

# Maximum-likelihood approaches reveal signatures of positive selection in IL genes in mammals

Fabiana Neves<sup>1,2</sup>, Joana Abrantes<sup>1,3</sup>, John W Steinke<sup>4</sup>  
and Pedro J Esteves<sup>1,5</sup>

Innate Immunity  
2014, Vol. 20(2) 184–191  
© The Author(s) 2013  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1753425913486687  
ini.sagepub.com



## Abstract

ILs are part of the immune system and are involved in multiple biological activities. ILs have been shown to evolve 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: IL-1A and B; IL-2, IL-4 to IL-10, IL-12A and B; IL-14 to IL-17A and C; IL-18, IL-20 to IL-22, IL-25, IL-26, IL-27B, IL-31, IL-34, IL-36A; and G. Codons under positive selection varied between 1 and 15. No evidence of positive selection was detected in IL-13; IL-17B and F; IL-19, IL-23, IL-24, IL-27A; or IL-29. 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, which should be pursued by using mutagenesis and crystallographic approaches.

## Keywords

ILs, IL receptors, maximum likelihood, positive selection

Date received: 15 October 2012; revised: 8 January 2013; accepted: 27 January 2013

## 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, cks 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.<sup>1–6</sup> When cks are secreted by leukocytes and act on other cells they are called IL.<sup>3,7–10</sup> ILs are polypeptides of low molecular mass involved in several biological activities, including immunity, inflammation, inflammatory diseases, hematopoiesis, oncogenesis, neurogenesis and fertility, among many others. Each IL is normally involved in a multiple biological processes.<sup>9,11–13</sup> Currently, 37 ILs 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.<sup>8,9,14–20</sup> 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 ILs, which are able to transmit intracellular signals through different signaling pathways.<sup>1,11,13,21</sup> Sometimes, for optimal function, interaction between complementary

<sup>1</sup>CIBIO/UP – Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Porto, Portugal

<sup>2</sup>UMIB/UP – Unidade Multidisciplinar de Investigação Biomédica, Universidade do Porto, Porto, Portugal

<sup>3</sup>INSERM, U892, Université de Nantes, Nantes, France

<sup>4</sup>University of Virginia, Charlottesville, VA, USA

<sup>5</sup>CITS – Centro de Investigação em Tecnologias de Saúde, CESPU, Gandra, Portugal

## Corresponding author:

Pedro J. Esteves, CIBIO-UP, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas, no. 7, 4485-661 Vairão, Portugal.  
Email: pjesteves@cibio.up.pt

ILs is necessary.<sup>22</sup> These interactions can be synergistic, additive or antagonistic, and exhibit both negative and positive regulatory effects.<sup>9,11,22–24</sup>

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.<sup>5,25</sup> ILs such as IL-2, IL-3, IL-4, IL-5, IL-13, IL-23A, IL-28A, IL-28B and IL-29<sup>1,2,4,26–31</sup> 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,<sup>2,4,27,29–31</sup> with the exception of IL-4<sup>26</sup> in which 15 residues were detected as being under positive selection and located on sites responsible for binding to receptors. With these results in mind, we have extended the search for signatures of positive selection to include mammalian ILs 1–37 by using different maximum-likelihood approaches (ML).

## 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 to 47 (some of the IL genes were not found in all mammalian genomes) and included representatives of most mammalian 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 Appendix 1 (Supplementary material). Each of the 37 ILs were aligned using ClustalW, as implemented in the software program BioEdit version 7.1.3,<sup>32</sup> and adjusted manually.

### Codon-based analyses of positive selection

According to Poon and collaborators,<sup>6</sup> and in order to reliably identify codons under positive selection, only ILs that were represented in at least 10 species were analyzed. Hence, IL-17D, IL-28A and B, IL-32, IL-36B and IL-37 were not considered. IL-3 and IL-33 were also excluded as the alignments produced were not reliable and prone to affect the detection of positive selection by leading to false predictions.<sup>33</sup> Signatures of positive selection are inferred if the ratio of non-synonymous substitutions per non-synonymous 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 ( $\omega$ ) were compared using two ML frameworks, the HyPhy package implemented in the Data Monkey Web Server (<http://www.datamonkey.org/>)<sup>6</sup> and CODEML implemented in PAML version 4,<sup>34,35</sup> 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. 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.<sup>36–38</sup>

In CODEML two opposing models, M7 and M8, were compared using likelihood ratio tests. While M7 assumes that  $\omega$  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  $\omega$  ratio freely estimated from the data allowing positive selection. M7 versus M8 were compared by computing twice the difference in the natural logs of the likelihoods ( $2\Delta\ln L$ ). 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.<sup>34,35,39–41</sup> For each gene, a neighbor-joining tree was constructed in MEGA5<sup>42</sup> 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,<sup>43,44</sup> only sites detected to be under positive selection in more than one ML method were considered.

