Immunogenetics (2016) 68:719–731
DOI 10.1007/s00251-016-0922-1

ORIGINAL ARTICLE

Identification of interleukin genes in *Pogona vitticeps* using a de novo transcriptome assembly from RNA-seq data

Alexandra Livernois[1] · Kristine Hardy[2] · Renae Domaschenz[3] · Alexie Papanicolaou[4] · Arthur Georges[1] · Stephen D Sarre[1] · Sudha Rao[2] · Tariq Ezaz[1] · Janine E Deakin[1]

Received: 3 March 2016 / Accepted: 24 May 2016 / Published online: 2 June 2016

Abstract Interleukins are a group of cytokines with complex immunomodulatory functions that are important for regulating immunity in vertebrate species. Reptiles and mammals last shared a common ancestor more than 350 million years ago, so it is not surprising that low sequence identity has prevented divergent interleukin genes from being identified in the central bearded dragon lizard, *Pogona vitticeps*, in its genome assembly. To determine the complete nucleotide sequences of key interleukin genes, we constructed full-length transcripts, using the Trinity platform, from short paired-end read RNA sequences from stimulated spleen cells. De novo transcript reconstruction and analysis allowed us to identify interleukin genes that are missing from the published *P. vitticeps* assembly. Identification of key cytokines in *P. vitticeps* will provide insight into the essential molecular mechanisms and evolution of interleukin gene families and allow for characterization of the immune response in a lizard for comparison with mammals.

Electronic supplementary material The online version of this article (doi:10.1007/s00251-016-0922-1) contains supplementary material, which is available to authorized users.

✉ Janine E Deakin
janine.deakin@canberra.edu.au

[1] Institute for Applied Ecology, University of Canberra, Canberra, ACT 2617, Australia

[2] Discipline of Biomedical Sciences, Faculty of Education, Science, Technology and Mathematics, University of Canberra, Canberra, ACT 2617, Australia

[3] John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia

[4] The Hawkesbury Institute for the Environment, University of Western Sydney, Penrith, NSW 2751, Australia

Keywords *Pogona vitticeps* · De novo assembly · Cytokine · Interleukin

Introduction

Interleukins (ILs) are small glycoproteins that, as members of the large cytokine protein group, play a key role in modulating growth, differentiation and activation of T and B lymphocytes and haematopoietic stem cells during an immune response (Commins et al. 2010). They evoke a wide range of responses in cells and tissues more generally, initiated by binding to a high-affinity cell surface receptor (Brocker et al. 2010). ILs can act as inflammatory and anti-inflammatory molecules, and some can act as chemo-attractants for helper T cells. In general, ILs are important for responding to infection and are involved in a range of physiological disorders (reviewed in Brocker et al. 2010).

The ILs are classed as cytokines and divided into four major families. The IL-1-like cytokines which comprise 11 proteins (including IL-1α, IL-1β, IL-1ra [receptor antagonist], IL-18 and IL-33) (Dinarello and Wolff 1993) that are rich in β-strands; the class 1 helical cytokines (IL-4-like) which have a common structure of four tightly packed α-helices, or a 'four-helix bundle' arranged in an 'up-up-down-down' orientation (including IL-3, IL-4, IL-5, IL-13 and CSF2); the class II helical cytokines (IL-10, IL-19, IL-20, IL-22, IL-24 and IL-26) also have a 'bundled helix' structure, but of six or seven helices (Kotenko 2002); and the IL-17A homologues, which are distinct from the other IL groups, form a neurotrophin-like cysteine-knot fold (reviewed in Brocker et al. 2010). IL family members are often found tightly clustered in the genome. For example, nine of the IL-1 family members are found within 430 kbp on human chromosome 2q13 (whereas IL-18 is on chromosome 11 and IL-33 is on chromosome 9) (Dunn et al. 2001;

Gibson et al. 2014), and IL-2 and IL-21 are separated by only 244 kbp on human 4q26. The close proximity of IL family members suggests that clustered genes originated from a common ancestral gene that underwent gene duplications (reviewed in Brocker et al. 2010; Dunn et al. 2001; Gibson et al. 2014).

Although many ILs are the product of gene duplication events, they show relatively low sequence identity, which suggests rapid divergent evolution. This is likely the result of host-parasite co-evolution, where advantageous mutations are heavily favoured (Dunn et al. 2001). Furthermore, it has been hypothesized that only a few key residues are required to maintain the secondary structures required for IL function (Brocker et al. 2010).

ILs play a critical role in regulating and controlling the immune response in mammals and birds but it is unknown how this holds across vertebrates, especially non-endothermic ones. IL genes will likely be present across vertebrates because all vertebrates have the cells that produce them, which include T lymphocytes, innate immune cells (mast cells, eosinophils and basophils) and innate lymphoid cells (Eberl et al. 2015). Much of what we know about ILs comes from studies in mammals, and to a lesser extent, birds. We know comparably little about ILs in reptiles, or the reptilian immune system for that matter. Reptiles have unique features that make them an important model group for comparative study of IL gene content and immune system function. They have humoral and cellular immunity as well as innate immune systems and possess lymphoid tissues, but unlike mammals and birds, most reptiles range between being poikilothermic and ectothermic. Reptiles are imprecise facultative homeotherms when active, drawing upon heat sources and sinks in the environment to manipulate their body temperatures, and essentially poikilothermic when inactive. As immune responses vary with body temperature, reptiles are therefore likely to differ from mammals in how they allocate resources to their immune response (Zimmerman et al. 2010).

The recent sequencing and annotation of the genome of the central bearded dragon lizard, *Pogona vitticeps* (Georges et al. 2015), provides a new opportunity to study the immune system of a reptile and identify the IL gene content. A handful of interleukins have already been annotated in reptiles, most of them in the lizard, *Anolis carolinensis*, the Chinese softshell turtle, *Pelodiscus sinensis*, and the Chinese alligator, *Alligator sinensis*. However, the more divergent IL genes, such as *IL2*, which is important for the initiation and regulation of the immune response (Morgan et al. 1976), and *IL4*, *IL5* and *IL13*, which are essential in the differentiation of naïve CD4+ helper T (Th) cells, are not detected by automated annotation pipelines. Two divergent ILs have been predicted in reptiles; *IL2* was predicted in the *A. carolinensis* genome (AnoCar2.0) by Dijkstra et al. (Dijkstra et al. 2014), and *IL4* was annotated in *A. sinensis* during the genome sequencing project (GenBank AVPB00000000). The location of the predicted *IL4* gene in the *A. sinensis* genome is the same as in mammals, but it is also possible that this predicted gene is *IL13* since these two genes are adjacent. The identification of IL gene content across all vertebrates will give us a better understanding of the evolution of IL gene families and insight into when duplication events occurred (Dunn et al. 2001).

There has been a particular focus on chromatin changes in the coding and regulatory sequences of IL genes during an immune response in mammals. For example, Rao and colleagues demonstrated that increased *IL2* gene transcription in murine T cells is preceded by changes in the chromatin structure across the promoter region (Rao et al. 2003, 2001). Recent evidence has linked the expression of *IL8* in humans with a decrease in DNA methylation (Takahashi et al. 2015). Thus, identification of divergent IL genes in reptiles will allow us to study the relationship between changes in their chromatin structure and expression. Future comparison of the expression patterns and chromatin dynamics of ILs in mammals with those in *P. vitticeps* will be particularly interesting because of the non-endothermic nature of reptiles. This type of comparison will be insightful for the evolution of regulatory mechanisms in vertebrates and may increase our understanding of the vertebrate immune response more generally.

The published annotation of the *P. vitticeps* genome identified phylogenetically conserved interleukins, but because automatic annotation relies on homology, it is not surprising that highly divergent ILs were not identified (Brocker et al. 2010; Crawford et al. 2010). For example, amino acid identity between mammalian IL-4 molecules is only 42–57 %. To fully ascertain the IL gene family in the *P. vitticeps* genome, we used the Trinity version 2.1.1 (Grabherr et al. 2011) platform to perform de novo reconstruction of transcripts from RNA sequenced (RNA-seq) reads from phorbol-myristrate-acetate and ionomycin (PMA/I)-stimulated splenic cells. Stimulation of lymphocytes from *P. vitticeps* spleen was crucial to elicit expression of IL genes involved in the immune response, thereby enabling us to obtain their sequences from RNA. Here, we conducted the most extensive manual annotation to date of the IL gene family in lizards using the *P. vitticeps* genome, vertebrate orthologues and a combination of synteny and alignment analysis. As a result, 26 ILs have now been identified in *P. vitticeps*, the highest in any lizard species, eight of which we describe here for the first time in a lizard. As the community continues with the sequencing of other reptilian taxa, our data and methods will enable studying the evolutionary history of these key immunity proteins.

