addition, Mexican cottontails possess a frameshift mutation in the first exon making it also a pseudogene. In the Ochotonidae family, with the exception of Hoffmann's pika, CCL16 seems functional. In Hoffmann's pika, CCL16 presents two alleles, one functional and the other with the same mutation in the CC motif observed in leporids. Interestingly, in rodents there are also some species where CCL16 is a pseudogene while in others it is functional, however due to mutations different than those described for lagomorphs.

3. MATERIALS AND METHODS

Genomic DNA was extracted using the EasySpin Genomic DNA Minipreps Tissue Kit (Citomed, Torun, Poland) from tissue samples of European rabbit (*Oryctolagus cuniculus cuniculus and Oryctolagus cuniculus algirus),* riverine rabbit (*Bunolagus monticularis*), Amami rabbit (*Pentalagus furnessi*), pygmy rabbit (*Brachylagus idahoensis),* Mexican cottontail (*Sylvilagus cunicularis*), forest cottontail (*Sylvilagus brasiliensis*), Eastern cottontail (*Sylvilagus floridanus*), European brown hare (*Lepus europaeus),* Iberian hare (*Lepus granatensis*), volcano rabbit (*Romerolagus diazi*), American pika

(*Ochotona princeps*), Northern pika (*O. hyperborea*), manchurian pika (*O. mantchurica*), steppe pika (*O. pusilla*), Hoffmann's pika (*O. hoffmanni*), Palla's pika (*O. pallasi*) and turuchan pika (*O. turuchanensis*) according to manufacturer's instructions. Total RNA was extracted from one specimen of American pika and three of Eastern cottontail by using the RNeasy Mini Kit according to the manufacturer's instructions (Qiagen, Hilden, Germany). cDNA was synthesized using oligo(dT) as primers and SuperScriptIII reverse transcriptase (Invitrogen, Carlsbad, CA, USA). The European rabbit and American pika sequences available in GenBank were used for primer design. PCR amplification from gDNA was performed by amplification of several overlapping fragments with Multiplex PCR Kit (Qiagen, Hilden, Germany), according to the manufacturer's protocol (PCR conditions are available upon request). Sequencing was performed on an ABI PRISM 310 Genetic Analyser (PE Applied Biosystems) and PCR products were sequenced in both directions. Sequences were submitted to GenBank under the following accession numbers: xxxxxx.

The sequences obtained in this work were aligned with other CCL16 sequences available in Genbank. For Rodentia all sequences available in Genbanl were collected. Additionally we added CCL16 for the most representative Primates, Artyodactyla, Carnivores, etc. Sequences were aligned using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) available at http://www.ebi.ac.uk/ (Edgar, 2004) and translated using BioEdit (Hall, 1999). Furthermore, we retrieved sequences available for all the 28 chemokine ligands for Primates, Lagomorpha, Rodentia, Arctiodactyla and Carnivora and performed a phylogenetic analysis. The Maximum Likelihood tree was estimated in MEGA6 by using the best-fit nucleotide substitution model predicted by the same software and 1000 bootstrap replicates (Tamura et al., 2013).

Splicing sites were predicted by using the NetGene2 server available at http://www.cbs.dtu.dk/services/NetGene2/ (Brunak et al., 1991; Hebsgaard et al., 1996).

4. RESULTS AND DISCUSSION

In both the European rabbit and mouse, CCL16 is described as being a pseudogene located in the MIP region of chromosomes 19 and 11, respectively. In this study we amplified and sequenced the CCL16 gene for eleven leporids: both subspecies of European rabbit (*Oryctolagus cuniculus cuniculus* and *Oryctolagus cuniculus algirus),* riverine rabbit (*Bunolagus monticularis*), Amami rabbit (*Pentalagus furnessi*), pygmy rabbit (*Brachylagus idahoensis),* Mexican cottontail (*Sylvilagus cunicularis*), forest cottontail (*Sylvilagus brasiliensis*), Eastern cottontail (*Sylvilagus floridanus*), European brown hare (*Lepus europaeus),* Iberian hare (*Lepus granatensis*) and volcano rabbit (*Romerolagus diazi*). In addition, we sequenced the CCL16 gene for seven pika species: American pika (*Ochotona princeps*), Northern pika (*O. hyperborea*), manchurian pika (*O. mantchurica*), steppe pika (*O. pusilla*), Hoffmann's pika (*O. hoffmanni*), palla's pika (*O. pallasi*) and turuchan pika (*O. turuchanensis*).

An amino acid alignment of the CCL16 sequences for lagomorphs, rodents and other mammalian groups (Artiodactyla, Carnivores, Chiroptera, Primates, etc.) is depicted in Figure 4.5. The CCL16 sequences obtained for leporids were quite different from those from the remaining mammals. Indeed, when comparing the cottontails' CCL16 sequences with the human sequence they only share 48 amino acids. However, C-C chemokines are similar between themselves. Thus, we performed a phylogenetic analysis of the sequences obtained along with all the C-C chemokine ligand sequences available for Primates, Lagomorpha, Rodentia, Arctiodactyla and Carnivora. In the phylogenetic analysis our CCL16 sequences clustered with the other CCL16 sequences with good bootstrap values (Figure 4.6).

Leporidae

Our results further (Figure 4.7) showed that, like in the European rabbit, in the riverine rabbit, Amami rabbit, pygmy rabbit, European brown hare, Iberian hare and volcano rabbit, CCL16 is a pseudogene. In all these leporids, a non-synonymous mutation (C>A) at codon 45 leads to a premature stop codon (TGC>TGA) and disrupts the typical C-C chemokines juxtaposed Cysteines (Cys45-Cys46). Interestingly, in the Mexican, forest and Eastern cottontail rabbits, the Cys45 also suffered a mutation, but it encodes a Lysine (K).

Figure 4.5.

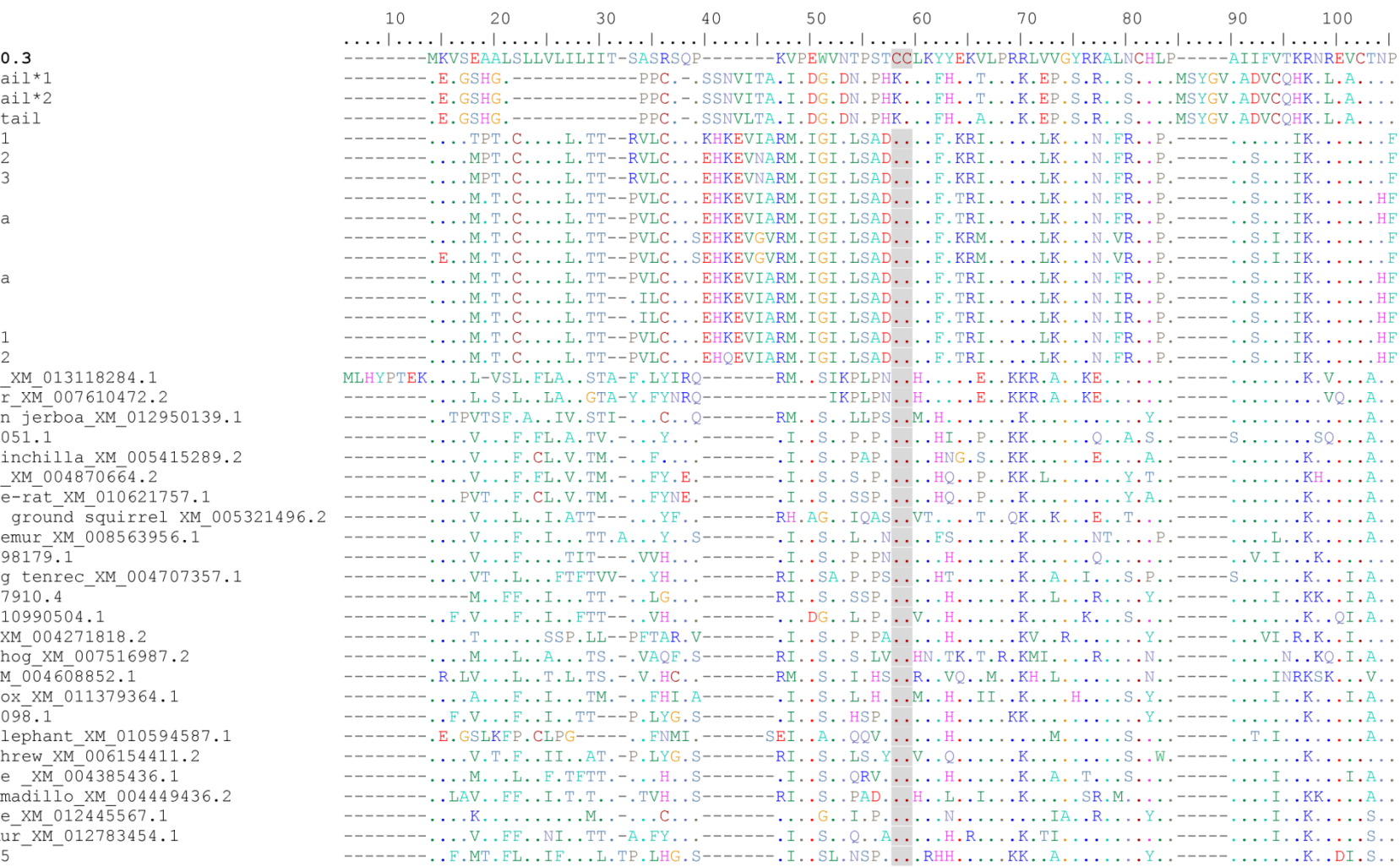

10 20 30 40 50 60 70 80 90 100
Human_NM_004590.3
Forest cottontail*1
Forest cottontail*2
Eastern cottontail
American pika*1
American pika*2
American pika*3
Northern pika
Manchurian pika
Steppe pika*1
Steppe pika*2
Hoffmann's pika
Palla's pika*1
Palla's pika*2
Turuchan pika*1
Turuchan pika*2
Golden hamster_XM_013118284.1
Chinese hamster_XM_007610472.2
Lesser Egyptian jerboa_XM_012950139.1
Degu XM_004643051.1
Long-tailed chinchilla_XM_005415289.2
Naked mole-rat_XM_004870664.2
Damaraland mole-rat_XM_010621757.1
Thirteen-lined ground squirrel XM_005321496.2
Sunda flying lemur_XM_008563956.1
Cattle_XM_010798179.1
Lesser hedgehog tenrec_XM_004707357.1
Horse_XM_001917910.4
Dromedary_XM_010990504.1
Killer whale_ XM_004271818.2
European hedgehog_XM_007516987.2
Common shrew_XM_004608852.1
Large flying fox_XM_011379364.1
Cat_ XM_006940098.1
African bush elephant_XM_010594587.1
Chinese tree shrew_XM_006154411.2
Florida manatee _XM_004385436.1
Nine-banded armadillo_XM_004449436.2
Aotus_nancymaae_XM_012445567.1
Gray_mouse_lemur_XM_012783454.1
Dog_XM_537724.5

Figure 4. 5. Amino acid alignment of CCL16 for several mammalian species. (*) represent normal stop codons; (-) represent indels; *1, *2 and *3 represent different alleles.