## Results

Positive selection was found in 28 of the 46 ILs: IL-1 (A and B), IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12 (A and B), IL-14, IL-15, IL-16, IL-17 (A and C), IL-18, IL-20, IL-21, IL-22, IL-25, IL-26, IL-27B, IL-31, IL-34 and IL-36 (A and G), as shown in Table 1. Codons under positive selection detected by

**Table 1.** Phylogenetic tests of positive selection.<sup>a</sup>

Gene	No. of aa residues	No. of species	ln L M7	ln L M8	2Δln L	PAML	SLAC <sup>b</sup>	FEL <sup>c</sup>	REL <sup>d</sup>	Total no. of sites	% of sites
IL-1 A	825	33	-10,987, 711	-10,939, 544	96,335*	<u>104, 122, 132, 147, 148, 173, 174, 175, 178, 180, 190, 227, 230, 248, 258</u>	<u>173, 174</u>	<u>66, 173, 174, 180, 188, 215, 227, 230, 265</u>	<u>54, 99, 104, 116, 122, 125, 127, 129, 132, 139, 158, 162, 173, 174, 175, 179, 180, 215, 227, 230, 248, 250</u>	11	1.333
IL-1 B	801	37	-9578, 964	-9565, 064	27,800*	79, 150, 205	<u>8, 84</u>	<u>8, 36, 71, 84, 93, 96, 104, 129, 153</u>	<u>84</u>	2	0.25
IL-2	513	47	-5429, 515	-5417, 492	24,045*	<u>100, 101, 124</u>	-	<u>28, 39, 42, 58, 69</u>	<u>43, 46, 66, 68, 91, 98, 100, 101, 102, 115, 122, 124, 152</u>	3	0.585
IL-4	465	44	-6436, 977	-6424, 860	24,235*	<u>84, 105, 106, 120</u>	<u>84</u>	<u>5, 59, 105, 120, 128, 147</u>	<u>84, 105, 106, 120</u>	4	0.86
IL-5	408	36	-4996, 583	-4994, 044	5078	-	-	<u>4, 93, 96, 100</u>	<u>100</u>	1	0.245
IL-6	657	42	-9824, 485	-9805, 607	37,755*	<u>46, 55, 62, 65, 75, 120, 141, 145, 149, 158, 161, 162, 177</u>	<u>8, 145</u>	<u>8, 162, 168, 174, 193</u>	<u>30, 38, 46, 55, 56, 62, 75, 76, 96, 113, 120, 141, 145, 149, 153, 158, 161, 162, 168, 171, 173, 174, 177</u>	15	2.283
IL-7	540	21	-3887, 223	-3861, 408	51,630*	<u>47, 51, 54, 60, 92, 102, 113, 126, 153, 172, 174</u>	<u>92, 144</u>	<u>38, 43, 51, 53, 55, 92, 105, 106, 144</u>	<u>30, 38, 46, 47, 51, 53, 54, 55, 92, 101, 105, 106, 113, 143, 144</u>	11	2.037
IL-8	297	33	-3256, 753	-3252, 791	7924*	<u>71</u>	-	<u>22, 30</u>	<u>22, 30, 64, 71, 74</u>	3	1010
IL-9	423	20	-4237, 433	-4217, 874	39,118*	<u>5, 24, 25, 84, 86</u>	<u>84</u>	<u>79, 84, 106</u>	<u>5, 20, 22, 24, 25, 28, 31, 35, 49, 79, 82, 84, 86, 90, 98, 106, 131, 136, 140</u>	7	1.655
IL-10	534	44	-6188, 435	-6184, 527	7815*	<u>26</u>	<u>18</u>	<u>18, 63, 71, 121</u>	<u>18, 20, 26, 35, 63, 68, 71, 121</u>	5	0.936
IL-12A	669	27	-5853, 066	-5847, 159	11,813*	<u>161</u>	-	<u>24, 35, 39, 91, 154</u>	<u>39, 161, 172</u>	2	0.299
IL-12B	1032	33	-9173, 438	-9170, 760	5357	-	<u>83</u>	<u>83, 147</u>	<u>21, 83, 182, 185, 188, 216, 289, 302, 306</u>	1	0.097
IL-14	1698	18	-8069, 857	-8069, 766	0.183	-	-	<u>501</u>	<u>501</u>	1	0.059