Methods

Wild-caught male *P. vitticeps* (Id #UC<Aus>:111880108089) was euthanized by injection of 0.1 mL sodium pentobarbitone (60 mg/mL) according to ethics approval CE-13-10 at the

University of Canberra. The spleen was immediately immersed in Dulbecco's modified Eagle's medium (DMEM) (Thermofisher, Australia) and placed on ice.

Splenic cell disassociation was performed through a cell strainer using the plunger of a 1-mL syringe as a pestle. Cells were spun (1500 rpm, 5 min) and splenocytes isolated using the Lympholyte M density gradient (Cedar Lane, Canada). After washing in DMEM, cells (2×10^6 in 4 mL of DMEM with foetal bovine serum (10 %, FBS) (Thermofisher)) were left unstimulated or stimulated with phorbol myristate acetate (PMA; 10 ng/mL) (Boehringer Mannheim, Germany) and ionomycin (I; 1 uM; Sigma-Aldrich) for 4 h at 28 °C.

RNA isolation

Total RNA was isolated from PMA/I-stimulated (ST) and non-stimulated (NS) splenic cells using the RNeasy Plus Mini Kit (Qiagen, Australia) according to manufacturer's instructions. The purified RNA was treated with DnaseI using the manufacturer's protocol (Ambion, USA). RNA quality was assessed using an Agilent 2100 Bioanalyzer at the Australian Cancer Research Foundation (ACRF) Biomolecular Resource Facility (BRF) at the Australian National University, Canberra, Australia.

Sequencing

NS and ST libraries were generated by the BRF using the paired-end TruSeq RNA library prep kit (Illumina) and sequenced on the Illumina HiSeq 2500 platform (paired-end, 100 bp reads).

Reads without a mate-pair were removed. Reads were adapter/quality-trimmed using Trimmomatic (v0.32 with options PE:2:30:10 LEADING:20 TRAILING:20 MINLEN = 40) (Bolger et al. 2014).

De novo assembly

Trimmed reads from NS and ST samples were pooled to increase the likelihood of detecting transcripts then assembled into transcripts using Trinity-RNASeq with default settings (Haas et al. 2013). Trinity de novo-assembled transcripts were aligned to the *P. vitticeps* genome using GMAP (Wu and Watanabe 2005) and visualized using the Integrative Genomics Viewer (IGV) (Robinson et al. 2011; Thorvaldsdottir et al. 2013).

To identify sequences previously undiscovered in the published genome annotation (Georges et al. 2015), we used the position of the IL gene of interest relative to more conserved genes within the genome (synteny). Sequences were extracted from de novo-assembled transcripts that aligned to the syntenic location of unidentified interleukins. All six translated frames were aligned using ClustalOmega (Sievers et al. 2011) to amino acid sequences of the interleukin of interest to determine identity and conservation of biologically significant sites for structure and function in their eutherian, avian and amphibian (where possible) orthologues.

When only one novel transcript was present at the syntenic location of two genes (which was the case for *IL1A*:*IL1B*, *IL4*:*IL13*, and *IL3*:*CSF2*), it was aligned to orthologues of both possible identities. Alignments were manually examined for conservation of motifs and altered if the ClustalOmega programme missed the motif. Accession numbers of sequences used in alignments are shown in figure legends. We checked for gaps in the genome by manual examination of the extracted sequence.

Since ILs are secreted proteins, the presence of a signal peptide in de novo transcripts was predicted by SignalP 4.1 (Petersen et al. 2011) and glycosylation sites were estimated with NetOGlyc 4.0 (http://www.cbs.dtu.dk/services/NetOGlyc/) and NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/). When the identity of interleukins was ambiguous, we attempted amino acid alignments with presumptive homologues from those taxa for which relevant sequence was available through GenBank. Where the alignments were successful, we constructed neighbour-joining trees with 1000 bootstrap repetitions using Geneious (Version R6) to identify genetic association of *P. vitticeps* sequence with target candidate genes and resolve the ambiguities.

Differential expression

For differential expression, we used the DEW pipeline (http://dx.doi.org/10.4225/08/54AB816944E07). Briefly, reads from each library were mapped to our assembly with Bowtie2 (Langmead and Salzberg 2012) and quantitated with eXpress (Roberts and Pachter 2013). EdgeR was used to perform trimmed mean normalization estimate fragments per kilobase per million (FPKM) reads (Robinson et al. 2010).

Results and discussion

We generated nearly 200 million reads from each of the ST and NS splenic cell libraries. These reads are available on the SRA database under the Bioproject identifier PRJNA299832. After filtering, 197,945,818 paired reads (98 % of input) from the NS library and 182,583,658 paired reads (98 % of input) from the ST library were assembled with Trinity to find novel transcripts, that is, transcripts not identified as ILs in the published annotation.

De novo transcript reconstruction

Overall, 242,519 contigs were assembled from the combined NS and ST reads. This is expected given the degree of nucleotide polymorphism in this species. The average contig length was 829.74 bp with 44.7 % GC content and an N50 of 1490 bp.

We searched for novel transcripts that could represent divergent ILs based on their relative location to highly conserved genes in mammals and birds and then used structural analyses to confirm their identities. Using this approach, we identified for the first time in a lizard *IL1B*, *IL3*/*CSF2*, *IL4*/*IL13*, *IL7*, *IL9*, *IL12A*, *IL27* and *IL31*. We also identified for the first time in *P. vitticeps IL2*, *IL21* and *IL15*, which have been predicted in the anole lizard (*A. carolinensis*) (Dijkstra et al. 2014). Additionally, we confirmed and corrected the annotations of 10 additional ILs in the published *P. vitticeps* genome annotation using our novel transcripts. This new information is available at https://genomics.canberra.edu.au/gbrowse/gbrowse/pogona_pvi1.1.

The importance of stimulating splenic cells from *P. vitticeps* to identify highly divergent ILs was illustrated by their differential expression in NS and ST libraries. All ILs were expressed at a higher level in ST splenic cells except for *IL7*, *IL15* and the novel transcript in the syntenic region of *IL5* (Fig. 1). Because of very low expression in NS splenic cells, without PMA/I stimulation, it is unlikely that we would have identified *IL2*, *IL21*, *IL3*/*CSF2*, *IL4*/*IL13*, *IL9* and *IL31* (Fig. 1). This is consistent, at least for *IL2*, with studies in human and mouse in which the expression of *IL2* is not detected until activation of T cells with stimuli such as PMA/I (Kumagai et al. 1987).

Manually curated interleukins

IL2 has proven difficult to isolate in non-eutherian vertebrates because of the larger evolutionary distance between clades than within eutherians (Harrison and Wedlock 2000) as well as the high rate of evolution of IL genes. *IL2* is related to *IL21* and *IL15*, and they likely evolved through duplication of an ancestral gene (Bird et al. 2005a). Failed efforts to isolate *IL2* in chicken and wallaby revolved around designing heterologous primers for RT-PCR based on eutherian sequences to clone genes (Harrison and Wedlock 2000; Zelus et al. 2000). *IL2* was eventually identified in chicken by functional screening (Beck 1998) and in wallaby by manually curating syntenic regions of interest of the wallaby genome (Wong et al. 2006).

We identified *IL2* in *P. vitticeps* by searching for novel transcripts in its syntenic region between annotated genes *ADAD1* and *CETN4* (Fig. 2). The amino acid sequence from the predicted *IL2* transcript was extracted and aligned to other *IL2* sequences (Fig. 3), which revealed low identity with human and chicken orthologues (28 %) (Table 1). We regard this as a valid identity because ILs generally share low identity across lineages due to their rapid evolutionary rates (Harrison and Wedlock 2000). For example, chicken and human *IL2* orthologues share only 24 % identity.

Importantly, *P. vitticeps IL2* contains several conserved residues vital for biological activity including amino acid aspartic acid in helix A and glutamine in helix D, which are crucial for binding to the *IL2* receptor in human (Fig. 3) (Wong et al. 2006). An essential structural disulphide bond between cysteines 58 and 105 (in human) is also conserved in *P. vitticeps*. This disulphide bond connects the second helix to the region between the third and fourth helices (Gaffen and Liu 2004). Secondary structure prediction using Jpred 4 (Drozdetskiy et al. 2015) shows that the *P. vitticeps IL2* molecule belongs to the type I cytokines, described as four alpha helical bundles (Supplemental Fig. 1). The *IL2*/*IL15* family signature in vertebrates, identified as 'L-X-C-X(3)-E-[LVI]-X(2)-[LV]' by Diaz-Rosales et al. (2009), is modified with the inclusion of lizards to '[LM]-X-C-X(3)-E-[LVI]-X(2)-[LVI]'. The 3′UTR of the *IL2* transcript in *P. vitticeps* contains two instability motifs (ATTTA), which is typical of genes coding for inflammatory mediators (Caput et al. 1986) and suggests this gene is involved in inflammatory responses (Bird et al. 2005b).