Human (*Homo sapiens*), forest cottontail (*Sylvilagus brasiliensis*), Eastern cottontail (*Sylvilagus floridanus*), American pika (*Ochotona princeps*), Northern pika (*O. hyperborea*), manchurian pika (*O. mantchurica*), steppe pika (*O. pusilla*), Hoffmann's pika (*O. hoffmanni*), Palla's pika (*O. pallasi*), turuchan pika (*O. turuchanensis*), golden hamster (*Mesocricetus auratus*), Chinese hamster (*Cricetulus griseus*), lesser Egyptian jerboa (*Jaculus jaculus*), ord's kangaroo rat (*Dipodomys ordii*), guinea pig (*Cavia porcellus*), degu (*Octodon degus*), long tailed chinchilla (*Chinchilla lanigera*), naked mole-rat (*Heterocephalus glaber*), damaraland mole-rat (*Fukomys damarensis*), thirteen lined ground squirrel (*Ictidomys tridecemlineatus*), alpine marmot (*Marmota marmota*), sunda flying lemur (*Galeopterus variegatus*), cattle (*Bos taurus*), lesser hedgehog tenrec (*Echinops telfairi*), horse (*Equus caballus*), Arabian camel (*Camelus dromedarius*), killer whale (*Orcinus orca*), European hedgehog (*Erinaceus europaeus*), common shrew (*Sorex araneus*), large flying fox (*Pteropus vampyrus*), cat (*Felis catus*), African bush elephant (*Loxodonta africana*), Chinese tree shrew (*Tupaia belangeri chinensis*), Florida manatee (*Trichechus manatus latirostris*), nine-banded armadillo (*Dasypus novemcinctus*), nancy Ma's night monkey (*Aotus nancymaae*), gray mouse lemur (*Microcebus murinus*), dog (*Canis lupus familiaris*). Numbering used in the text is according to human CCL16 sequence (GenBank accession number; NM_004590.3), with signal peptide and indels (indicated as (-)) being included in the numbering.

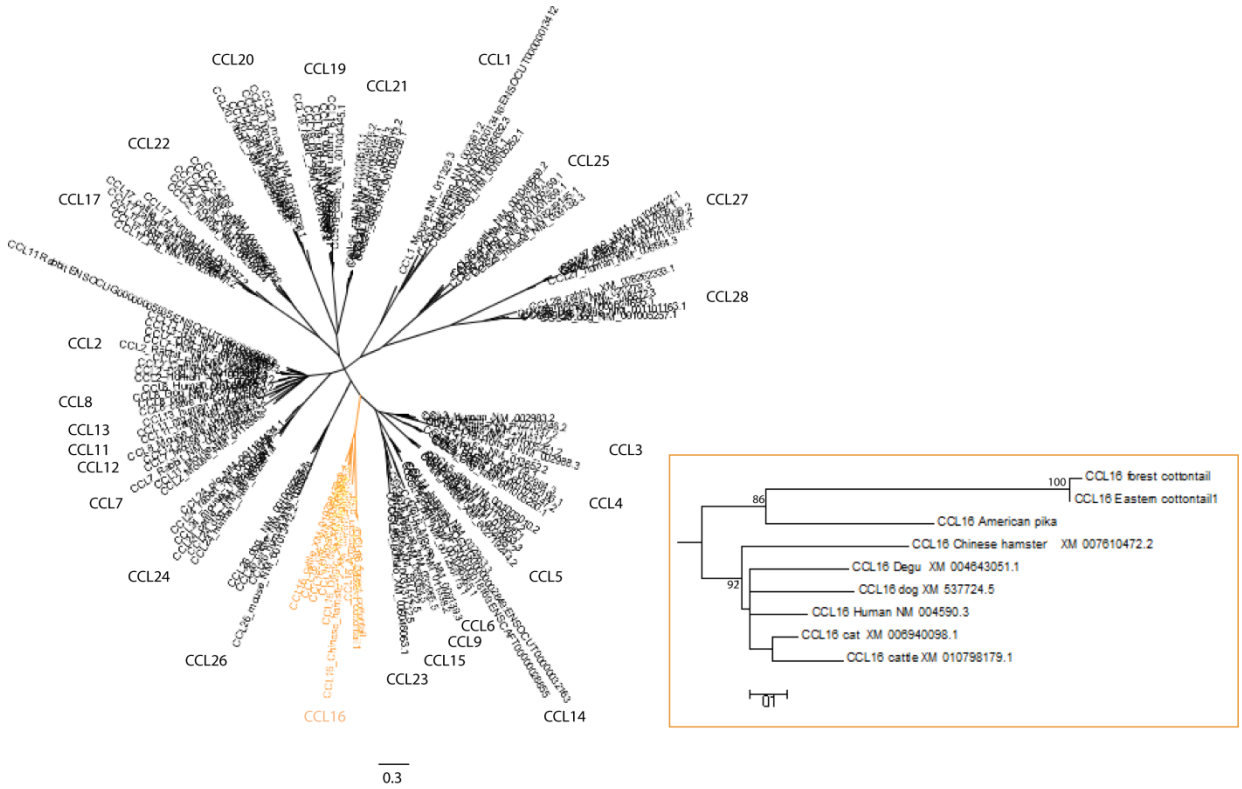


Figure 4. 6. Maximum likelihood (ML) tree of the mammalian C-C chemokine ligands CCL1-CCL28. The clade composed of CCL16 sequences is shown in detail. Only bootstrap values ≥ 75% are shown.

In addition, in the Mexican cottontail, the CCL16 gene presents a deletion of 95-base pairs (bp) between nucleotides 22 and 117 (see Figure 4.7) that further leads to a frameshift.

To assess whether the CCL16 gene was expressed in cottontails, we attempted its amplification from cDNA of three samples of Eastern cottontail. These cDNA samples had been previously successfully amplified for other genes of the innate and adaptive immune system (Lavazza et al., 2015; Neves et al., 2015a; Neves et al., 2015b; Pinheiro et al., 2014). Despite all attempts, amplification failed, suggesting that CCL16 is not expressed in the Eastern cottontail although no disrupting mutations were detected.

The exchange of one of the Cysteines for a Lysine might have suggested the possibility that the protein acquired a new function. However, the unsuccessful amplification from cDNA of the Eastern cottontail along with the frameshift observed in the Mexican cottontail may indicate that CCL16 is not functional and is probably undergoing a pseudogenization process in cottontails. We hypothesize that the Cys>stop codon mutation was already present in the ancestor of leporids and that, for some unknown reason, was

“restored” into a Lysine in the cottontail branch at ~9.2 mya. In addition, in the cottontail lineage two distinct scenarios are possible: some species possess an intact CCL16 gene, however it is not expressed at the mRNA level, while in others it is not functional due to other pseudogenization mutations. Thus, we suggest that although the CCL16 gene suffered mutations that could make it functional, in cottontails it “remains” nonfunctional.

Other species-specific mutations that can lead to pseudogenization were also observed. Indeed, *Lepus* spp., riverine rabbit, Amami rabbit and volcano rabbit have a bp deletion at nucleotide positions 4, 100, 142 and 275, respectively (Figure 4.7). In addition, *Lepus*, volcano rabbit and pygmy rabbit also present deletions of 4, 10 and 29 bp, respectively. Furthermore, pygmy rabbit and riverine rabbit present other mutations that lead to stop codons. These mutations occur in the pygmy rabbit at nucleotide position 148 (GAG (Glu) > TGA) and in the riverine rabbit in position 184 (AGA (Arg) > TGA). In addition, amplification of CCL16 from gDNA of the Amami rabbit revealed an insertion of 24 nucleotides at the end of the first exon (from position 206 to 234). All these deletions were probably due to independent events that occurred in different moments in the evolution of leporids, likely being lineage-specific.

Ochotonidae

We successfully amplified the American pika CCL16 from both cDNA and gDNA. The American pika sequence obtained from cDNA presented some differences when comparing with the sequence available in Ensembl (ENSOPRG00000012019) (Figure 4.8a). The sequence available in Ensembl presents several indels and misses the stop codon suggesting a non-functional CCL16. However, our sequences (gDNA and cDNA) seem to be functional and present an insertion of twenty one nucleotides that correspond to an insertion of seven amino acids in the beginning of exon 2. The complete sequence of the CCL16 gene (three exons and the two introns) showed that this insertion derives from intron 1 (Figure 4.8b). This insertion might have resulted from the appearance of an alternative splicing site in the American pika CCL16 gene. This alternative splicing site occurs in a CA motif located in the intron 21 bp upstream of the CA motif that in human CCL16 gene immediately flanks the exon 2. These results were further confirmed by comparing the human and

American pika CCL16 sequences in NetGene2. Indeed, for the American pika, NetGene2 predicted that the splicing occurs at nucleotide position 15 while in human it corresponds to nucleotide position 36 (according to the American pika sequence; Figure 4b). Studies on other chemokines showed that human CXCL12, CCL4, CCL20, CCL23 and CCL27 also exhibit alternative splicing, leading to novel and functional proteins (Colobran et al., 2007). Alternative splicing is a crucial step in the mature mRNA production (Shakola et al., 2015). This process leads to protein diversity and can occur by exon skipping (38%); alternative 5' or 3' spliced sites (26%); intron retention (3%); mutually exclusive exons, alternative promoters or multiple polyadenylation sites (33%) (Keren et al., 2010; Sahoo and Im, 2010). In the American pika CCL16 gene, alternative splicing occurred by intron retention (seven amino acids).

For pikas, we observed that, with the exception of Hoffmann's pika, the CCL16 gene seems to encode a functional protein. Interestingly, in Hoffmann's pika we identified two alleles, one coding for an intact gene and the other, similar to leporids, presenting a mutation in the Cys45 that leads to a premature stop codon. This stop codon has the same location as found for leporids suggesting that this region may be prone to suffer mutations which is at odd as this site is important for disulfide bond formation, and thus alterations in this motif may alter protein structure and consequently, its function.

C-C chemokines are characterized by two juxtaposed Cysteines that in CCL16 correspond to amino acids 45 and 46. The loss of one of these Cysteines due to a mutation that encodes a premature stop codon leads to pseudogenization of this chemokine. Moreover, the mutation into an amino acid different than a Cys impairs the protein to exert its functions. This is the case for almost all leporids studied and one allele of Hoffmann's pika. The presence of the same mutation in the two families of the order Lagomorpha might be explained by parallel evolution in two different lineages, leporids and Hoffmann's pika, such that the same mutation occurred independently at different moment in the lagomorphs' evolution. Alternatively, this Cys-stop codon mutation was already present in the lagomorphs' ancestor and was later "distributed" stochastically with some species presenting the stop mutation whilst others do not.