(continued)

**Table 1.** Continued.

Gene	No. of aa residues	No. of species	ln L M7	ln L M8	2Δln L	PAML	SLAC <sup>b</sup>	FEL <sup>c</sup>	REL <sup>d</sup>	Total no. of sites	% of sites
IL-15	483	29	-4003, 633	-3987, 300	32,666*	<u>51, 105, 106, 108</u>	-	4, <u>59, 116</u>	<u>51, 52, 55, 59, 62, 96, 105, 106, 108, 116, 152</u>	6	1.242
IL-16	4095	17	-27,560, 331	-27,552, 931	14,800*	622	335	<u>118, 188, 335, 633, 708, 734, 745, 801, 930, 1075</u>	<u>118, 188, 335, 705, 745, 801, 930, 1075</u>	8	0.195
IL-17A	483	27	-5623, 567	-5618, 970	9193*	-	<u>18, 98</u>	<u>18, 98</u>	-	2	0.414
IL-17C	600	18	-4833, 872	-4829, 855	8034*	1	<u>11, 17</u>	<u>11, 17</u>	<u>11, 17, 55</u>	2	0.333
IL-18	588	24	-4474, 258	-4458, 874	30,768*	<u>4, 5, 29, 45, 75, 105, 176, 192</u>	<u>180</u>	<u>6, 32, 37, 45, 159, 180</u>	<u>4, 6, 45, 75, 82, 105, 180, 192</u>	7	1.19
IL-20	525	25	-4976, 365	-4974, 835	3060	-	-	<u>56, 62</u>	<u>56, 62</u>	2	0.381
IL-21	456	19	-3201, 764	-3191, 105	21,319*	<u>98, 119</u>	<u>86, 112</u>	<u>86, 112</u>	<u>86, 98, 112, 119</u>	4	0.877
IL-22	543	25	-5167, 215	-5159, 805	14,821*	86, 173	<u>53</u>	<u>8, 53, 84</u>	<u>3, 8, 53, 65, 84</u>	3	0.552
IL-25	510	20	-4452, 843	-4441, 182	23,323*	<u>1, 2, 137</u>	-	<u>53, 137</u>	<u>36, 40, 55</u>	1	0.196
IL-26	504	20	-3693, 866	-3687, 927	11,877*	2	2	2	-	1	0.198
IL-27B	708	16	-5613, 682	-5608, 707	9949*	-	<u>114</u>	<u>114, 139</u>	-	1	0.141
IL-31	489	14	-4093, 760	-4063, 715	60,091*	<u>31, 33, 38, 67, 103, 133,</u>	-	<u>30, 33, 38, 43, 56, 67, 141</u>	<u>31, 33, 38, 56, 66, 98, 133</u>	6	1.227
IL-34	717	17	-4918, 901	-4917, 438	2926	-	<u>134</u>	<u>134</u>	<u>134</u>	1	0.139
IL-36A	477	18	-4805, 698	-4766, 449	78,496*	<u>2, 4, 5, 8, 9, 12, 38, 133</u>	-	<u>8, 133</u>	<u>1, 2, 3, 5, 7, 8, 9, 12, 71, 133, 156</u>	6	1.258
IL-36G	471	21	-4966, 855	-4916, 882	99,946*	<u>2, 3, 5, 6, 104, 140</u>	<u>2, 6, 104</u>	<u>2, 6, 104, 141</u>	<u>2, 5, 6, 18, 26, 41, 52, 55, 69, 73, 91, 104, 128, 144, 148</u>	4	0.849

<sup>a</sup>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 disulfide bonds are boxed.