We were able to identify *IL21* in *P. vitticeps* based on alignment with the recently predicted (Dijkstra et al. 2014) *IL21* in *A. carolinensis* and *P. sinensis* (Chinese softshell turtle) (Supplemental Fig. 2) in conjunction with the partial transcripts the de novo assembly identified in its syntentic location (neighbouring *IL2*, see Fig. 2).

IL5 is the only gene annotated in mammals in the region between the genes *RAD50* and *IRF5* (Fig. 4). We identified a novel transcript in this syntenic region on *P. vitticeps* scaffold

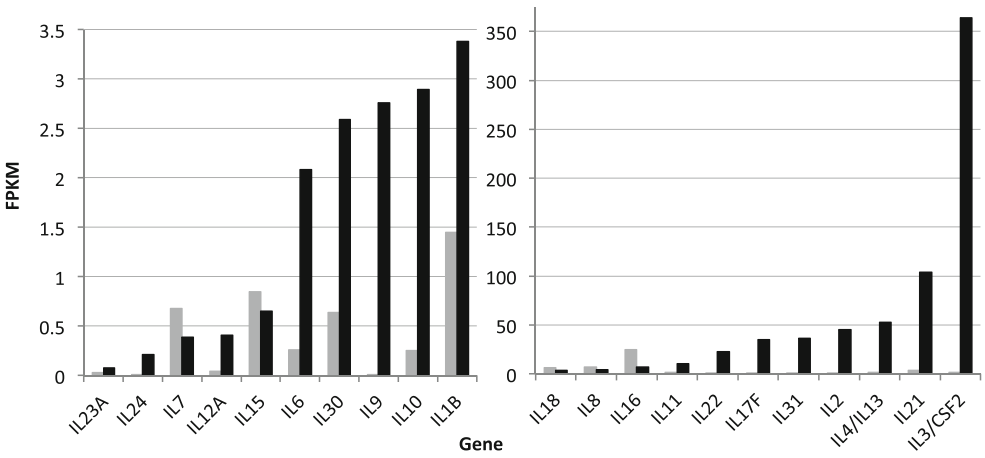

Fig. 1 Expression levels (fragments per kilobase of transcript per million mapped reads: FPKM) of interleukin genes in non-stimulated (*grey*) and PMA/I-stimulated (*black*) splenic cells. The *graph on the left* includes ILs that had low expression levels and the *graph on the right* has include the ILs that were highly expressed

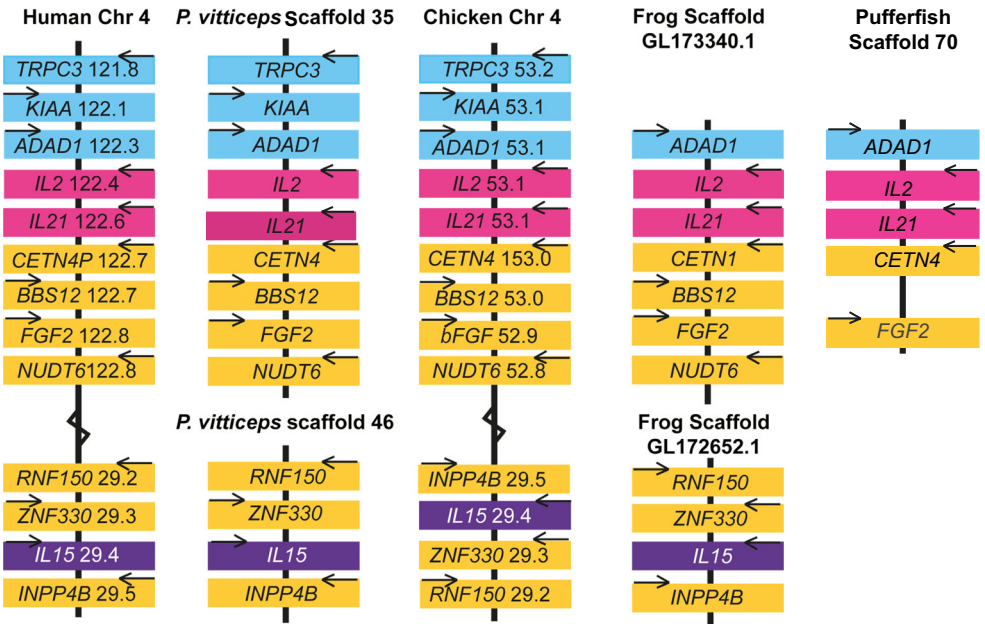

Fig. 2 Syntenic maps of genes in the regions containing *IL2*, *IL21* and *IL15* across vertebrates. The information for *X. tropicalis*, human, chicken and pufferfish (*Fugu rubripes*) is from the Ensembl databases JGI 4.2, GR38.p3, Galgal4 and FUGU 4.0, respectively. *IL2* and *IL21* are highlighted in *pink*, and *IL15* is *purple*. Positions on the chromosomes are noted for human and chicken. *Squiggles* denote an area of the chromosome not included in the synteny map. *Arrows* denote transcriptional orientation

485. It is the only transcript in this location and is in the same transcriptional orientation as in mammals. The amino acid sequence of this transcript shares low identity with human and chicken *IL5* orthologues (14 and 16 %, respectively), although human and chicken *IL5* orthologues only share 20 % identity. The 3′UTR contains a polyadenylation site (AATAAA), indicative of its role in inflammatory response (Bird et al. 2005b). However, this transcript does not share conserved cysteine residues required to form the two disulphide bridges present in the human protein, and the secondary structure prediction is not consistent with *IL5* structure in other vertebrates (Supplemental Fig. 1; Supplemental Fig. 3) (Milburn et al. 1993; Wells et al. 1994).

Since *IL5* evolved from the same common ancestor as *IL3* and *CSF2* (these three genes remain within a 1-Mb cluster on human chromosome 5) (Baxter et al. 2012; Frazer et al. 1997), it is possible that *IL5* did not evolve in the reptile lineage. Prior to identification of this gene in other reptiles, and functional studies, we are unable to assign an identity to this novel transcript in *P. vitticeps*.

Additional ILs identified using a syntenic approach include *IL7*, *IL9*, *IL12A*, *IL15*, *IL21*, *IL27* and *IL31*. The syntenic regions of these six ILs genes are illustrated in Fig. 2, (*IL15*), Fig. 4 (*IL9*) and Supplemental Fig. 4 (*IL7*, *IL12A*, *IL27* and *IL31*), and their characteristics are outlined in Table 1.

Unique homologous syntenic regions in *P. vitticeps*

There were three occurrences of syntenic regions containing only one novel transcript in *P. vitticeps* where two IL genes were expected from synteny in mammals: *IL1A* and *IL1B* (Supplemental Fig. 5), *IL4* and *IL13* (Fig. 4), and *IL3* and *CSF2* (Fig. 4). Because many ILs arose from duplication events (Bao and Reinhardt 2015), the presence of only one gene in *P. vitticeps* could represent either ancestral genes that duplicated in other lineages or loss of one IL in *P. vitticeps*, in each of the three cases. Importantly, there are no gaps in the genome large enough to account for a missing gene in these three regions.

Mammals carry *IL1A* and *IL1B* whereas only *IL1B* has been identified in non-mammal vertebrates (Bird et al. 2002; Ogryzko et al. 2014). *IL18* has been documented in mammals and birds and fish (Bird et al. 2002; Secombes et al. 2011). *IL1A* and *IL1B* are thought to have split from *IL1RN* 350 Mya (Eisenberg et al. 1991), with the divergence of *IL1A* and *IL1B* occurring 250–300 Mya, at approximately the same time as the emergence of mammals (Dunn et al. 2001; Mulero et al. 2000; Young and Sylvester 1989). The presence of both *IL1A* and *IL1B* is therefore a synapomorphy uniting the mammals, with the single *IL1* gene of *P. vitticeps* a retained ancestral state (plesiomorphy).