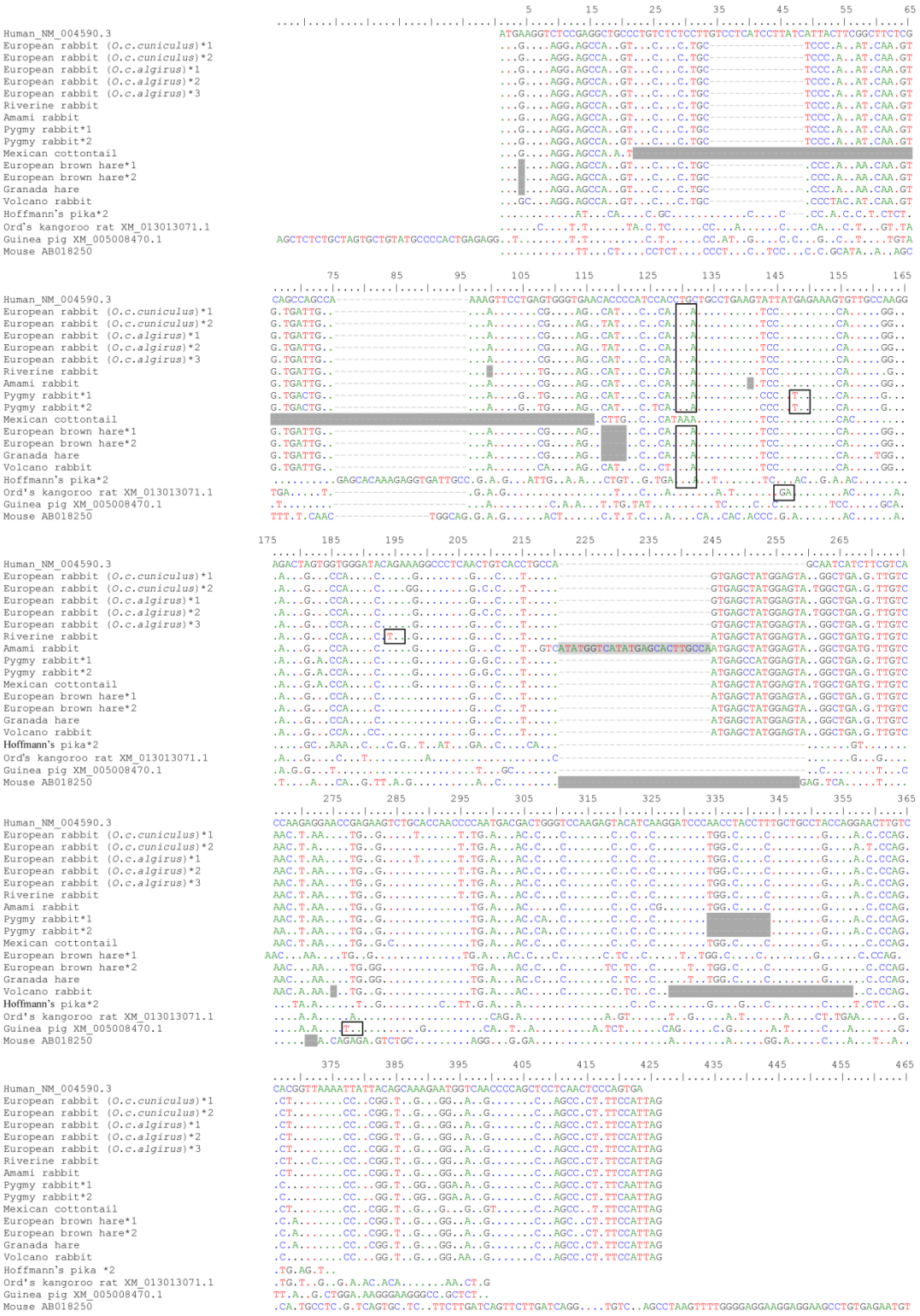


Figure 4. 7. Detail of the nucleotide alignment for the different CCL16 pseudogenes. Insertions in the coding sequence are shaded in light grey while frameshifts mutations are shaded in dark grey. A black dashed box represents the extra stop codons; (*) represent normal stop codons; (-) represent indels; *1, *2 and *3 represent different alleles. Human (*Homo sapiens*), European rabbit (*Oryctolagus cuniculus cuniculus* and *Oryctolagus cuniculus algirus),* riverine rabbit (*Bunolagus monticularis*), Amami rabbit (*Pentalagus furnessi*), pygmy rabbit (*Brachylagus idahoensis),* Mexican cottontail (*Sylvilagus cunicularis*), Hoffmann's pika (*O. hoffmanni*), ord's kangaroo rat (*Dipodomys ordii*), guinea pig (*Cavia porcellus*), mouse (*Mus musculus*). Numbering used in the text is according to human CCL16 sequence, with signal peptide and indels (indicated as (-)) being included in the numbering.

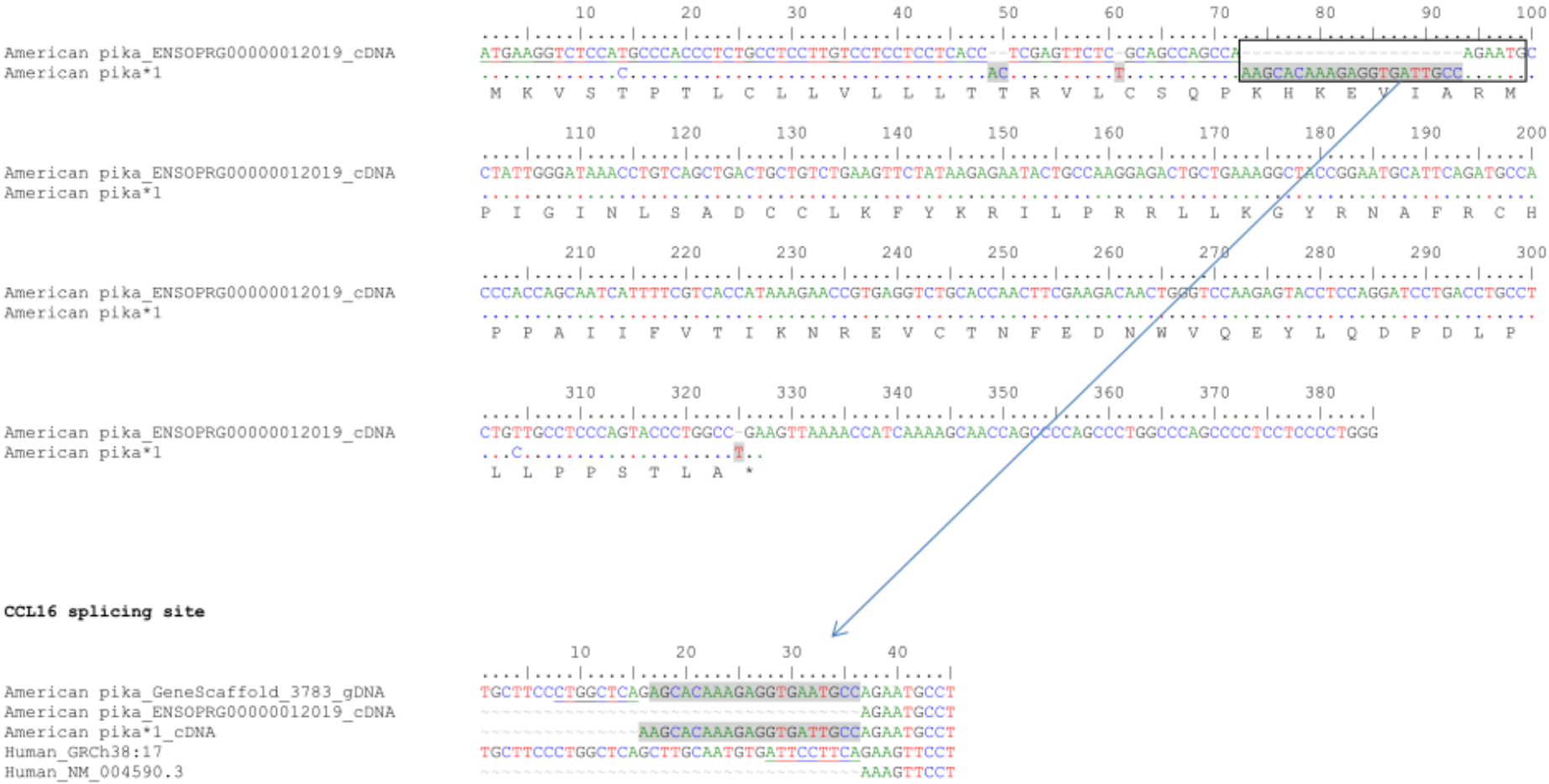


Figure 4. 8 a) Comparison of the American pika CCL16 sequences from Ensembl and obtained in this study. b) Detail of the alternative splicing site in the American pika CCL16 gene. Ensembl accession numbers is indicated in bold for the American pika retrieved sequence. Insertions in the coding sequence are shaded in light grey. (*) represent normal stop codons; (-) represent indels. Human (*Homo sapiens*), American pika (*Ochotona princeps*).

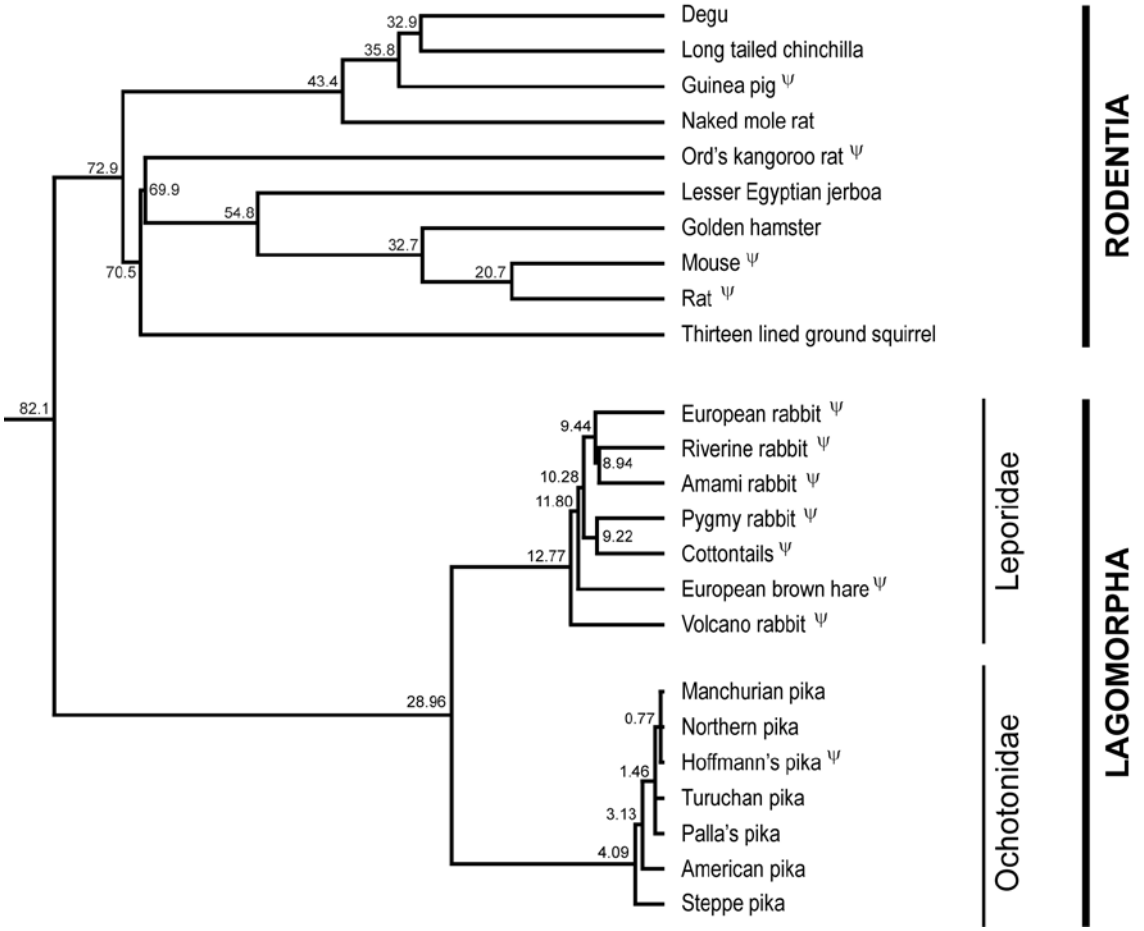


Figure 4. 9. Phylogenetic relationships within the clade Glires. Divergence times (in million years) are indicated in the nodes. Relationships within the Leporidae family are based on a molecular super matrix (adapted from Mathee et al. 2015, while for the Ochotonidae family it is based in a multilocus coalescent approach (adapted from (Melo-Ferreira et al., 2015). Within Rodentia, relationships are according to (Hedges et al., 2015). ψ indicate the pseudogenes.