<sup>b,c</sup>Codons with *P*-values < 0.05.

<sup>d</sup>Codons with Bayes factor > 50.

\**P* < 0.05.

at least two ML approaches ranged between 1 and 15, with IL-6 being the IL with the highest percentage of positively selected codons (2.28% with 15 codons under positive selection) followed by IL-7 (2.04% and 11 codons under positive selection). The smallest percentages of positively selected codons were 0.06% in IL-14 and 0.10% in IL-12B, each with 1 codon. No positively selected codons were detected in IL-13, IL-17B and F, IL-19, IL-23, IL-24, IL-27A and IL-29.

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.<sup>33</sup> From our physiochemical study of the amino acids under positive selection (see Appendix 2 in Supplementary material) we observed that for the majority of codons under positive selection changes in charge and polarity occur. For example, in IL-36G, 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.

## Discussion

Mammalian genes involved in the immune response are among the most rapidly evolving genes<sup>25,45,46</sup> as they include proteins with biological activities designed to protect the host (antibacterial, antiviral, antifungal or antiparasitic). As 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 ILs with their receptors occurs are essential for their function.<sup>26,47</sup> To verify a cause-effect relationship between ILs and their receptors, we extended our positive selection analyses to IL receptors (ILR), in particular focusing on sites that interact with ILs at sites detected as being under positive selection (Table 2). From all ILs where positive selection was detected, only in IL-4 and IL-18 was that correlation found. For IL-4, and by focusing on the residues under positive selection known to interact with IL-4 receptors, mainly IL-4R $\alpha$  and IL-13R $\alpha$  (see Figure 1), our results are in agreement with previous studies.<sup>48–51</sup> Indeed, of the IL-4 residues detected under positive selection, Glu84 is located between sites of interaction with IL-4R $\alpha$  and IL-13R $\alpha$ , while Arg105 and Phe106 are located in interaction sites of IL-4 with IL-4R $\alpha$ . No codons under positive selection were detected at the sites of IL-4: $\gamma$ c interaction, as suggested by others.<sup>1,26</sup> Of the two residues of IL-4R $\alpha$  under positive selection, only Ser95 has a direct interaction with IL-4 (Asn113).

IL-4 positively selected sites Arg105 and Phe106 are located near sites where IL-4 interacts with IL-4R $\alpha$ .<sup>26,49,52</sup> For IL-13R $\alpha$ , none of the seven residues detected as under positive selection interacted with Glu84 of IL-4, but the positively selected Ile254 from IL-13R $\alpha$  interacted with IL-4. The sites detected are in accordance with those described by Koyanagi et al.,<sup>26</sup> but not with those detected by O'Connel et al.<sup>1</sup> One explanation for this discordance is the number of species used, which differs between studies. For IL-18, of the seven residues detected as under positive selection, only Leu45 is located near the site of interaction with IL-18R $\alpha$  (Lys44) and IL-18BP (Glu42). The signatures of positive selection in these sites might also be involved in modulating signaling intensity.<sup>26</sup>

In their evolution, pathogens have evolved an arsenal of immune-evading strategies, which include antagonists for the host immune-related proteins<sup>28</sup> 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 IL-4 have been suggested to be associated with escaping from pathogen-encoded antagonistic proteins.<sup>26</sup> BP are polypeptides that have the ability to interact with corresponding proteins and neutralize them. In the ILs studied, there are many examples of such proteins. Indeed, IL-18BP has the ability to bind to IL-18 with high affinity preventing IL-18–IL-18R interaction and, therefore, neutralizing their biological activities. Some authors have demonstrated that IL-18BP prevents LPS-induced IFN- $\gamma$  production.<sup>53–55</sup> In some models, it has been described that blocking IL-18 results in a reduction in disease severity in injuries related to IL-18 increased production, namely injuries of heart, kidneys, liver, arthritis, etc.<sup>56</sup> By studying IL-18, Leu45 which was detected as under positive selection is located close to the site of interaction of IL-18 (Glu42) with IL-18BP. In addition, when we studied IL-18BP, we detected that Cys131 is under positive selection next to the site of interaction of IL-18BP (Lys130) with IL-18.