Sequence identity searches with *IL1A* (21–28 % identity) and *IL1B* (23–42 % identity) orthologues, along with secondary protein structure prediction, suggested that the novel transcript in the region in *P. vitticeps* syntenic with the mammalian region containing *IL1A* and *IL1B* is more similar to *IL1B* than to *IL1A* (Supplemental Fig. 2 and Supplemental Fig. 3). This result was also supported by neighbour-joining trees, which showed clustering of the *P. vitticeps IL1* novel transcript with *IL1β* orthologues of other reptiles (turtle) and birds (Supplemental Fig. 6) as one might expect, whereas the tree relating *Pogona IL1* to *IL1A* had the *P. vitticeps* transcript as a distant outgroup to the remaining vertebrates (Supplemental Fig. 7). Furthermore, the top hit from a protein blast was IL1β-like protein in *Python bivittatus* (XP_007441780.1) (59 % identity). *IL1A* is absent from the *P. bivittatus* genome, as it is in other non-mammalian genomes such as chicken and frog, which have only *IL1B* annotated. Therefore, this transcript, which

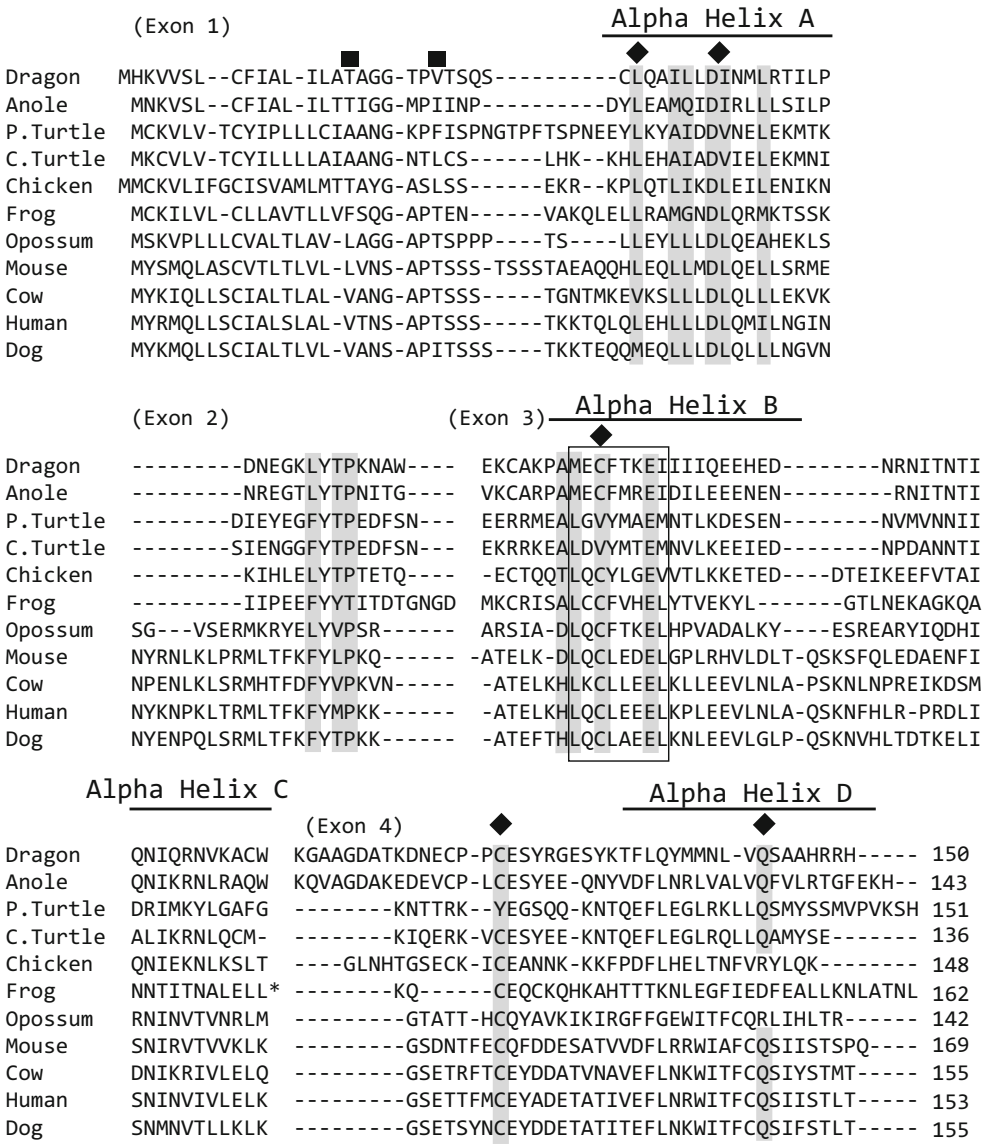

Fig. 3 Multiple alignment of the predicted *P. vitticeps IL2* amino acid sequence with other known *IL2* amino acid sequences. The *IL2/IL15* family signature is *boxed*. Locations of the four alpha helices in human are indicated. Functionally important residues and conserved cysteine residues involved in disulphide bridges are indicated with *diamonds*. Predicted glycosylation sites in *P. vitticeps* are indicated with *squares*. Similar residues are highlighted in *grey*. Sequence length is indicated at the *end of each sequence*. Sequences used for the alignment: human *IL2* (uniprot P04351), chicken *IL2* (uniprot O73883), frog *IL2* (predicted by Dijkstra et al. 2014), anole IL2 (predicted by Dijkstra et al. 2014), painted turtle *IL2* (predicted by Dijkstra et al. 2014), Chinese softshell turtle *IL2* (predicted by Dijkstra et al. 2014), opossum *IL2* (http://bioinf.wehi.edu.au/opossum/seq/cytokines.fa), mouse *IL2* P04351), cow *IL2* (uniprot P05016), dog *IL2* (uniprot Q29416), and dragon *IL2* (*P. vitticeps*). The *asterisk* (*) indicates additional sequence at the end of exon 3 in frog: KLLDASYSKTEPSQYL

we assign to *IL1B* in *P. vitticeps* and *IL1B* in mammals, has deviated less over time from the ancestral state than has mammalian *IL1A*. This result is not novel; we know in mammals that *ILA* is evolving faster than *IL1B* (Young and Sylvester 1989).

IL4 and *IL13* lie in a head-to-tail organization in mammalian genomes. They are functionally closely related (Chomarat and Banchereau 1998; Minty et al. 1993; Moy et al. 2001; Wynn 2003; Zurawski and de Vries 1994) and have highly similar secondary structures: the human orthologues of these two proteins both form four-helix bundles with overall similar structures (Supplemental Fig. 3), typical of their classification in the *IL4*-like family (Brocker et al. 2010; Eisenmesser et al. 2001). We identified only one novel transcript in the syntenic region of *IL4* and *IL13* in *P. vitticeps* (Fig. 4). IL-4 and IL-13 have been described as functionally redundant based on their concurrent expression and use of shared receptors and they are both central to type 2 inflammation (Bao and Reinhardt 2015). Thus, it is unlikely that we only identified one transcript in this region because of an absence of expression of the other.

The novel transcript in the syntenic region of *IL4* and *IL13* in *P. vitticeps* shares low sequence identity with both *IL4* orthologues (15–18 %) and *IL13* orthologues (13–26 %),

Table 1 Comparison of novel interleukins identified in *P. vitticeps* with known human and chicken interleukin sequences

Identity	Scaffold	Number of exons		Number of amino acids		% amino acid identity with		Predicted glyc. sites		Conserved cysteines	Signal peptide identified	3′UTR inst. motif
		P. vitticeps	Human	*P. vitticeps*	Human	Human	Chicken	N-glyc	O-glyc			
IL1B	27	8	7	271	269	37	42	0	0	N/A	No	0
IL2	35	4	4	150	153	28	28	1	4	2 of 2	Yes	2
IL3/CSF2	485	6	5/4	111	152/144	14/18	18/16	8	17	No	No	10
IL4/IL13	815	4	4	148	153/146	16/22	16/13	2	8	Yes	Yes	12
IL7	87	6	3	169	177	25	27	1	1	5 of 6	Yes	3
IL9	43	6	5	142	144	20	20	0	1	N/A	Yes	3
IL12A	147	6	7	202	253	22	28	0	4	4 of 4	Yes	1
IL15	46	6	8	154	162	28	22	3	1	4 of 4	No	1
IL21	35	5	4	145	155	22	26	1	2	4 of 4	Yes	5
IL30	154	5	5	215	243	24	N/A	0	0	N/A	Yes	5
IL31	329	3	3	188	164	23	N/A	0	0	N/A	No	16

N/A not applicable

although identity between taxa for these two proteins is generally low (identity between human and chicken *IL4* and *IL13* orthologues is only 27 and 20 %, respectively). Cysteine residues involved in the formation of disulphide bridges in both *IL4* and *IL13* in human are present in *P. vitticeps* (Fig. 5), and the predicted secondary structure of the novel *IL4/IL13* transcript is consistent with an *IL4*-like family member (Supplemental Fig. 1; Supplemental Fig. 3).