Rodentia

Considering that in rodents some species have a functional CCL16 while in others CCL16 is a pseudogene (Shibata et al., 2013), we retrieved twelve rodent CCL16 sequences available in public databases (NCBI, Ensembl and UNIPROT). We observed that besides mouse, rat and guinea pig, CCL16 is also a pseudogene in the ord's kangaroo rat (Figure 4.10). In these species, CCL16 is a pseudogene due to different mutations. Mouse CCL16 has been reported as a pseudogene due to mutations that lead to the loss of the characteristic juxtaposed conserved Cysteines and an insertion of a Long Interspersed Element–1 (L1) in the third exon (Fukuda et al., 1999). Rat CCL16 is also a pseudogene (Nomiyama et al., 2010; Shibata et al., 2013), but there is no further information on what led to its pseudogenization and no sequence is

available in the public databases. For the ord's kangaroo rat and guinea pig, CCL16 is a pseudogene due to a premature stop codon at positions 145 and 277, respectively. In the remaining sequences available, CCL16 seems to encode a functional protein.

The different mutations observed in different lineages in rodents may indicate that the CCL16 gene was functional in the rodents' ancestor and became posteriorly pseudogenized (Figure 4.10). Indeed, we observed that Muridae (mouse and rat), Heteromyidae (ord's kangaroo rat) and Cavioidea (guinea pig) have a pseudogenized CCL16 gene while in members of the Sciuroidea (thirteen lined ground squirrel and alpine marmot), Cricetidae (Chinese and golden hamsters), Dipodidae (lesser Egyptian jerboa), Bathyergidae (naked mole-rat and damaraland mole-rat), Chinchilloidea (long tailed chinchilla), and Octodontoidea (degu), CCL16 is intact. Thus, the CCL16 pseudogenization occurs stochastically along the Rodentia order.

5. CONCLUSIONS

Overall these results suggest that in Glires (rodents and lagomorhps), CCL16 suffered several independent pseudogenization events, with some species presenting one or both alleles disrupted. Thus, in Glires, we observe distinct scenarios for CCL16: in the order Lagomorpha, some species present an intact CCL16 while in others it is a pseudogene. The same occurs in the Rodentia order, where some members of the Muridae, Heteromyidae and Cavioidea have a pseudogenized CCL16 gene while in members of the Sciuroidea, Cricetidae, Dipodidae, Bathyergidae, Chinchilloidea, and Octodontoidea, CCL16 is intact. This may indicate that although CCL16 was present and functional in the ancestor of the Glires clade, it was later inactivated in some species. This may have occurred stochastically or in specific lineages at different moments in the CCL16 evolution.

6. REFERENCES

Abrantes, J., Carmo, C.R., Matthee, C.A., Yamada, F., van der Loo, W., Esteves, P.J., 2011. **A shared unusual genetic change at the chemokine receptor type 5 between Oryctolagus, Bunolagus and Pentalagus**. Conserv Genet 12, 325-330.

Blanga-Kanfi, S., Miranda, H., Penn, O., Pupko, T., DeBry, R.W., Huchon, D., 2009. **Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades**. BMC Evol Biol 9, 71.

Brunak, S., Engelbrecht, J., Knudsen, S., 1991. **Prediction of human mRNA donor and acceptor sites from the DNA sequence**. J Mol Biol 220, 49-65.

Carmo, C.R., Esteves, P.J., Ferrand, N., van der Loo, W., 2006. **Genetic variation at chemokine receptor CCR5 in leporids: alteration at the 2nd extracellular domain by gene conversion with CCR2 in Oryctolagus, but not in Sylvilagus and Lepus species**. Immunogenetics 58, 494-501.

Colobran, R., Pujol-Borrell, R., Armengol, M.P., Juan, M., 2007. **The chemokine network. II. On how polymorphisms and alternative splicing increase the number of molecular species and configure intricate patterns of disease susceptibility**. Clin Exp Immunol 150, 1-12.

Crump, M.P., Spyracopoulos, L., Lavigne, P., Kim, K.S., Clark-lewis, I., Sykes, B.D., 1999. **Backbone dynamics of the human CC chemokine eotaxin: fast motions, slow motions, and implications for receptor binding**. Protein Sci 8, 2041-2054.

de Matos, A.L., Lanning, D.K., Esteves, P.J., 2014. **Genetic characterization of CCL3, CCL4 and CCL5 in leporid genera Oryctolagus, Sylvilagus and Lepus**. Int J Immunogenet 41, 154-158.

Ding, Y., Zhou, Q., Wang, W., 2012. **Origins of new genes and evolution of their novel functions**. Annu. Rev. Ecol. Evol. Syst. 43, 345-363.

Edgar, R.C., 2004. **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. Nucleic Acids Res 32, 1792-1797.

Esteves, P.J., Abrantes, J., van der Loo, W., 2007. **Extensive gene conversion between CCR2 and CCR5 in domestic cat (Felis catus)**. Int J Immunogenet 34, 321-324.

Fukuda, S., Hanano, Y., Iio, M., Miura, R., Yoshie, O., Nomiyama, H., 1999. **Genomic organization of the genes for human and mouse CC chemokine LEC**. DNA Cell Biol 18, 275-283.

Hall, T.A., 1999. **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT**. Nucl. Acids. Symp. Ser. 41, 95-98.

Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P., Brunak, S., 1996. **Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information**. Nucleic Acids Res 24, 3439-3452.

Hedges, S.B., Marin, J., Suleski, M., Paymer, M., Kumar, S., 2015. **Tree of life reveals clock-like speciation and diversification**. Mol Biol Evol 32, 835-845.

Juan, M., Pedrosa, E., Colobran, R., 2009. **Chemokines and Chemokine Receptors**. eLS. John Wiley & Sons Ltd.

Keren, H., Lev-Maor, G., Ast, G., 2010. **Alternative splicing and evolution: diversification, exon definition and function**. Nat Rev Genet 11, 345-355.

Lavazza, A., Cavadini, P., Barbieri, I., Tizzani, P., Pinheiro, A., Abrantes, J., Esteves, P.J., Grilli, G., Gioia, E., Zanoni, M., Meneguz, P., Guitton, J.S., Marchandeau, S., Chiari, M., Capucci, L., 2015. **Field and experimental data indicate that the eastern cottontail (Sylvilagus floridanus) is susceptible to infection with European brown hare syndrome (EBHS) virus and not with rabbit haemorrhagic disease (RHD) virus**. Vet Res 46, 13.

Levasseur, A., Pontarotti, P., 2011. **The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics**. Biology direct 6, 11.

Lissovsky, A.A., 2014. **Taxonomic revision of pikas Ochotona (Lagomorpha, Mammalia) at the species level**. Mammalia 78, 199-216.

Matthee, C.A., van Vuuren, B.J., Bell, D., Robinson, T.J., 2004. **A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene**. Syst Biol 53, 433-447.

Melo-Ferreira, J., Lemos de Matos, A., Areal, H., Lissovsky, A.A., Carneiro, M., Esteves, P.J., 2015. **The phylogeny of pikas (Ochotona) inferred from a multilocus coalescent approach**. Mol Phylogenet Evol 84, 240-244.

Neves, F., Abrantes, J., Almeida, T., de Matos, A.L., Costa, P.P., Esteves, P.J., 2015a. **Genetic characterization of interleukins (IL-1alpha, IL-1beta, IL-2, IL-4, IL-8, IL-10, IL-12A, IL-12B, IL-15 and IL-18) with relevant biological roles in lagomorphs**. Innate immunity 21, 787-801.

Neves, F., Abrantes, J., Esteves, P.J., 2016. **Evolution of CCL11: genetic characterization in lagomorphs and evidence of positive and purifying selection in mammals**. Innate immunity 22, 336-343.

Neves, F., Abrantes, J., Lissovsky, A.A., Esteves, P.J., 2015b. **Pseudogenization of CCL14 in the Ochotonidae (pika) family**. Innate immunity 21, 647-654.

Nomiyama, H., Hieshima, K., Nakayama, T., Sakaguchi, T., Fujisawa, R., Tanase, S., Nishiura, H., Matsuno, K., Takamori, H., Tabira, Y., Yamamoto, T., Miura, R., Yoshie, O., 2001. **Human CC chemokine liver-expressed chemokine/CCL16 is a functional ligand for CCR1, CCR2 and CCR5, and constitutively expressed by hepatocytes**. International immunology 13, 1021-1029.

Nomiyama, H., Hieshima, K., Osada, N., Kato-Unoki, Y., Otsuka-Ono, K., Takegawa, S., Izawa, T., Yoshizawa, A., Kikuchi, Y., Tanase, S., Miura, R., Kusuda, J., Nakao, M., Yoshie, O., 2008. **Extensive expansion and diversification of the chemokine gene family in zebrafish: identification of a novel chemokine subfamily CX**. BMC genomics 9, 222.

Nomiyama, H., Osada, N., Yoshie, O., 2010. **The evolution of mammalian chemokine genes**. Cytokine Growth Factor Rev 21, 253-262.

Nomiyama, H., Osada, N., Yoshie, O., 2013. **Systematic classification of vertebrate chemokines based on conserved synteny and evolutionary history**. Genes to cells : devoted to molecular & cellular mechanisms 18, 1-16.

Ono, S.J., Nakamura, T., Miyazaki, D., Ohbayashi, M., Dawson, M., Toda, M., 2003. **Chemokines: roles in leukocyte development, trafficking, and effector function**. J Allergy Clin Immunol 111, 1185-1199; quiz 1200.

Perelygin, A.A., Zharkikh, A.A., Astakhova, N.M., Lear, T.L., Brinton, M.A., 2008. **Concerted evolution of vertebrate CCR2 and CCR5 genes and the origin of a recombinant equine CCR5/2 gene**. J Hered 99, 500-511.

Pinheiro, A., Woof, J.M., Almeida, T., Abrantes, J., Alves, P.C., Gortazar, C., Esteves, P.J., 2014. **Leporid immunoglobulin G shows evidence of strong selective pressure on the hinge and CH3 domains**. Open Biol 4, 140088.

Sahoo, A., Im, S.H., 2010. **Interleukin and interleukin receptor diversity: role of alternative splicing**. International reviews of immunology 29, 77-109.

Shakola, F., Suri, P., Ruggiu, M., 2015. **Splicing Regulation of Pro-Inflammatory Cytokines and Chemokines: At the Interface of the Neuroendocrine and Immune Systems**. Biomolecules 5, 2073-2100.

Shibata, K., Nomiyama, H., Yoshie, O., Tanase, S., 2013. **Genome diversification mechanism of rodent and Lagomorpha chemokine genes**. Biomed Res Int 2013, 856265.

Shields, D.C., 2000. **Gene conversion among chemokine receptors**. Gene 246, 239-245.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. **MEGA6: Molecular Evolutionary Genetics Analysis version 6.0**. Mol Biol Evol 30, 2725-2729.

Teran, L.M., 2000. **CCL chemokines and asthma**. Immunology today 21, 235-242.