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 examining Table 1 two of the genes with the largest number of positively selected residues are IL-1A and IL-6. IL-1A is the initial cytokine produced upon inflammasome activation, which directly leads to increased expression of IL-6.<sup>9</sup> 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, knowledge on the functional role of IL-1A and IL-6 derives mostly from



**Table 2.** Phylogenetic tests of positive selection for receptors and binding proteins.<sup>a</sup>

Gene	No. of aa residues	No. of Species	In L M7	In L M8	2Δln L	PAML	SLAC <sup>b</sup>	FEL <sup>c</sup>	REL <sup>d</sup>	Total no. of sites	% of sites
IL-4Rα	2529	10	-13,357, 029	-13346,832	20,393*	<u>95</u>	-	471, 479, 733, <u>764</u>	67, 68, 93, 95, 529, 590, 677, 689, <u>764</u>	2	0.079
IL-13Rα	1296	19	-9056, 699	-9038, 203	36,992*	<u>254, 276</u>	<u>254</u>	83, <u>196</u> , <u>225</u> , <u>237</u> , <u>254</u> , <u>259</u>	16, 19, 25, 31, 83, 171, <u>196</u> , <u>225</u> , 235, 237, 239, 242, <u>254</u> , <u>259</u> , 272, 274, <u>276</u> , <u>289</u> , 334, 335	7	0.540
IL-18BP	573	21	-5034, 333	-5016, 838	34,991*	<u>54, 60, 126, 131, 156, 176</u>	<u>11</u>	9, <u>11</u> , 24, 60, 61	11, 60, 63, <u>126</u> , <u>131</u> , <u>153</u> , <u>156</u>	5	0.873

<sup>a</sup>Codons identified by more than one ML method are underlined.<sup>b,c</sup>Codons with *P*-values < 0.05.<sup>d</sup>Codons with Bayes factor > 50.\**P* < 0.05.**Figure 1.** Positively selected sites in the three-dimensional structures of IL-4-IL-4Rα-IL-13Rα (marked in yellow). IL-4 appears in pink, IL-4Rα in blue and IL-13Rα in brown.

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 throughout the organism, interaction with cell surface receptors and subsequent biological activity.<sup>57–59</sup> This modification is an effective way to generate diversity and modulate protein properties due to inherent structural variations of the glycans.<sup>57,59,60</sup> For IL-1A, IL-9, IL-22, IL-25 and IL-31, residues under positive selection were found in close proximity to sites where *N*-glycosylation occurs (Table 1). Disulfide bonds play an important role in folding, stability and function of the protein, being, according to some authors,<sup>61,62</sup> associated with functional differentiation of the proteins. These bonds are thought to be well conserved in proteins.<sup>61</sup> Positive selection was detected at sites where disulfide bonds exist for IL-2 and IL-17A (Table 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.

## Conclusions

Positive selection may function to maintain a host response to pathogens, diseases and environmental

conditions. Here, we have detected positive selection in 28 ILs 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 IL-s 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 full comprehension of the role of the observed variation in mammalian ILs.

### Funding

The Portuguese Foundation for Science and Technology (FCT-Portugal) supported the doctoral fellowship of Fabiana Neves (SFRH/BD/81916/2011), the post-doctoral fellowships of Joana Abrantes (SFRH/BPD/73512/2010) and Pedro J. Esteves (SPRH/BPD/27021/2006). This study was performed under the project conceded by FCT (PTDC/BIA-BEC/103158/2008).