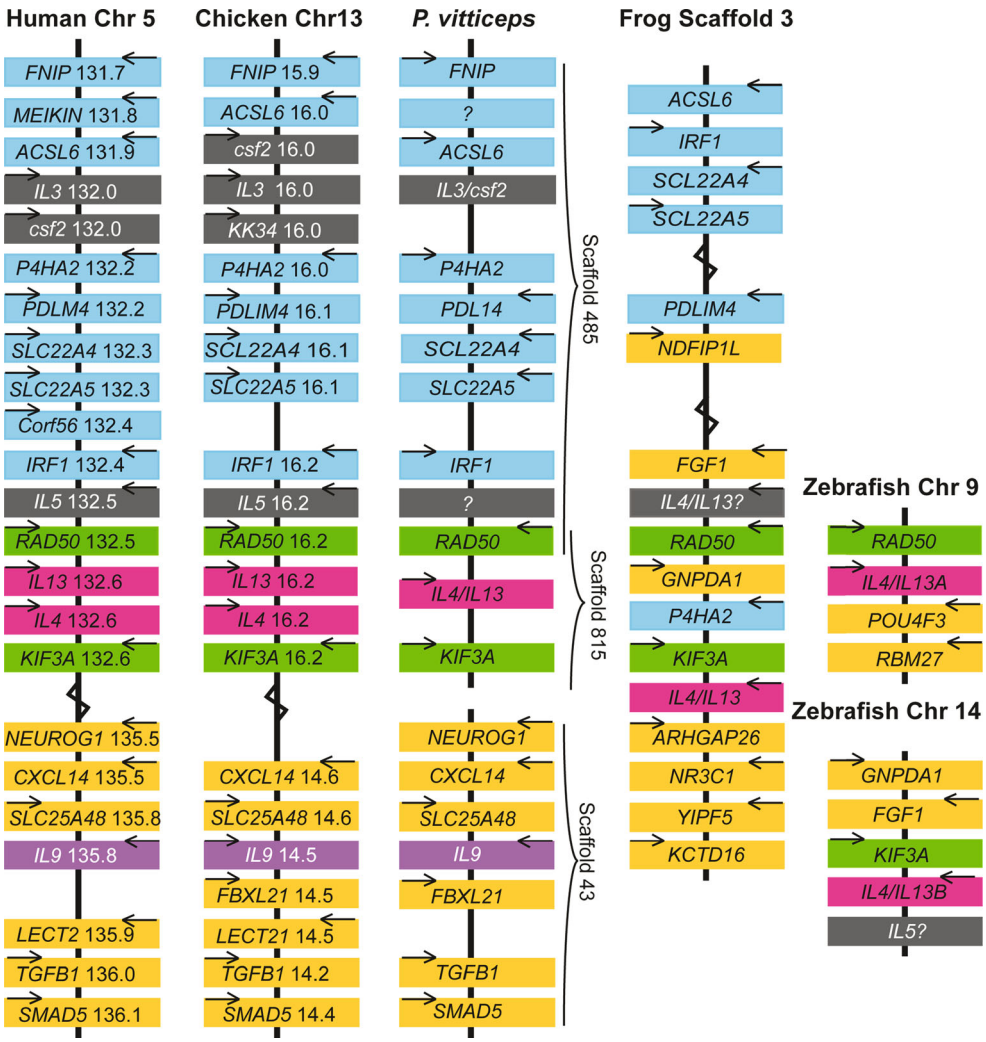

Fig. 4 Syntenic maps of genes in the *IL4/IL13* region, the *IL3/CSF2* region and the *IL9* region across vertebrates. The information for the *X. tropicalis KIF3A/IL4/IL13/RAD50* locus is from Wang and Secombes (2015). The information for zebrafish is from Ohtani et al. (2008) and the Ensembl database GRCz10. Information for human and chicken is from the Ensembl databases GR38.p3 and Galgal4, respectively. *RAD50* and *KIF3A* are highlighted in *green*, *IL4*, *IL13* and *IL4/IL13* are *pink*, *IL9* is *purple* and related genes *IL5*, *IL3*, *CSF2* and *KK35* are *grey*. *IL3*, *CSF2* and *IL9* have not been annotated in frog and zebrafish. Positions on the chromosomes are noted for human, chicken. *Squiggles* denote an area of the chromosome not included in the synteny map. *Arrows* denote transcriptional orientation

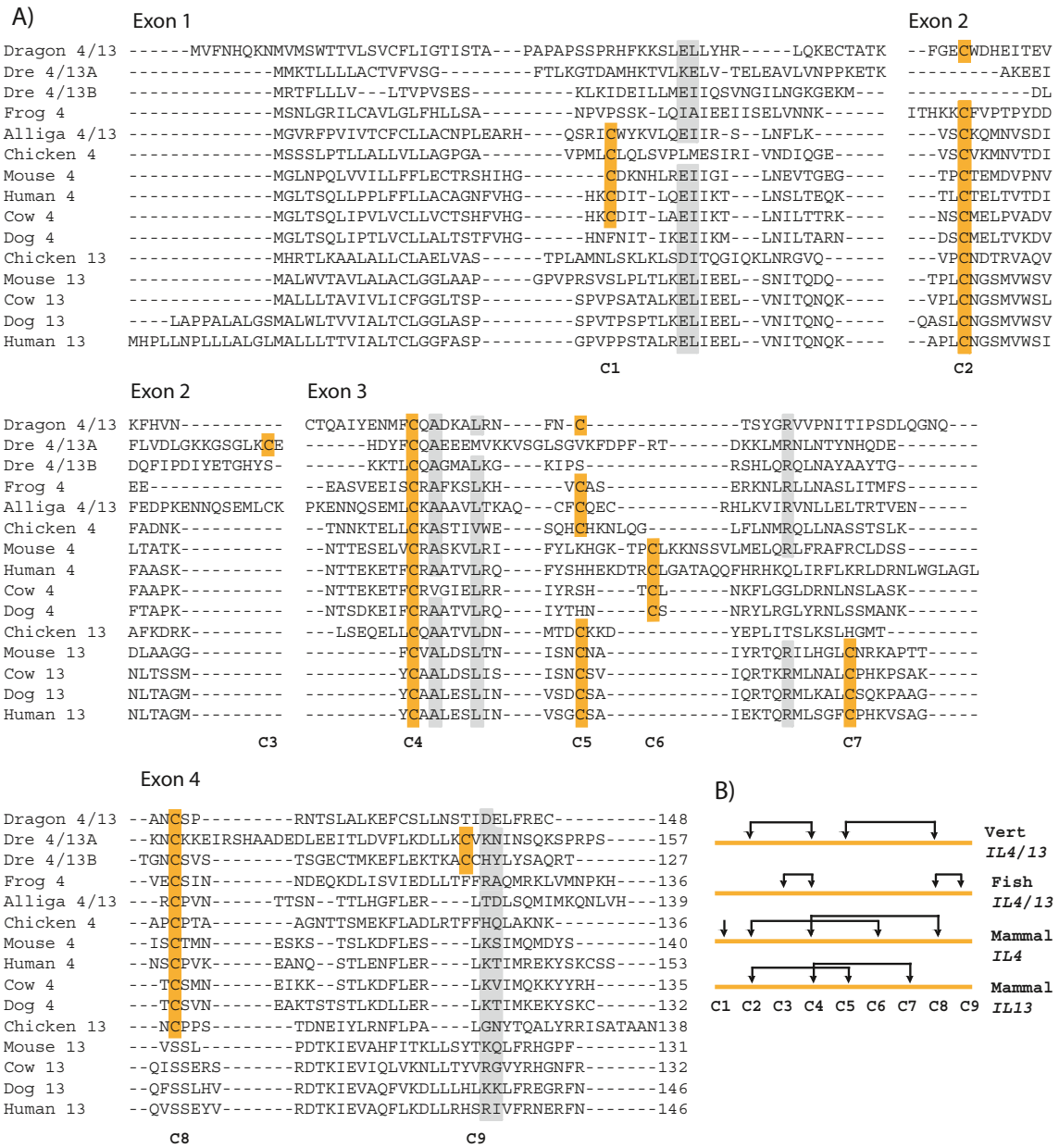

Fig. 5 **a** Clustal Omega Multiple alignment of vertebrate *IL13*, *IL4* and *IL4*/*IL13* protein sequences illustrating the location of cysteine residues known to be required for the formation of disulphide bonds in human *IL4* and *IL13*. Conserved cytosines are highlighted in *orange*. Similar residues are highlighted in *grey*. **b** Cartoon summarizing the location of cysteine residues in vertebrate IL4/IL13, mammalian *IL4*, mammalian *IL13* and fish *IL4*/*IL13*. This cartoon was inspired by Fig. 3 in Wang et al. (2016). Sequences used in this alignment include the following: dragon *IL4*/*IL13*, zebrafish (Dre) *IL4*/*IL13A* (uniprot B3IWZ9), zebrafish (Dre) *IL4*/*IL13B* (uniprot D1YSM1), alligator (Alliga) *IL4*/*IL13* (XP_006023629), chicken *IL13* (uniprot Q5W4U0), mouse *IL13* (uniprot P20109), dog *IL13* (uniprot Q9N0W9), human *IL13* (uniprot P35225), cow *IL13* (uniprot Q9XSV9), frog *IL4*/*IL13* (uniprot A9JPI4), chicken *IL4* (uniprot Q5W4U1), mouse *IL4* (uniprot P07750), human *IL4* (P05112), cow *IL4* (uniprot P30367), and dog *IL4* (uniprot O77762)

Furthermore, the exon/intron organization is consistent with orthologues of both *IL4* and *IL13* (Fig. 5), and transcriptional orientation of the *P. vitticeps IL4*/*IL13* transcript relative to adjacent genes *KIF3A* and *RAD50* is the same as seen in mammals and birds. Alignment of *P. vitticeps IL4*/*IL13* with either *IL4* or *IL13* homologues in other taxa proved uninformative, a reflection of very low sequence identity and precluding phylogenetic analysis.