Van Coillie, E., Van Damme, J., Opdenakker, G., 1999. **The MCP/eotaxin subfamily of CC chemokines**. Cytokine Growth Factor Rev 10, 61-86.

van der Loo, W., Afonso, S., de Matos, A.L., Abrantes, J., Esteves, P.J., 2012. **Pseudogenization of the MCP-2/CCL8 chemokine gene in European rabbit (genus Oryctolagus), but not in species of Cottontail rabbit (Sylvilagus) and Hare (Lepus)**. BMC Genet 13, 72.

van der Loo, W., Magalhaes, M.J., de Matos, A.L., Abrantes, J., Yamada, F., Esteves, P.J., 2016. **Adaptive Gene Loss? Tracing Back the Pseudogenization of the Rabbit CCL8 Chemokine**. J Mol Evol 83, 12-25.

Vazquez-Salat, N., Yuhki, N., Beck, T., O'Brien, S.J., Murphy, W.J., 2007. **Gene conversion between mammalian CCR2 and CCR5 chemokine receptor genes: a potential mechanism for receptor dimerization**. Genomics 90, 213-224.

Youn, B.S., Zhang, S., Broxmeyer, H.E., Antol, K., Fraser, M.J., Jr., Hangoc, G., Kwon, B.S., 1998. **Isolation and characterization of LMC, a novel lymphocyte and monocyte chemoattractant human CC chemokine, with myelosuppressive activity**. Biochem Biophys Res Commun 247, 217-222.

Zhang, J., 2003. **Evolution by gene duplication: an update**. TRENDS in Ecology and Evolution 18, 292-298.

Zlotnik, A., Yoshie, O., 2012. **The chemokine superfamily revisited**. Immunity 36, 705-716.

Zlotnik, A., Yoshie, O., Nomiyama, H., 2006. **The chemokine and chemokine receptor superfamilies and their molecular evolution**. Genome biology 7, 243.

CHAPTER 5

Final considerations

General discussion

Future perspectives

1. GENERAL DISCUSSION

The major goal of this dissertation was to contribute to the current knowledge on the lagomorphs' immune system. The European rabbit has been widely used as a laboratory model in immunological studies. Since the 1950s, the European rabbit population has decreased abruptly, mostly due to the arrival of two viral diseases, myxomatosis and rabbit hemorrhagic disease (RHD). The emergence of these diseases has lead to the co-evolution between the host and the respective pathogens. Myxomatosis is caused by a poxvirus, the myxoma virus (MYXV), which has its natural host in the forest cottontail (*Sylvilagus brasiliensis*) where it only causes a cutaneous lesion at the inoculation site, while in the European rabbit it is fatal (Kerr et al., 2015; Spiesschaert et al., 2011). Myxomatosis is caused by a poxvirus, the MYXV, responsible for high mortality in the European rabbit, being useful as a model to study host-pathogen interactions. RHD is caused by a calicivirus, the rabbit hemorrhagic disease virus (RHDV), which is also responsible for high mortality in the European rabbit. RHDV and RHD were also used as a model to better understand acute liver failure, a human critical illness with high mortality rates (Garcia-Lastra et al., 2010; San-Miguel et al., 2014; Tunon et al., 2011a; Tunon et al., 2011b; Vallejo et al., 2014). With larger incidence in humans with no pre-existing liver diseases, the acute liver failure leads to a rapid deterioration of the liver functions (Bernal and Wendon, 2013). Thus, the European rabbit and its viral diseases are important models to further understand similar human diseases. However, most of these studies focused on the pathogens rather than in the host. Genes of the innate and adaptive host immune system are candidate genes to explore the processes involving the host-pathogen co-evolution. In order to further improve our current knowledge on the host immune system genes that may have important roles against these viral pathogens, we investigated the evolutionary and genetic aspects of some of these genes. Our results on the C-C chemokine ligands (CCLs) 11, 14 and 16 and interleukins (ILs) identified unique characteristics in lagomorphs. These findings should be further complemented with functional studies to determine their biological relevance.

Innate Immune System

The innate immune system is the first line of defense against pathogens and includes, among others, the Toll-like receptors (TLRs), peptidoglycan recognition proteins, NOD-like receptors, cytokines and chemokines (Owen et al., 2013). Previous studies on lagomorphs' innate immune system genes included TLRs (Abrantes et al., 2013; Chen et al., 2014; Elfeil et al., 2016; Vaure and Liu, 2014), cytokines (Gertz et al., 2011; Godornes et al., 2007; Kerr et al., 2004; Liu et al., 2009; Mage and Mage, 2012; Perkins et al., 2000; Schnupf and Sansonetti, 2012; Siewe et al., 2010; Trzeciak-Ryczek et al., 2016) and chemokines receptors (Abrantes et al., 2011; Abrantes et al., 2008; Carmo et al., 2006; de Matos et al., 2014; Schnupf and Sansonetti, 2012; van der Loo et al., 2012; van der Loo et al., 2016).

CCR5-CCR2 gene conversion is well documented in lagomorphs, where a characteristic motif of the second extracellular loop of CCR2 (HTIMRN) replaces the CCR5 sequence motif QTLKMT (Carmo et al., 2006), characterized as crucial for ligand specificity (Samson et al., 1997). Interestingly, this event was not observed in *Sylvilagus* or *Lepus* species (Abrantes et al., 2011; Carmo et al., 2006). This gene conversion event prompted the study of the CCR5 ligands (Zlotnik and Yoshie, 2012). Previous studies revealed that CCL3, CCL4 and CCL5 genes are functional in leporids and under strong purifying selection (de Matos et al., 2014) while CCL8 is pseudogenized in the European rabbit, riverine rabbit and Amami rabbit, while intact in hares and Eastern cottontails (van der Loo et al., 2012; van der Loo et al., 2016). Thereby, our main goal was to genetically characterize the remaining CCR5 ligands in lagomorphs, expecting to better understand the implications of the CCR5-CCR2 gene conversion in *Oryctolagus*, *Sylvilagus* and *Lepus.*

Considering the role of interleukins in RHDV (IL1, IL2, IL6, IL8, and IL10) (Garcia-Lastra et al., 2010; Marques et al., 2012; Marques et al., 2014; Teixeira et al., 2012; Trzeciak-Ryczek et al., 2016); MYXV (IL4, IL12, IL15 and IL18) (Johnston and McFadden, 2004; Kerr et al., 2004; Liu et al., 2009; Stanford and McFadden, 2005; Tosic et al., 2014; Vande Walle and Lamkanfi, 2011) and bacterial (IL17 and IL22) infections (Schnupf and Sansonetti, 2012; Skyberg et

al., 2013), we performed a genomic characterization of these interleukins in several lagomorph species.

A dynamic immune system requires an efficient recognition of the invading pathogens. For this, a wide array of antigens must be efficiently recognized and cells of the immune system must be able to quickly adapt and respond to changes in the normal conditions. Therefore, such an intricate system involves highly complex networks and relies heavily on several mechanisms of gene regulation to succeed by obtaining the required diversity and flexibility. The immune system genes studied here revealed some of the mechanisms that ensure the plasticity of the immune system and that ultimately leads to successfully eradicate the invading pathogens.

Evolutionary forces

Genes of the immune system are constantly under pressure from invading microorganisms. In vertebrates, the innate immune response precedes and shapes the adaptive immunity. Therefore, alterations in the vertebrate innate immunity genes may have crucial consequences in the development of an appropriate adaptive immune response (Quintana-Murci and Clark, 2013). There is some controversy regarding the selective pressures acting in the immune system. It is well-known that genes associated with the immune system present high evolutionary rates and signatures of positive selection are commonly detected due to adverse environmental conditions and the arms race established between pathogens and the host (Barreiro and Quintana-Murci, 2010; Ferrer-Admetlla et al., 2008). Nevertheless, in order to maintain their conformation and biological role, immune system genes are also under purifying selection (Metzger and Thomas, 2010). Previous studies on human innate immunity genes evidenced signals of both positive (Vallender and Lahn, 2004) and purifying selection (Mukherjee et al., 2009). Our results show that in lagomorphs, the interleukins studied and CCL14 are also under strong positive selection. On the other hand, and as observed for other chemokine ligands (de Matos et al., 2014), CCL11 is under both positive and purifying selection. Interestingly, we detected both selective pressures in regions responsible for the interaction with other proteins such as receptors and binding proteins.

Mutations in immune system genes

Gene mutation is a permanent change in the DNA sequence that can be inherited or acquired due to distinct selective pressures or owing to errors during DNA replication. Studies in different mammals, particularly in humans, showed that mutations in immune system genes might be responsible for the development of diseases (Buckley, 2005; Carvalho et al., 2010; Ochs et al., 2014). For example, mutations in human IL2 are associated with an increase of Mendelian susceptibility to mycobacterial disease (MSMD), while mutations in the C-terminal of CXCR4 are associated with Myelokathexis (Rosenzweig and Holland, 2011). In this work, several different mutations that lead to different alterations in innate immune system genes of lagomorphs were detected. Indeed, CCL11 is functional in lagomorphs, but the pygmy rabbit presents a longer protein due to a mutation in the stop codon. Despite being functional in leporids, in some *Ochotona* species (*O. princeps*, *O. pallasi*, *O. alpina* and *O. turuchanensis*, *O. hyperborea* and *O. pusilla*) the CCL14 is a pseudogene due to distinct disrupting mutations. For CCL16, some lagomorphs (the European rabbit, riverine rabbit, Amami rabbit, pygmy rabbit, European brown hare, Iberian hare, volcano rabbit and Hoffmann's pika) present a mutation in the characteristic juxtaposed cysteines leading to a premature stop codon. Interestingly, the only leporids that present a possible functional CCL16 are the cottontails (with the exception of the Mexican cottontail) where this mutation codes for a Lysine. Nevertheless, and despite all the attempts, CCL16 could not be amplified from Eastern cottontail cDNA, suggesting that the CCL16 may not be expressed. For IL6, the European rabbit and the Amami rabbit, when compared with other mammals, present a mutation that leads to longer proteins with extra 27 and 17 amino acids, respectively.

Alternative Splicing

Alternative splicing is an important feature of gene regulation being responsible for tissue and species-specific differentiation patterns (Keren et al., 2010; Kornblihtt et al., 2013). Alternative splicing includes the recognition and removal of introns and the attachment of different exons by several splicing proteins to form the mature mRNA (Keren et al., 2010). This mechanism can

increase protein diversity by producing distinct mature transcripts that translate into different proteins. Over 90% of human genes are subject to alternative splicing (Wang et al., 2008). Alternative splicing may also result in proteins with different biological roles and with different outcomes within cells, alterations in the biochemical properties of the proteins and ultimately, in the loss of function (Shakola et al., 2015; Yabas et al., 2015). Only a few studies have focused on the importance of alternative splicing in regulating the functions of the immune system and there are still many unanswered questions. Thus, in order to fully comprehend immunity-related diseases, it is crucial to understand the mechanisms underlining the alternative splicing and also the mechanisms that regulate the splicing events in the normal immune response (Martinez and Lynch, 2013). Indeed, some isoforms of cytokine receptors mediate antagonistic functions in cytokine signaling leading, to alterations in cytokine function (Sahoo and Im, 2010; Shakola et al., 2015). Moreover, cytokines modulate the expression of several genes associated with the alternative splicing machinery (Ortis et al., 2010). Previous studies have highlighted the existence of different isoforms in the European rabbit interleukins IL1Ra, which suppress the IL1 anti-inflammatory activity (Cominelli et al., 1994; Cominelli et al., 1992); IL2, IL4 and IL10, where the new isoforms lack some exons (Perkins et al., 2000), and IL7II that may act as an antagonist of IL7 (Siewe et al., 2010). Our results further showed the existence of alternative splicing in the second exon of American pika IL8 and CCL16 that originated from the 5' untranslated regions (UTRs).