### References

- O'Connell MJ and McInerney JO. Gamma chain receptor interleukins: evidence for positive selection driving the evolution of cell-to-cell communicators in the mammalian immune system. *J Mol Evol* 2005; 61: 608–619.
- Zelus D, Robinson-Rechavi M, Delacore M, et al. Fast evolution of interleukin-2 in mammals and positive selection in ruminants. *J Mol Evol* 2000; 51: 234–244.
- Kaiser P, Rothwell L, Avery S and Balu S. Evolution of the interleukins. *Develop Comp Immunol* 2004; 28: 375–394.
- Zhang J and Nei M. Positive selection in the evolution of mammalian interleukin-2 genes. *Mol Biol Evol* 2000; 17: 1413–1416.
- Brocker C, Thompson D, Matsumoto A, et al. Evolutionary divergence and functions of the human interleukin (IL) gene family. *Hum Genomics* 2010; 5: 30–55.
- Poon AF, Frost SD and Pond SL. Detecting signatures of selection from DNA sequences using Datamonkey. *Methods Mol Biol* 2009; 537: 163–183.
- Schrader JW. Interleukin is as interleukin does. *J Immunol Methods* 2003; 276: 1–3.
- Beadling C and Slifka MK. Regulation of innate and adaptive immune responses by the related cytokines IL-12, IL-23, and IL-27. *Arch Immunol Ther Exp (Warsz)* 2006; 54: 15–24.
- Akdis M, Burgler S, Cramer R, et al. Interleukins, from 1 to 37, and interferon-gamma: receptors, functions, and roles in diseases. *J Allergy Clin Immunol* 2011; 127: 701–721 e1–e70.
- Ozaki K and Leonard WJ. Cytokine and cytokine receptor pleiotropy and redundancy. *J Biol Chem* 2002; 277: 29355–29358.
- Afzal N, Tahir R and Jahan S. Cytokines: an ever expanding area. *Biol Biomed Rep* 2012; 2: 37–43.
- Heinrich PC, Behrmann I, Haan S, et al. Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem J* 2003; 374: 1–20.
- Ishihara K and Hirano T. Molecular basis of the cell specificity of cytokine action. *Biochim Biophys Acta* 2002; 1592: 281–296.
- Borish LC and Steinke JW. 2. Cytokines and chemokines. *J Allergy Clin Immunol* 2003; 111: S460–S475.
- Commins SP, Borish L and Steinke JW. Immunologic messenger molecules: cytokines, interferons, and chemokines. *J Allergy Clin Immunol* 2010; 125: S53–S72.
- Steinke JW and Borish L. 3. Cytokines and chemokines. *J Allergy Clin Immunol* 2006; 117: S441–S445.
- Sabat R. IL-10 family of cytokines. *Cytokine Growth Factor Rev* 2010; 21: 315–324.
- Krause CD and Pestka S. Evolution of the Class 2 cytokines and receptors, and discovery of new friends and relatives. *Pharmacol Ther* 2005; 106: 299–346.
- Leonard WJ and Lin JX. Cytokine receptor signaling pathways. *J Clin Immunol* 2000; 105: 877–888.
- Secombes CJ, Wang T and Bird S. The interleukins of fish. *Develop Immunol* 2011; 35: 1336–1345.
- Fumagalli M, Pozzoli U, Cagliani R, et al. Parasites represent a major selective force for interleukins genes and shape the genetic predisposition to autoimmune conditions. *J Exp Med* 2009; 206: 1395–1408.
- Commins S, Steinke JW and Borish L. 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* 2008; 121: 1108–1111.
- Feghali CA and Wright TM. Cytokines in acute and chronic inflammation. *Front Biosci* 1997; 2: d12–d26.
- Haddad JJ. Cytokines and related receptor-mediated signaling pathways. *Biochem Biophys Res Commun* 2002; 297: 700–713.
- Barreiro LB and Quintana-Murci L. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat Rev Genet* 2010; 11: 17–30.
- Koyanagi M, Kerns JA, Chung L, et al. Diversifying selection and functional analysis of interleukin-4 suggests antagonism-driven evolution at receptor-binding interfaces. *BMC Evol Biol* 2010; 10: 223.
- Makalowski W and Boguski MS. Synonymous and nonsynonymous substitution distances are correlated in mouse and rat genes. *J Mol Evol* 1998; 47: 119–121.
- PIL-lai MR and Bix M. Evolution of IL4 and pathogen antagonism. *Growth Factors* 2011; 29: 153–160.
- Manry J, Laval G, Patin E, et al. Evolutionary genetic dissection of human interferons. *J Exp Med* 2011; 208: 2747–2759.
- Tindall EA and Hayes VM. Comprehensive sequence analysis of the human IL23A gene defines new variation content and high rate of evolutionary conservation. *DNA Res* 2010; 17: 117–122.
- Zhou G, Zhai Y, Dong X, et al. Haplotype structure and evidence for positive selection at the human IL13 locus. *Mol Biol Evol* 2004; 21: 29–35.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 1999; 41: 95–98.
- Majewski J and Ott J. Amino acid substitutions in the human genome: evolutionary implications of single nucleotide polymorphisms. *Gene* 2003; 305: 167–173.
- Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 1997; 13: 555–556.
- Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 2007; 24: 1586–1591.
- Kosakovsky Pond SL and Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* 2005; 22: 1208–1222.
- Pond SL and Frost SD. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 2005; 21: 2531–2533.
- Poon AF, Frost SD and Pond SL. Detecting signatures of selection from DNA sequences using Datamonkey. *Methods Mol Biol* 2009; 537: 163–183.
- Bielawski JP and Yang Z. Maximum likelihood methods for detecting adaptive evolution after gene duplication. *J Struct Funct Genomics* 2003; 3: 201–212.
- Yang Z. Inference of selection from multiple species alignments. *Curr Opin Genet Dev* 2002; 12: 688–694.