Two *IL4*/*IL13*-like genes have been identified in several teleost fish (Bird and Secombes 2006; Li et al. 2007; Wang et al. 2016). These two genes in fish have been called *IL4*/*13A* and *IL4*/*13B* because the similarity of each to *IL4* and *IL13* is unclear, and the two genes are likely a result of fish-specific whole genome duplication of an ancestral *IL4*/*IL13* gene (Ohtani et al. 2008; Wang et al. 2016). At least two IL4/*IL13*-like genes have been identified in the elephant shark

Callorhinchus milii in the conserved syntenic location between *KIF3A* and *RAD50*, and potentially two *IL4*/*IL13* genes have been identified in an amphibian genome (the frog *Xenopus tropicalis*) (reviewed in Wang and Secombes 2015); one copy of *IL4*/*13* in frog is on the wrong side of *KIF3A*, but is in the correct syntenic transcriptional orientation.

Overall, it is likely that a single *IL4*/*IL13* gene in the ancestor of jawed vertebrates duplicated in different lineages, which is via whole genome duplication in fish and via tandem duplication events in other mammals, birds and amphibians, but it seems not in reptiles. The transcript we identified in *P. vitticeps* likely represents this single ancestral *IL4*/*IL13* gene. However, it is also possible that reptiles lost one copy of the *IL4*/*IL13* gene. Future studies that extensively analyze this region in additional reptilian and amphibian genomes will consolidate the evolutionary history of this pair of genes.

Like we saw for *IL4*/*IL13*, there was only one novel transcript in the syntenic region between *ACSL6* and *P4HA2* on *P. vitticeps* scaffold 485, which in humans contains *IL3* and *CSF2* (Fig. 4). *IL3* and *CSF2* are both granulocyte-macrophage colony-stimulating factors (GM-CSFs) that activate the same signalling pathways and have similar functions (Alexander 1998; Nicola 1989). *IL3* has a wider range of functions including stimulating the proliferation, differentiation and activation of haematopoietic progenitor cells (Frolova et al. 1991). Although they do not share considerable identity (Yang et al. 1986), these two genes, which are only separated by 10 kb on human chromosome 5, arose via duplication of a common ancestral gene that also gave rise to *IL5* (Cockerill 2004; Frolova et al. 1991). Interestingly, birds have a third gene in this region, *KK34*, which has been proposed to be a duplication of an ancestral *IL5* gene (Avery et al. 2004) and is likely unique to the avian lineage (Huising et al. 2006). This avian-specific duplication clearly occurred after the reptilian and avian lineage split since we detected only one gene whereas birds have three genes in this region.

The predicted secondary structure of the protein derived from the novel transcript in the syntenic location of *IL3* and *CSF2* in *P. vitticeps* is one composed of four alpha helices and three beta sheets (Supplemental Fig. 1). This is similar to the secondary structures of both *IL3* and *CSF2* in human which both exhibit characteristic four-helix bundles arranged in an up-up-down-down topology like that of *IL4* and *IL2* (Bazan and McKay 1992), with a fifth helix crossing over loop AB (Mirabella et al. 2010). The transcriptional orientation of the *IL3*/*CSF2* transcript in *P. vitticeps* relative to *ACSL6* and *P4HA2* is the same as seen in mammals and birds (Fig. 4), but it shares low identity (less than 20 %) with both *IL3* and *CSF2* (Table 1). Furthermore, this novel transcript in *P. vitticeps* is only 111 amino acids in length, whereas in human, it is 152 amino acids. However, low identity between taxa is common for ILs; human *IL3* and *CSF2* share only 8.4 % identity (Emboss needle alignment) (Goujon et al. 2010; Sievers et al. 2011).

Although this novel transcript is at the expected location of *IL3* or *CSF2* and is structurally related to these two proteins, we do not have strong arguments for its identification as either protein. It could also be related to *IL5*, which is part of the same family as *IL3* and *CSF2*. Indeed, alignment of *P. vitticeps IL3*/*CSF2* novel transcript with mammalian and avian amino acid sequences for *IL3* and *CSF2* (see "Methods") did not produce informative phylogenetic comparisons. To our knowledge, neither *IL3* nor *CSF2* has been identified in fish (Secombes et al. 2011; Yamaguchi et al. 2015), amphibians or reptiles, so it was not possible to include sequences from more closely related species in the alignments. We suggest that the novel transcript identified here in *P. vitticeps* in the syntenic region of *IL3* and *CSF2* represents an ancestral gene that has duplicated in more recent vertebrate evolution in mammals and birds. Identification of the gene(s) present at this locus in other lizard genomes and in amphibians will provide insight into the evolution of these two genes.

Future annotation of the above three syntenic regions in other lizards, amphibians and fish will provide clues about when duplication events of ancestral genes *IL1*, *IL4*/*IL13* and *IL5*/*IL3*/*CSF2* occurred, which may provide insight into the pressures driving the evolution of IL gene families in mammals.

Interleukins with corrected annotation

We used our novel transcripts to fine-tune the published annotations of ILs (Georges et al. 2015). Ten previously annotated ILs were represented in our novel transcripts: *IL6*, *IL8*, *IL10*, *IL11*, *IL16*, *IL17F*, *IL18*, *IL22*, *IL23A* and *IL24*; all of which had some level of error in their annotation (Table 2). For example, *IL8* (synonym: *CXCL8*) is annotated in the genome three times. Using the syntenic order of genes and alignment of novel transcripts in this region, we were able to identify the *IL8* sequence in *P. vitticeps* (Table 2). It shares relatively high identity with human and chicken *IL8* orthologues (50 and 63 %, respectively) and has a similar predicted secondary structure. Four cysteine residues involved in two disulphide bonds in human (in human Cys^{34}-Cys^{61} and Cys^{36}-Cys^{77}) are conserved throughout mammals, chicken and *P. vitticeps IL8* sequences. From this syntenic region in humans, the other two genes identified as *IL8* in the published annotation likely represent other members of the CXCL family, including *CXCL6*, *CXCL1*, *CXCL5* or *CXCL2*, which, along with *IL8*, are clustered between *ANKRD17* and *MTHFD2L* in human.

No additional *IL10* family members (*IL10*, *IL19*, *IL20*, *IL22*, *IL24* and *IL26*) were identified with the de novo reconstructed transcripts. *IL10*, *IL19* and *IL24* are clustered on *P. vitticeps* scaffold 613. In human, *IL20* is part of this syntenic region but it is missing from the published *P. vitticeps* annotation and our series of de novo transcripts, suggesting it arose via duplication from an ancestral *IL10*/*IL19*/*IL24* gene during vertebrate evolution. Interestingly,

Table 2 Comparison of published *P. vitticeps* interleukins with human, chicken and anole interleukin sequences. Corrections to current gene annotations are noted, as are the number of instability (inst.) motif sequences in the 3'UTR

IL with corrected annotation	Scaffold	Number of exons		Number of amino acids		% amino acid identity with			Problem with published annotation				3'UTR inst. motif
		P. vitticeps	Human	*P. vitticeps*	Human	Human	Chicken	Anole	Incorrect exon number	3'UTR truncation	5'UTR truncation	Other	
IL6	4	5	5	219	212	27	37	55	Missing exon 3			Exon sizes	1
IL8	855	5	4	100	99	50	63	73		Add 217 bp		Annotated 3 times	2
IL10	613	5	5	154	178	48	51	70		Add 823 bp	Add 163 bp		5
IL11	248	8	5	182	199	45	N/A	87		Add 1409 bp		Incorrect structure	3
IL16	121	7	7	709	631	53	57	65	3 extra exons			Incorrect structure	0
IL17F	129	5	3	175	163	35	47	N/A	Missing 1st exon	Add 441 bp			6
IL18	7	6	5	207	193	31	43	61	Missing 3 exons			Orientation	0
IL22	74	6	5	198	179	42	51	67	Missing last exon			Exon sizes, orientation	8
IL23A	4	4	4	162	189	32	N/A	N/A				Orientation	0
IL24	613	4	6	131	206	32	31	76				Shifted by two exons	0

N/A not applicable

the only *IL10* family members that have been identified in fish, *IL20* and *IL26*, are the only two that are not identified in *P. vitticeps*. It has been proposed that the ancestral gene that gave rise to *IL19*, *IL20* and *IL24* in mammals also gave rise to *IL20* in fish (Igawa et al. 2006; Wang et al. 2010).

The four-exon novel *IL24* transcript shifted the published annotation of *IL24* upstream by two exons, excluding exons III and IV, and suggesting that this was the only isoform expressed in this tissue under these conditions. The published annotation of *IL24* may represent a splice variant; human *IL24* has seven splice variants, or it could be incorrect; turtle has a single four-exon variant. *IL19* was not represented in the transcripts, suggesting that it was not expressed in NS cells or during the PMA/I stimulation (ST).