N-glycosylation

N-glycosylation consists in the attachment of a N-glycan to a nitrogen atom of the amino acid residue asparagine by a N-glycosidic bond. This process is considered crucial for protein function being involved in several biological mechanisms such as protein folding, ligand binding, protease protection, oligomerization, intrinsic stability, leukocyte trafficking, and expression (Marth and Grewal, 2008; Moremen et al., 2012; Varki and Lowe, 2009). In addition, glycans are present on the surface of mammalian cells, attached to the ectodomain of PRRs, being used as recognition motifs and therefore, are essential to mediate cell-to-cell interactions and in pathogen

recognition. Glycans are also responsible for most of the cell-specific inter-variability (Lee and Lee, 1995; van Kooyk and Rabinovich, 2008; Wolfert and Boons, 2013). Thus, alterations in glycosylation sites are variable and unpredictable and changes in these regions can have different effects in different cells (Varki and Lowe, 2009). For example, studies showed that alterations in N-glycosylation sites of interferon-γ receptor 2 disrupt the ability of this receptor to bind to its ligands, and were also associated with Mendelian susceptibility to mycobacterial disease (Marth and Grewal, 2008; Vogt et al., 2005). Furthermore, alterations in N-glycans of IL21R were linked to combined immunodeficiency syndrome (Kotlarz et al., 2013). In this work we detected alterations in N-glycosylation sites for IL1, IL6, IL10, IL12B, IL15 and IL17. Indeed, in some of these proteins these alterations are predicted to induce structural modifications, namely in the number of helixes.

Disulfide bonds

Formation of disulfide bonds is an important post-translational modification that can have crucial implications in the protein folding and integrity (Hogg, 2003). These bridges are formed between the thiol groups of cysteine residues being well conserved between related-proteins and species (Fass, 2012; Li et al., 2011; Thangudu et al., 2008). Some studies indicated that mutations in the cysteines that form the disulfide bonds do not significantly alter the proteins' structure, however, they might affect their activity and sometimes can lead to alterations in their function (Li et al., 2011; Thangudu et al., 2008). Disulfide bonds with a crucial role in the protein are usually highly conserved. The non-conservation of these bonds within homologous proteins can be related with mutations that favor other interactions or with other inconspicuous structural alterations that compensate this loss, being also associated with differentiation and specialization of the protein function (Thangudu et al., 2008). Nevertheless, studies on IL2 and its receptors showed that alterations in these proteins' disulfide bonds lead to the lack of interaction, loss of function and were also associated with development of disease (Metcalfe et al., 2012; Niemela et al., 2000; Wang et al., 2005). Our results show that alterations in cysteines important for disulfide bonds lead to an extra helix in the American pika IL1β

and in the European rabbit IL6. Moreover, in CCL16 our results evidenced that in the majority of leporids and in Hoffman's pika one of the two N-terminal cysteine residues that form disulfide bridges with the two C-terminal cysteines of the molecule is mutated into a stop codon.

Since the 1980s the European rabbit has been used as an animal model and important advances were made on human health, namely for rabies and syphilis (Knell, 2004; Pasteur, 1885). However, the European rabbit was gradually replaced as an animal model for human diseases by mouse (*Mus musculus*). However, from an immunological point of view, it is now known that the European rabbit presents some unique characteristics such as organ's size, longer lifetime which allows to study chronic aspects of a disease and are carriers or reservoirs of different pathogens that are able to cause zoonotic diseases (Fontanesi et al., 2016; Kamaruzaman et al., 2013). These specific features make the European rabbit a better model to study human diseases. Indeed several authors stated that for some immunological studies the rabbit may be a better model than mouse (Burkhardt and Zlotnik, 2013; Kamaruzaman et al., 2013; Keir and Page, 2008; Mullane and Williams, 2014; Perkins et al., 2000; Pinheiro et al., 2016; Seok et al., 2013; Shay et al., 2015; Takao and Miyakawa, 2015; Webb, 2014). Our results show that for interleukins, the European rabbit is closer to human further supporting that the European rabbit might be a better animal model to study the role of human interleukins in the inflammatory process.

2. FUTURE PERSPECTIVES

The work developed in this thesis shed light on some aspects of the lagomorphs' innate immune system. However, many questions croped up and remain unanswered. Future lines of research should take them into account.

Despite the knowledge gathered with this thesis regarding different innate immune system genes, the biological implications of the alterations found in lagomorphs and in the development of an effective immune response are still largely speculative. Several reasons might underlie our findings, including the "acquisition" of specific biological functions and the maintenance of crucial structural features. However, the limited knowledge of the role of specific

codons in the functions of the proteins encoded by these genes hampers the complete understanding of our observations. In addition, the information available for the genes studied is mostly restricted to human with almost no information available for other mammals. Our results provide essential contributions to the knowledge of the innate immune response in lagomorphs that could have implications for the initiation and perpetuation of immune-mediated disorders. It is clear that a multitude of apparently non-redundant regulatory mechanisms act in concert to fine-tune and modulate the initiation, duration and magnitude of the immune response. An extensive knowledge of all of these mechanisms can provide significant information on our understanding of the European rabbit diseases such as myxomatosis and RHDV and lead to the development of new therapeutics.

Functional and structural studies by using mutagenesis and crystallographic approaches should be performed for a full comprehension of the role and consequences of the observed alterations. With a site-directed mutagenic approach we will be able to induce specific mutations into a protein with a known sequence (Carter, 1986; Kunkel, 1985). Crystallographic studies are used to obtain the structure of a protein, being also helpful to determine how some interactions influence the target of the protein and what structural changes occur upon interaction (Scapin, 2006). Thus, the application of these techniques to the proteins studied will allows us to determine the impact of the mutations described in the protein function, structure and interactions. An effort should also be made to integrate functional transcriptomic and proteomic analyses in different organs and tissues in order to link the observed mutations with their emerging new roles in lagomorphs.

Several aspects of lagomorphs' immune system remain unclear. Many other host immune system genes might be responsible for susceptibility/resistance to the different diseases and outcomes in leporid species. Indeed, Toll-like receptors (TLRs) are important proteins with the ability to detect a wide range of pathogens, being important for both innate and adaptive immunity. These pattern recognition receptors play a central role in response against danger signals (Christmas, 2010) and might be important for RHDV infection (Abrantes et al., unpublished observations). In lagomorphs, studies are only available for TLR1, TLR2, TLR3 and TLR4 (Abrantes et al.,

2013; Chen et al., 2014; Elfeil et al., 2016; Vaure and Liu, 2014), but should be further extended to the other TLRs.

Protein kinase R (PKR) is an antiviral protein that can be found in most vertebrates and can be induced by type I IFN. When activated this protein phosphorylates the α subunit of the eukaryotic initiation factor 2 (eIF2) that leads to the inhibition of virus replication (Elde et al., 2009; Sadler and Williams, 2007). Interestingly, recent studies evidenced that myxoma virus variant MYVX156 is homologous to eIF2 and is able to inhibit the rabbit PKR (Peng et al., 2016). In addition, a single mutation in this variant blocks this inhibition (Burgess and Mohr, 2016; Peng et al., 2016). These results might be related with the MYXV-rabbit co-evolution.

Immunoglobulins (Igs) are proteins able to recognize and bind to several different antigens as well as recruit effector functions for the pathogen removal. Thus, in order to successfully infect the host, pathogens evolved several strategies to escape from the Igs activity (Owen et al., 2013). Due to its important functions in the European rabbit, Igs have been widely studied, with the main focus in the Igs heavy chain (Esteves et al., 2006; Esteves et al., 2005; Gertz et al., 2013; Lavinder et al., 2014; Pinheiro et al., 2011; Pinheiro et al., 2013; Pinheiro et al., 2014; Ros et al., 2004). There are few studies on the Immunoglobulin light chain, with emphasis on the kappa (κ) locus (Mage et al., 2006; Popkov et al., 2003; Ros et al., 2005). Indeed, Gertz et al. (2013) is the only study performed in recent years that focus on the immunoglobulin light chain lambda locus (IgLλ) and summarizes the available information for the European rabbit.

3. REFERENCES

Abrantes, J., Areal, H., Esteves, P.J., 2013. **Insights into the European rabbit (Oryctolagus cuniculus) innate immune system: genetic diversity of the toll-like receptor 3 (TLR3) in wild populations and domestic breeds**. BMC Genet 14, 73.

Abrantes, J., Carmo, C.R., Matthee, C.A., Yamada, F., van der Loo, W., Esteves, P.J., 2011. **A shared unusual genetic change at the chemokine receptor type 5 between Oryctolagus, Bunolagus and Pentalagus**. Conserv Genet 12, 325-330.

Abrantes, J., Esteves, P.J., Carmo, C.R., Muller, A., Thompson, G., van der Loo, W., 2008. **Genetic characterization of the chemokine receptor CXCR4 gene in lagomorphs: comparison between the families Ochotonidae and Leporidae**. Int J Immunogenet 35, 111-117.

Barreiro, L.B., Quintana-Murci, L., 2010. **From evolutionary genetics to human immunology: how selection shapes host defence genes**. Nature reviews. Genetics 11, 17-30.

Bernal, W., Wendon, J., 2013. **Acute liver failure**. N Engl J Med 369, 2525-2534.

Buckley, R.H., 2005. **Variable phenotypic expression of mutations in genes of the immune system**. J Clin Invest 115, 2974-2976.

Burgess, H.M., Mohr, I., 2016. **Evolutionary clash between myxoma virus and rabbit PKR in Australia**. Proc Natl Acad Sci U S A 113, 3912-3914.

Burkhardt, A.M., Zlotnik, A., 2013. **Translating translational research: mouse models of human disease**. Cell Mol Immunol 10, 373-374.

Carmo, C.R., Esteves, P.J., Ferrand, N., van der Loo, W., 2006. **Genetic variation at chemokine receptor CCR5 in leporids: alteration at the 2nd extracellular domain by gene conversion with CCR2 in Oryctolagus, but not in Sylvilagus and Lepus species**. Immunogenetics 58, 494-501.

Carter, P., 1986. **Site-directed mutagenesis**. The Biochemical journal 237, 1-7.

Carvalho, A., Cunha, C., Pasqualotto, A.C., Pitzurra, L., Denning, D.W., Romani, L., 2010. **Genetic variability of innate immunity impacts human susceptibility to fungal diseases**. Int J Infect Dis 14, e460-468.

Chen, C., Zibiao, H., Ming, Z., Shiyi, C., Ruixia, L., Jie, W., SongJia, L., 2014. **Expression pattern of Toll-like receptors (TLRs) in different organs and effects of lipopolysaccharide on the expression of TLR 2 and 4 in reproductive organs of female rabbit**. Dev Comp Immunol 46, 341-348.

Christmas, P., 2010. **Toll-Like Receptors: Sensors that Detect Infection**. Nature Education 3, 85.

Cominelli, F., Bortolami, M., Pizarro, T.T., Monsacchi, L., Ferretti, M., Brewer, M.T., Eisenberg, S.P., Ng, R.K., 1994. **Rabbit interleukin-1 receptor antagonist. Cloning, expression, functional characterization, and regulation during intestinal inflammation**. J Biol Chem 269, 6962-6971.