41. Yang Z, Nielsen R, Goldman N and Pedersen AM. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 2000; 155: 431–449.
42. Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011; 28: 2731–2739.
43. Areal H, Abrantes J and Esteves PJ. Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evol Biol* 2011; 11: 368.
44. Wlasiuk G and Nachman MW. Adaptation and constraint at Toll-like receptors in primates. *Mol Biol Evol* 2010; 27: 2172–2186.
45. Kosiol C, Vinar T, da Fonseca RR, et al. Patterns of positive selection in six Mammalian genomes. *PLoS Genet* 2008; 4: e1000144.
46. Ferrer-Admetlla A, Bosch E, Sikora M, et al. Balancing selection is the main force shaping the evolution of innate immunity genes. *J Immunol* 2008; 181: 1315–1322.
47. Smith KA. The structure of IL2 bound to the three chains of the IL2 receptor and how signaling occurs. *Med Immunol* 2006; 5: 3.
48. Hage T, Sebald W and Reinemer P. Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface. *Cell* 1999; 97: 271–281.
49. LaPorte SL, Juo ZS, Vaclavikova J, et al. Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system. *Cell* 2008; 132: 259–272.
50. Zarlenga DS, Dawson H, Kringel H, et al. 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. *Vet Immunol Immunopathol* 2004; 101: 223–234.
51. Zhang JL, Simeonowa I, Wang Y and Sebald W. The high-affinity interaction of human IL-4 and the receptor alpha chain is constituted by two independent binding clusters. *J Mol Biol* 2002; 315: 399–407.
52. Mueller TD, Zhang JL, Sebald W and Duschl A. Structure, binding, and antagonists in the IL-4/IL-13 receptor system. *Biochim Biophys Acta* 2002; 1592: 237–250.
53. Dinarello CA. Targeting interleukin 18 with interleukin 18 binding protein. *Ann Rheum Dis* 2000; 59(Suppl. 1): i17–i20.
54. He Z, Lu L, Altmann C, et al. Interleukin-18 binding protein transgenic mice are protected against ischemic acute kidney injury. *Am J Physiol Renal Physiol* 2008; 295: F1414–F1421.
55. Mallat Z, Corbaz A, Scoazec A, et al. Interleukin-18/interleukin-18 binding protein signaling modulates atherosclerotic lesion development and stability. *Circ Res* 2001; 89: E41–E45.
56. Dinarello CA. Interleukin 1 and interleukin 18 as mediators of inflammation and the aging process. *Am J Clin Nutr* 2006; 83: 447S–455S.
57. Shental-Bechor D and Levy Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc Natl Acad Sci USA* 2008; 105: 8256–8261.
58. Szabo TG, Palotai R, Antal P, et al. Critical role of glycosylation in determining the length and structure of T cell epitopes. *Immunome Res* 2009; 5: 4.
59. Waetzig GH, Chalaris A, Rosenstiel P, et al. N-linked glycosylation is essential for the stability but not the signaling function of the interleukin-6 signal transducer glycoprotein 130. *J Biol Chem* 2010; 285: 1781–1789.
60. Chamorey AL, Magne N, Pivot X and Milano G. Impact of glycosylation on the effect of cytokines. A special focus on oncology. *Eur Cytokine Netw* 2002; 13: 154–160.
61. Li XQ, Zhang T and Donnelly D. Selective loss of cysteine residues and disulphide bonds in a potato proteinase inhibitor II family. *PLoS One* 2011; 6: e18615.
62. Fass D. Disulfide bonding in protein biophysics. *Annu Rev Biophys* 2012; 41: 63–79.