A few of the annotated ILs had the incorrect number of exons, such as *IL16*, which is annotated in the genome with 10 exons. We identified a seven-exon novel transcript, suggesting that this was the only isoform expressed in this tissue under these conditions, or, since human *IL16* also has seven exons (Table 2), that the 10-exon annotation is incorrect. Similarly, *IL18* is annotated with only three exons whereas the novel transcript has six exons, adding three exons to the start of the annotation. The annotations of *IL17F* and *IL22* are missing the first and last exons, respectively, in the published annotation, added here using the novel transcript annotation (Table 2).

Missing interleukins

Five ILs (*IL17B*, *IL17C*, *IL17D*, *IL19* and *IL25*) are annotated in the *P. vitticeps* genome but were not represented by transcripts in this study, suggesting they were not expressed during the PMA/I stimulation. However, this could reflect our use of splenic cells instead of sorted T cells. Repeating these experiments with sorted immune cell subsets may reveal expression of these five annotated ILs.

ILs that are not annotated in the genome nor were they expressed under the conditions of this study include *IL17A*, *IL20*, *IL26* and *IL33*. Importantly, there are no gaps in the genome in the region where *IL20* is expected. There are small gaps in genome in the region expected to contain *IL26*, although not large enough to miss an entire gene. However, gaps in the genome where *IL33* and *IL17A* could be are large enough to account for omission of these two genes.

To be sure these ILs are absent from the *P. vitticeps* genome, expression from different tissues or conditions as well as from pure immune cell subsets must be examined. For example, stimulation with different mitogens phytohemagglutinin (PHA), lipopolysaccharide (LPS), concanavalin A (Con A) or pokeweed mitogen (PWM) may elicit expression of a different subset of ILs (Baran et al. 2001). Indeed, it has been demonstrated that different stimulants produce unique cytokine profiles from rat blood immune cells (Wenchao et al.

2013); LPS significantly promoted the expression of TNF-α and *IL10*, but only weakly promoted the production of *IL2*. In contrast, PHA promoted the productions of *IL2*, *RANTES* and *TGF-β*, and weakly increased the levels of *IL4* and *IL6*. PMA/I promoted expression of *IL2*, interferon-gamma (IFN-γ), TNF-α, TNF-β and *RANTES* (Wenchao et al. 2013). Different mitogens also have been shown to affect the cytokine production from CD4 T cell clones. Stimulation with PMA/I produced a significantly higher amount of IFN-γ and *IL17* compared with anti-CD3/anti-CD28 stimulation, whereas the reverse was the case for *IL10* production (Olsen and Sollid 2013).

Furthermore, lizards undergo seasonal physiological variation associated with environmental temperatures (Zimmerman et al. 2010), including structural changes of their lymphoid tissues (El Ridi 1981; Hussein et al. 1978, 1979; Zapata et al. 1992), which affects the immune system (reviewed in Hsu 1998). Allograft rejection (Afifi et al. 1993) and T cell proliferation have been shown to vary with season. For example, Farag and El Ridi (1985) observed seasonal patterns in the proliferation of leukocytes from striped sand snake (*Psammophis sibilans*) and diadem snake (*Spalerosophis diadema*), with the greatest proliferation occurring during spring and autumn (Farag and El Ridi 1985, 1990). In the Caspian pond turtle (*Mauremys caspica*), lymphocyte proliferation in response to Con A and PHA was strongest in spring but was significantly reduced in summer, autumn and winter (Munoz and De la Fuente 2001). We observed differential expression of most of the ILs we identified in *P. vitticeps* in this study, indicating that the splenic cells from the wild-caught male used here, who was killed during late spring, had the capacity for an immune response. It is conceivable that we would observe a modified profile of ILs expressed from PMA/I-stimulated splenic cells if we studied animals from different seasons. The response to PMA/I-stimulation must be studied in more lizards from all seasons before making any further speculations on the IL gene repertoire in *P. vitticeps*.

Conclusion

Here, we manually curated IL reptilian genes central to the immune response in mammals. Our identification of eight new IL genes in *P. vitticeps* suggests that lizards are capable of a complex immune response comparable to those seen in mammals. Our results bring the total number of annotated IL genes to 26, the highest in any lizard species, which lays the groundwork for future studies of the immune response in lizards. The development of lizard-specific immunological reagents, such as CD4 and CD8 cell markers, will enable better characterization of reptile immunity, including expression and epigenetic profiling, and allow for comparison with mammals. Revealing similarities and differences between the mammalian and reptilian immune responses will provide insight into the evolution of the vertebrate immune system.

Acknowledgments This work was supported by a University of Canberra postdoctoral fellowship (awarded to Tariq Ezaz, Sudha Rao, Stephen Sarre, Janine Deakin, Kris Hardy and Arthur Georges, and supporting Renae Domaschenz and Alexandra Livernois). Tariq Ezaz is supported by an Australian Research Council Future Fellowship (FT110100733).

Compliance with ethical standards Wild-caught male *P. vitticeps* (Id #UC<Aus>:111880108089) was euthanized according to ethics approval by the Animal Ethics Committee (reference number CE-13-10) at the University of Canberra.

References

Afifi A, Mohamed ER, El Ridi R (1993) Seasonal conditions determine the manner of rejection in reptiles. J Exp Zool 265:459–468

Alexander WS (1998) Cytokines in hematopoiesis. Int Rev Immunol 16: 651–682

Avery S, Rothwell L, Degen WD, Schijns VE, Young J, Kaufman J, Kaiser P (2004) Characterization of the first nonmammalian T2 cytokine gene cluster: the cluster contains functional single-copy genes for IL-3, IL-4, IL-13, and GM-CSF, a gene for IL-5 that appears to be a pseudogene, and a gene encoding another cytokinelike transcript, KK34. J Interf Cytokine Res 24:600–610

Bao K, Reinhardt RL (2015) The differential expression of IL-4 and IL-13 and its impact on type-2 immunity. Cytokine 75:25–37

Baran J, Kowalczyk D, Ozog M, Zembala M (2001) Three-color flow cytometry detection of intracellular cytokines in peripheral blood mononuclear cells: comparative analysis of phorbol myristate acetate-ionomycin and phytohemagglutinin stimulation. Clin Diagn Lab Immunol 8:303–313

Baxter EW, Mirabella F, Bowers SR, James SR, Bonavita AM, Bertrand E, Strogantsev R, Hawwari A, Bert AG, Gonzalez de Arce A, West AG, Bonifer C, Cockerill PN (2012) The inducible tissue-specific expression of the human IL-3/GM-CSF locus is controlled by a complex array of developmentally regulated enhancers. J Immunol 189:4459–4469

Bazan JF, McKay DB (1992) Unraveling the Structure of IL-2. Science 257:410–413

Beck G (1998) Macrokines: invertebrate cytokine-like molecules? Front Biosci 3:d559–d569

Bird S, Secombes CJ (2006) *Danio rerio* partial mRNA for interkeukin-4 (il-4 gene). GenBank Accession No. AM403245

Bird S, Zou J, Kono T, et al. (2005a) Characterisation and expression analysis of interleukin 2 (IL-2) and IL-21 homologues in the Japanese pufferfish, *Fugu rubripes*, following their discovery by synteny. Immunogenetics 56:909–923

Bird S, Zou J, Savan R, Kono T, Sakai M, Woo J, Secombes C (2005b) Characterisation and expression analysis of an interleukin 6 homologue in the Japanese pufferfish, *Fugu rubripes*. Dev Comp Immunol 29:775–789

Bird S, Zou J, Wang T, Munday B, Cunningham C, Secombes CJ (2002) Evolution of interleukin-1beta. Cytokine Growth Factor Rev 13: 483–502

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120

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

Caput D, Beutler B, Hartog K, Thayer R, Brown-Shimer S, Cerami A (1986) Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. Proc Natl Acad Sci U S A 83:1670–1674

Chomarat P, Banchereau J (1998) Interleukin-4 and interleukin-13: their similarities and discrepancies. Int Rev Immunol 17:1–52

Cockerill PN (2004) Mechanisms of transcriptional regulation of the human IL-3/GM-CSF locus by inducible tissue-specific promoters and enhancers. Crit Rev Immunol 24:385–408

Commins SP, Borish L, Steinke JW (2010) Immunologic messenger molecules: cytokines, interferons, and chemokines. J Allergy Clin Immunol 125:S53–S72

Crawford JE, Guelbeogo WM, Sanou A, Traoré A, Vernick KD, Sagnon N, Lazzaro BP (2010) De novo transcriptome sequencing in *Anopheles funestus* using Illumina RNA-seq technology. PLoS One 5:e14202