Cominelli, F., Nast, C.C., Duchini, A., Lee, M., 1992. **Recombinant interleukin-1 receptor antagonist blocks the proinflammatory activity of endogenous interleukin-1 in rabbit immune colitis**. Gastroenterology 103, 65-71.

de Matos, A.L., Lanning, D.K., Esteves, P.J., 2014. **Genetic characterization of CCL3, CCL4 and CCL5 in leporid genera Oryctolagus, Sylvilagus and Lepus**. Int J Immunogenet 41, 154-158.

Elde, N.C., Child, S.J., Geballe, A.P., Malik, H.S., 2009. **Protein kinase R reveals an evolutionary model for defeating viral mimicry**. Nature 457, 485-489.

Elfeil, W.M., Algammal, A.M., Abouelmaatti, R.R., Gerdouh, A., Abdeldaim, M., 2016. **Molecular characterization and analysis of TLR-1 in rabbit tissues**. Cent Eur J Immunol 41, 236-242.

Esteves, P.J., Carmo, C., Godinho, R., van der Loo, W., 2006. **Genetic diversity at the hinge region of the unique immunoglobulin heavy gamma (IGHG) gene in leporids (Oryctolagus, Sylvilagus and Lepus)**. Int J Immunogenet 33, 171-177.

Esteves, P.J., Lanning, D., Ferrand, N., Knight, K.L., Zhai, S.K., van der Loo, W., 2005. **The evolution of the immunoglobulin heavy chain variable region (IgVH) in Leporids: an unusual case of transspecies polymorphism**. Immunogenetics 57, 874-882.

Fass, D., 2012. **Disulfide bonding in protein biophysics**. Annu Rev Biophys 41, 63-79.

Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marques-Bonet, T., Ramirez-Soriano, A., Muntasell, A., Navarro, A., Lazarus, R., Calafell, F., Bertranpetit, J., Casals, F., 2008. **Balancing selection is the main force shaping the evolution of innate immunity genes**. J Immunol 181, 1315-1322.

Fontanesi, L., Di Palma, F., Flicek, P., Smith, A.T., Thulin, C.G., Alves, P.C., Lagomorph Genomics, C., 2016. **LaGomiCs-Lagomorph Genomics Consortium: An International Collaborative Effort for Sequencing the Genomes of an Entire Mammalian Order**. J Hered 107, 295-308.

Garcia-Lastra, R., San-Miguel, B., Crespo, I., Jorquera, F., Alvarez, M., Gonzalez-Gallego, J., Tunon, M.J., 2010. **Signaling pathways involved in liver injury and regeneration in rabbit**

hemorrhagic disease, an animal model of virally-induced fulminant hepatic failure. Vet Res 41, 2.

Gertz, E.M., Agarwala, R., Mage, R.G., Schaffer, A.A., 2011. **Comparative analysis of genome sequences of the Th2 cytokine region of rabbit (Oryctolagus cuniculus) with those of nine different species**. Immunol Immunogenet Insights 3, 59-82.

Gertz, E.M., Schaffer, A.A., Agarwala, R., Bonnet-Garnier, A., Rogel-Gaillard, C., Hayes, H., Mage, R.G., 2013. **Accuracy and coverage assessment of Oryctolagus cuniculus (rabbit) genes encoding immunoglobulins in the whole genome sequence assembly (OryCun2.0) and localization of the IGH locus to chromosome 20**. Immunogenetics 65, 749-762.

Godornes, C., Leader, B.T., Molini, B.J., Centurion-Lara, A., Lukehart, S.A., 2007. **Quantitation of rabbit cytokine mRNA by real-time RT-PCR**. Cytokine 38, 1-7.

Hogg, P.J., 2003. **Disulfide bonds as switches for protein function**. Trends in biochemical sciences 28, 210-214.

Johnston, J.B., McFadden, G., 2004. **Technical knockout: understanding poxvirus pathogenesis by selectively deleting viral immunomodulatory genes**. Cellular microbiology 6, 695-705.

Kamaruzaman, N.A., Kardia, E., Kamaldin, N., Latahir, A.Z., Yahaya, B.H., 2013. **The rabbit as a model for studying lung disease and stem cell therapy**. Biomed Res Int 2013, 691830.

Keir, S., Page, C., 2008. **The rabbit as a model to study asthma and other lung diseases**. Pulm Pharmacol Ther 21, 721-730.

Keren, H., Lev-Maor, G., Ast, G., 2010. **Alternative splicing and evolution: diversification, exon definition and function**. Nat Rev Genet 11, 345-355.

Kerr, P.J., Liu, J., Cattadori, I., Ghedin, E., Read, A.F., Holmes, E.C., 2015. **Myxoma virus and the Leporipoxviruses: an evolutionary paradigm**. Viruses 7, 1020-1061.

Kerr, P.J., Perkins, H.D., Inglis, B., Stagg, R., McLaughlin, E., Collins, S.V., Van Leeuwen, B.H., 2004. **Expression of rabbit IL-4 by recombinant myxoma viruses enhances virulence and overcomes genetic resistance to myxomatosis**. Virology 324, 117-128.

Knell, R.J., 2004. **Syphilis in renaissance Europe: rapid evolution of an introduced sexually transmitted disease?** Proc Biol Sci 271 Suppl 4, S174-176.

Kornblihtt, A.R., Schor, I.E., Allo, M., Dujardin, G., Petrillo, E., Munoz, M.J., 2013. **Alternative splicing: a pivotal step between eukaryotic transcription and translation**. Nat Rev Mol Cell Biol 14, 153-165.

Kotlarz, D., Zietara, N., Uzel, G., Weidemann, T., Braun, C.J., Diestelhorst, J., Krawitz, P.M., Robinson, P.N., Hecht, J., Puchalka, J., Gertz, E.M., Schaffer, A.A., Lawrence, M.G., Kardava, L., Pfeifer, D., Baumann, U., Pfister, E.D., Hanson, E.P., Schambach, A., Jacobs, R., Kreipe, H., Moir, S., Milner, J.D., Schwille, P., Mundlos, S., Klein, C., 2013. **Loss-of-function mutations in the IL-21 receptor gene cause a primary immunodeficiency syndrome**. J Exp Med 210, 433-443.

Kunkel, T.A., 1985. **Rapid and efficient site-specific mutagenesis without phenotypic selection**. Proc Natl Acad Sci U S A 82, 488-492.

Lavinder, J.J., Hoi, K.H., Reddy, S.T., Wine, Y., Georgiou, G., 2014. **Systematic characterization and comparative analysis of the rabbit immunoglobulin repertoire**. PLoS One 9, e101322.

Lee, Y.C., Lee, R.T., 1995. **Carbohydrate-Protein interction: Basis of Glycobiology**. Acc. Chem. Res. 28, 321-327.

Li, X.Q., Zhang, T., Donnelly, D., 2011. **Selective loss of cysteine residues and disulphide bonds in a potato proteinase inhibitor II family**. PLoS One 6, e18615.

Liu, J., Wennier, S., Reinhard, M., Roy, E., MacNeill, A., McFadden, G., 2009. **Myxoma virus expressing interleukin-15 fails to cause lethal myxomatosis in European rabbits**. J Virol 83, 5933-5938.

Mage, R.G., Lanning, D., Knight, K.L., 2006. **B cell and antibody repertoire development in rabbits: the requirement of gut-associated lymphoid tissues**. Dev Comp Immunol 30, 137-153.

Mage, R.G., Mage, G.G., 2012. **Sequence of Rabbit (Oryctolagus cuniculus) DNA from the OryCun2.0 Donor does not Confirm a Frameshift in Exon 2 of IL4**. Immunology and Immunogenetics Insights 2012, 1-5.

Marques, R.M., Costa, E.S.A., Aguas, A.P., Teixeira, L., Ferreira, P.G., 2012. **Early inflammatory response of young rabbits attending natural resistance to calicivirus (RHDV) infection**. Veterinary immunology and immunopathology 150, 181-188.

Marques, R.M., Teixeira, L., Aguas, A.P., Ribeiro, J.C., Costa-e-Silva, A., Ferreira, P.G., 2014. **Immunosuppression abrogates resistance of young rabbits to Rabbit Haemorrhagic Disease (RHD)**. Vet Res 45, 14.

Marth, J.D., Grewal, P.K., 2008. **Mammalian glycosylation in immunity**. Nat Rev Immunol 8, 874-887.

Martinez, N.M., Lynch, K.W., 2013. **Control of alternative splicing in immune responses: many regulators, many predictions, much still to learn**. Immunol Rev 253, 216-236.

Metcalfe, C., Cresswell, P., Barclay, A.N., 2012. **Interleukin-2 signalling is modulated by a labile disulfide bond in the CD132 chain of its receptor**. Open Biol 2, 110036.

Metzger, K.J., Thomas, M.A., 2010. **Evidence of positive selection at codon sites localized in extracellular domains of mammalian CC motif chemokine receptor proteins**. BMC Evol Biol 10, 139.

Moremen, K.W., Tiemeyer, M., Nairn, A.V., 2012. **Vertebrate protein glycosylation: diversity, synthesis and function**. Nat Rev Mol Cell Biol 13, 448-462.

Mukherjee, S., Sarkar-Roy, N., Wagener, D.K., Majumder, P.P., 2009. **Signatures of natural selection are not uniform across genes of innate immune system, but purifying selection is the dominant signature**. Proc Natl Acad Sci U S A 106, 7073-7078.

Mullane, K., Williams, M., 2014. **Animal models of asthma: reprise or reboot?** Biochemical pharmacology 87, 131-139.

Niemela, J.E., Puck, J.M., Fischer, R.E., Fleisher, T.A., Hsu, A.P., 2000. **Efficient detection of thirty-seven new IL2RG mutations in human X-linked severe combined immunodeficiency**. Clin Immunol 95, 33-38.

Ochs, H.D., Smith, C.I.E., Puck, J., 2014. **Primary immunodeficiency diseases: a molecular and genetic approach**, 3rd ed. Oxford University Press, Oxford, New York.

Ortis, F., Naamane, N., Flamez, D., Ladriere, L., Moore, F., Cunha, D.A., Colli, M.L., Thykjaer, T., Thorsen, K., Orntoft, T.F., Eizirik, D.L., 2010. **Cytokines interleukin-1beta and tumor necrosis factor-alpha regulate different transcriptional and alternative splicing networks in primary beta-cells**. Diabetes 59, 358-374.

Owen, J.A., Punt, J., Stranford, S.A., Jones, P.P., Kuby, J., 2013. **Kuby immunology**, 7th ed. W.H. Freeman, New York.

Pasteur, L., 1885. **Méthode pour prévenir la rage après morsure**. Comptes rendus hebdomadaires des séances de l'Académie des sciences (Paris) 101, 765-774.

Peng, C., Haller, S.L., Rahman, M.M., McFadden, G., Rothenburg, S., 2016. **Myxoma virus M156 is a specific inhibitor of rabbit PKR but contains a loss-of-function mutation in Australian virus isolates**. Proc Natl Acad Sci U S A 113, 3855-3860.

Perkins, H.D., van Leeuwen, B.H., Hardy, C.M., Kerr, P.J., 2000. **The complete cDNA sequences of IL-2, IL-4, IL-6 AND IL-10 from the European rabbit (Oryctolagus cuniculus)**. Cytokine 12, 555-565.

Pinheiro, A., Lanning, D., Alves, P.C., Mage, R.G., Knight, K.L., van der Loo, W., Esteves, P.J., 2011. **Molecular bases of genetic diversity and evolution of the immunoglobulin heavy chain variable region (IGHV) gene locus in leporids**. Immunogenetics 63, 397-408.

Pinheiro, A., Neves, F., Lemos de Matos, A., Abrantes, J., van der Loo, W., Mage, R., Esteves, P.J., 2016. **An overview of the lagomorph immune system and its genetic diversity**. Immunogenetics 68, 83-107.

Pinheiro, A., Woof, J.M., Abi-Rached, L., Parham, P., Esteves, P.J., 2013. **Computational analyses of an evolutionary arms race between mammalian immunity mediated by immunoglobulin A and its subversion by bacterial pathogens**. PLoS One 8, e73934.

Pinheiro, A., Woof, J.M., Almeida, T., Abrantes, J., Alves, P.C., Gortazar, C., Esteves, P.J., 2014. **Leporid immunoglobulin G shows evidence of strong selective pressure on the hinge and CH3 domains**. Open Biol 4, 140088.

Popkov, M., Mage, R.G., Alexander, C.B., Thundivalappil, S., Barbas, C.F., 3rd, Rader, C., 2003. **Rabbit immune repertoires as sources for therapeutic monoclonal antibodies: the impact of kappa allotype-correlated variation in cysteine content on antibody libraries selected by phage display**. J Mol Biol 325, 325-335.

Quintana-Murci, L., Clark, A.G., 2013. **Population genetic tools for dissecting innate immunity in humans**. Nat Rev Immunol 13, 280-293.

Ros, F., Puels, J., Reichenberger, N., van Schooten, W., Buelow, R., Platzer, J., 2004. **Sequence analysis of 0.5 Mb of the rabbit germline immunoglobulin heavy chain locus**. Gene 330, 49-59.

Ros, F., Reichenberger, N., Dragicevic, T., van Schooten, W.C., Buelow, R., Platzer, J., 2005. **Sequence analysis of 0.4 megabases of the rabbit germline immunoglobulin kappa1 light chain locus**. Animal genetics 36, 51-57.

Rosenzweig, S.D., Holland, S.M., 2011. **Recent insights into the pathobiology of innate immune deficiencies**. Curr Allergy Asthma Rep 11, 369-377.

Sadler, A.J., Williams, B.R., 2007. **Structure and function of the protein kinase R**. Curr Top Microbiol Immunol 316, 253-292.

Sahoo, A., Im, S.H., 2010. **Interleukin and interleukin receptor diversity: role of alternative splicing**. International reviews of immunology 29, 77-109.

Samson, M., LaRosa, G., Libert, F., Paindavoine, P., Detheux, M., Vassart, G., Parmentier, M., 1997. **The second extracellular loop of CCR5 is the major determinant of ligand specificity**. J Biol Chem 272, 24934-24941.

San-Miguel, B., Crespo, I., Vallejo, D., Alvarez, M., Prieto, J., Gonzalez-Gallego, J., Tunon, M.J., 2014. **Melatonin modulates the autophagic response in acute liver failure induced by the rabbit hemorrhagic disease virus**. J Pineal Res 56, 313-321.

Scapin, G., 2006. **Structural biology and drug discovery**. Curr Pharm Des 12, 2087-2097.

Schnupf, P., Sansonetti, P.J., 2012. **Quantitative RT-PCR profiling of the rabbit immune response: assessment of acute Shigella flexneri infection**. PLoS One 7, e36446.

Seok, J., Warren, H.S., Cuenca, A.G., Mindrinos, M.N., Baker, H.V., Xu, W., Richards, D.R., McDonald-Smith, G.P., Gao, H., Hennessy, L., Finnerty, C.C., Lopez, C.M., Honari, S., Moore, E.E., Minei, J.P., Cuschieri, J., Bankey, P.E., Johnson, J.L., Sperry, J., Nathens, A.B., Billiar, T.R., West, M.A., Jeschke, M.G., Klein, M.B., Gamelli, R.L., Gibran, N.S., Brownstein, B.H., Miller-Graziano, C., Calvano, S.E., Mason, P.H., Cobb, J.P., Rahme, L.G., Lowry, S.F., Maier, R.V., Moldawer, L.L., Herndon, D.N., Davis, R.W., Xiao, W., Tompkins, R.G., 2013. **Genomic responses in mouse models poorly mimic human inflammatory diseases**. Proc Natl Acad Sci U S A 110, 3507-3512.

Shakola, F., Suri, P., Ruggiu, M., 2015. **Splicing Regulation of Pro-Inflammatory Cytokines and Chemokines: At the Interface of the Neuroendocrine and Immune Systems**. Biomolecules 5, 2073-2100.

Shay, T., Lederer, J.A., Benoist, C., 2015. **Genomic responses to inflammation in mouse models mimic humans: We concur, apples to oranges comparisons won't do**. Proc Natl Acad Sci U S A 112, E346.

Siewe, B.T., Kalis, S.L., Esteves, P.J., Zhou, T., Knight, K.L., 2010. **A novel functional rabbit IL-7 isoform**. Dev Comp Immunol 34, 828-836.

Skyberg, J.A., Rollins, M.F., Samuel, J.W., Sutherland, M.D., Belisle, J.T., Pascual, D.W., 2013. **Interleukin-17 protects against the Francisella tularensis live vaccine strain but not against a virulent F. tularensis type A strain**. Infection and immunity 81, 3099-3105.

Spiesschaert, B., McFadden, G., Hermans, K., Nauwynck, H., Van de Walle, G.R., 2011. **The current status and future directions of myxoma virus, a master in immune evasion**. Vet Res 42, 76.

Stanford, M.M., McFadden, G., 2005. **The 'supervirus'? Lessons from IL-4-expressing poxviruses**. Trends in immunology 26, 339-345.

Takao, K., Miyakawa, T., 2015. **Genomic responses in mouse models greatly mimic human inflammatory diseases**. Proc Natl Acad Sci U S A 112, 1167-1172.

Teixeira, L., Marques, R.M., Aguas, A.P., Ferreira, P.G., 2012. **Regulatory T cells are decreased in acute RHDV lethal infection of adult rabbits**. Veterinary immunology and immunopathology 148, 343-347.

Thangudu, R.R., Manoharan, M., Srinivasan, N., Cadet, F., Sowdhamini, R., Offmann, B., 2008. **Analysis on conservation of disulphide bonds and their structural features in homologous protein domain families**. BMC Struct Biol 8, 55.

Tosic, V., Thomas, D.L., Kranz, D.M., Liu, J., McFadden, G., Shisler, J.L., MacNeill, A.L., Roy, E.J., 2014. **Myxoma virus expressing a fusion protein of interleukin-15 (IL15) and IL15 receptor alpha has enhanced antitumor activity**. PLoS One 9, e109801.

Trzeciak-Ryczek, A., Tokarz-Deptuła, B., Deptuła, W., 2016. **Expression of IL-1β, IL-2, IL-10, TNF-β and GM-CSF in peripheral blood leukocytes of rabbits experimentally infected with rabbit haemorrhagic disease virus** Veterinary Microbiology 186, 71-81.

Tunon, M.J., San Miguel, B., Crespo, I., Jorquera, F., Santamaria, E., Alvarez, M., Prieto, J., Gonzalez-Gallego, J., 2011a. **Melatonin attenuates apoptotic liver damage in fulminant hepatic failure induced by the rabbit hemorrhagic disease virus**. J Pineal Res 50, 38-45.

Tunon, M.J., San Miguel, B., Crespo, I., Riezu-Boj, J.I., Larrea, E., Alvarez, M., Gonzalez, I., Bustos, M., Gonzalez-Gallego, J., Prieto, J., 2011b. **Cardiotrophin-1 promotes a high survival rate in rabbits with lethal fulminant hepatitis of viral origin**. J Virol 85, 13124-13132.

Vallejo, D., Crespo, I., San-Miguel, B., Alvarez, M., Prieto, J., Tunon, M.J., Gonzalez-Gallego, J., 2014. **Autophagic response in the Rabbit Hemorrhagic Disease, an animal model of virally-induced fulminant hepatic failure**. Vet Res 45, 15.

Vallender, E.J., Lahn, B.T., 2004. **Positive selection on the human genome**. Human molecular genetics 13 Spec No 2, R245-254.

van der Loo, W., Afonso, S., de Matos, A.L., Abrantes, J., Esteves, P.J., 2012. **Pseudogenization of the MCP-2/CCL8 chemokine gene in European rabbit (genus Oryctolagus), but not in species of Cottontail rabbit (Sylvilagus) and Hare (Lepus)**. BMC Genet 13, 72.

van der Loo, W., Magalhaes, M.J., de Matos, A.L., Abrantes, J., Yamada, F., Esteves, P.J., 2016. **Adaptive Gene Loss? Tracing Back the Pseudogenization of the Rabbit CCL8 Chemokine**. J Mol Evol 83, 12-25.

van Kooyk, Y., Rabinovich, G.A., 2008. **Protein-glycan interactions in the control of innate and adaptive immune responses**. Nat Immunol 9, 593-601.

Vande Walle, L., Lamkanfi, M., 2011. **Inflammasomes: caspase-1-activating platforms with critical roles in host defense**. Frontiers in microbiology 2, 3.

Varki, A., Lowe, J.B., 2009. **Biological role of glycans**, in: Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E. (Eds.), Essentials of Glycobyology, 2nd ed. Cold Spring Harbor, New York.

Vaure, C., Liu, Y., 2014. **A comparative review of toll-like receptor 4 expression and functionality in different animal species**. Front Immunol 5, 316.

Vogt, G., Chapgier, A., Yang, K., Chuzhanova, N., Feinberg, J., Fieschi, C., Boisson-Dupuis, S., Alcais, A., Filipe-Santos, O., Bustamante, J., de Beaucoudrey, L., Al-Mohsen, I., Al-Hajjar, S., Al-Ghonaium, A., Adimi, P., Mirsaeidi, M., Khalilzadeh, S., Rosenzweig, S., de la Calle Martin, O., Bauer, T.R., Puck, J.M., Ochs, H.D., Furthner, D., Engelhorn, C., Belohradsky, B., Mansouri, D., Holland, S.M., Schreiber, R.D., Abel, L., Cooper, D.N., Soudais, C., Casanova, J.L., 2005. **Gains of glycosylation comprise an unexpectedly large group of pathogenic mutations**. Nat Genet 37, 692-700.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., Burge, C.B., 2008. **Alternative isoform regulation in human tissue transcriptomes**. Nature 456, 470-476.

Wang, X., Rickert, M., Garcia, K.C., 2005. **Structure of the quaternary complex of interleukin-2 with its alpha, beta, and gammac receptors**. Science 310, 1159-1163.

Webb, D.R., 2014. **Animal models of human disease: inflammation**. Biochemical pharmacology 87, 121-130.

Wolfert, M.A., Boons, G.J., 2013. **Adaptive immune activation: glycosylation does matter**. Nat Chem Biol 9, 776-784.

Yabas, M., Elliott, H., Hoyne, G.F., 2015. **The Role of Alternative Splicing in the Control of Immune Homeostasis and Cellular Differentiation**. Int J Mol Sci 17.

Zlotnik, A., Yoshie, O., 2012. **The chemokine superfamily revisited**. Immunity 36, 705-716.