Diaz-Rosales P, Bird S, Wang TH, Fujiki K, Davidson WS, Zou J, Secombes CJ (2009) Rainbow trout interleukin-2: cloning, expression and bioactivity analysis. Fish Shellfish Immunol 27:414–422

Dijkstra JM, Takizawa F, Fischer U, Friedrich M, Soto-Lampe V, Lefevre C, Lenk M, Karger A, Matsui T, Hashimoto K (2014) Identification of a gene for an ancient cytokine, interleukin 15-like, in mammals; interleukins 2 and 15 co-evolved with this third family member, all sharing binding motifs for IL-15Ralpha. Immunogenetics 66:93–103

Dinarello CA, Wolff SM (1993) The role of interleukin-1 in disease. N Engl J Med 328:106–113

Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: a protein secondary structure prediction server. Nucleic Acids Res 43:W389–W394

Dunn E, Sims JE, Nicklin MJ, O'Neill LA (2001) Annotating genes with potential roles in the immune system: six new members of the IL-1 family. Trends Immunol 22:533–536

Eberl G, Colonna M, Di Santo JP, McKenzie AN (2015) Innate lymphoid cells. Innate lymphoid cells: a new paradigm in immunology. Science 348:aaa6566

Eisenberg SP, Brewer MT, Verderber E, Heimdal P, Brandhuber BJ, Thompson RC (1991) Interleukin-1 receptor antagonist is a member of the interleukin-1 gene family—evolution of a cytokine control mechanism. Proc Natl Acad Sci U S A 88:5232–5236

Eisenmesser EZ, Horita DA, Altieri AS, Byrd RA (2001) Solution structure of interleukin-13 and insights into receptor engagement. J Mol Biol 310:231–241

El Ridi R, Badir N, El Rouby S (1981) Effect of seasonal variations on the immune system of the snake, *Psammophis schokari*. J Exp Zool 216:357–365

Farag MA, El Ridi R (1985) Mixed leucocyte reaction (MLR) in the snake *Psammophis sibilans*. Immunology 55:173–181

Farag MA, El Ridi R (1990) Functional markers of the major histocompatibility gene complex of snakes. Eur J Immunol 20:2029–2033

Frazer KA, Ueda Y, Zhu Y, Gifford VR, Garofalo MR, Mohandas N, Martin CH, Palazzolo MJ, Cheng JF, Rubin EM (1997) Computational and biological analysis of 680 kb of DNA sequence from the human 5q31 cytokine gene cluster region. Genome Res 7:495–512

Frolova EI, Dolganov GM, Mazo IA, Smirnov DV, Copeland P, Stewart C, O'Brien SJ, Dean M (1991) Linkage mapping of the human CSF2 and IL3 genes. Proc Natl Acad Sci U S A 88:4821–4824

Gaffen SL, Liu KD (2004) Overview of interleukin-2 function, production and clinical applications. Cytokine 28:109–123

Georges A, Li Q, Lian J, O'Meally D, Deakin J, Wang Z, Zhang P, Fujita M, Patel HR, Holleley CE, Zhou Y, Zhang X, Matsurbara K, Waters P, Graves JAM, Sarre SD, Zhang G (2015) High-coverage sequencing and annotated assembly of the genome of the Australian dragon lizard *Pogona vitticeps*. Gigascience 4:45

Gibson MS, Kaiser P, Fife M (2014) The chicken IL-1 family: evolution in the context of the studied vertebrate lineage. Immunogenetics 66:427–438

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38:W695–W699

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol 29:644–652

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494–1512

Harrison GA, Wedlock DN (2000) Marsupial cytokines. Structure, function and evolution. Dev Comp Immunol 24:473–484 http://www.cbs.dtu.dk/services/NetNGlyc/N

Hsu E (1998) Mutation, selection, and memory in B lymphocytes of exothermic vertebrates. Immunol Rev 162:25–36

Huising MO, Kruiswijk CP, Flik G (2006) Phylogeny and evolution of class-I helical cytokines. J Endocrinol 189:1–25

Hussein MF, Badir N, el-Ridi R, Akef M (1978) Effect of seasonal variation on lymphoid tissues of the lizards, *Mabuya quinquetaeniata* Licht. and *Uromastyx aegyptia* Forsk. Dev Comp Immunol 2:469–478

Hussein MF, Badir N, el-Ridi R, el-Deeb S (1979) Effect of seasonal variation on immune system of the lizard *Scincus scincus*. J Exp Zool 209:91–96

Igawa D, Sakai M, Savan R (2006) An unexpected discovery of two interferon gamma-like genes along with interleukin (IL)-22 and -26 from teleost: IL-22 and -26 genes have been described for the first time outside mammals. Mol Immunol 43:999–1009

Kotenko SV (2002) The family of IL-10-related cytokines and their receptors: related, but to what extent? Cytokine Growth Factor Rev 13:223–240

Kumagai N, Benedict SH, Mills GB, Gelfand EW (1987) Requirements for the simultaneous presence of phorbol esters and calcium ionophores in the expression of human T lymphocyte proliferation-related genes. J Immunol 139:1393–1399

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Li JH, Shao JZ, Xiang LX, Wen Y (2007) Cloning, characterization and expression analysis of pufferfish interleukin-4 cDNA: the first evidence of Th2-type cytokine in fish. Mol Immunol 44:2078–2086

Milburn MV, Hassell AM, Lambert MH, Jordan SR, Proudfoot AE, Graber P, Wells TN (1993) A novel dimer configuration revealed by the crystal structure at 2.4 A resolution of human interleukin-5. Nature 363:172–176

Minty A, Chalon P, Derocq JM, Dumont X, Guillemot JC, Kaghad M, Labit C, Leplatois P, Liauzun P, Miloux B, et al. (1993) Interleukin-13 is a new human lymphokine regulating inflammatory and immune responses. Nature 362:248–250

Mirabella F, Baxter EW, Boissinot M, James SR, Cockerill PN (2010) The human IL-3/granulocyte-macrophage colony-stimulating factor locus is epigenetically silent in immature thymocytes and is progressively activated during T cell development. J Immunol 184:3043–3054

Morgan DA, Ruscetti FW, Gallo R (1976) Selective in vitro growth of T lymphocytes from normal human bone marrows. Science 193: 1007–1008

Moy FJ, Diblasio E, Wilhelm J, Powers R (2001) Solution structure of human IL-13 and implication for receptor binding. J Mol Biol 310: 219–230

Mulero JJ, Nelken ST, Ford JE (2000) Organization of the human interleukin-1 receptor antagonist gene IL1HY1. Immunogenetics 51:425–428

Munoz FJ, De la Fuente M (2001) The effect of the seasonal cycle on the splenic leukocyte functions in the turtle *Mauremys caspica*. Physiol Biochem Zool 74:660–667

Nicola NA (1989) Hemopoietic cell growth factors and their receptors. Annu Rev Biochem 58:45–77

Ogryzko NV, Renshaw SA, Wilson HL (2014) The IL-1 family in fish: swimming through the muddy waters of inflammasome evolution. Dev Comp Immunol 46:53–62

Ohtani M, Hayashi N, Hashimoto K, Nakanishi T, Dijkstra JM (2008) Comprehensive clarification of two paralogous interleukin 4/13 loci in teleost fish. Immunogenetics 60:383–397

Olsen I, Sollid LM (2013) Pitfalls in determining the cytokine profile of human T cells. J Immunol Methods 390:106–112

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786

Rao S, Gerondakis S, Woltring D, Shannon MF (2003) c-Rel is required for chromatin remodeling across the IL-2 gene promoter. J Immunol 170:3724–3731

Rao S, Procko E, Shannon MF (2001) Chromatin remodeling, measured by a novel real-time polymerase chain reaction assay, across the proximal promoter region of the IL-2 gene. J Immunol 167:4494–4503

Roberts A, Pachter L (2013) Streaming fragment assignment for real-time analysis of sequencing experiments. Nat Methods 10:71–73

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nat Biotechnol 29:24–26

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140

Secombes CJ, Wang T, Bird S (2011) The interleukins of fish. Dev Comp Immunol 35:1336–1345

Sievers F, Wilm A, Dineen D, Gibson T, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson J, Higgins D (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7

Takahashi A, de Andres MC, Hashimoto K, Itoi E, Oreffo RO (2015) Epigenetic regulation of interleukin-8, an inflammatory chemokine, in osteoarthritis. Osteoarthr Cartil 23:1946–1954

Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178–192

Wang T, Diaz-Rosales P, Martin SA, Secombes CJ (2010) Cloning of a novel interleukin (IL)-20-like gene in rainbow trout *Oncorhynchus mykiss* gives an insight into the evolution of the IL-10 family. Dev Comp Immunol 34:158–167

Wang T, Johansson P, Abos B, Holt A, Tafalla C, Jiang Y, Wang A, Xu Q, Qi Z, Huang W, Costa MM, Diaz-Rosales P, Holland JW, Secombes CJ (2016) First in-depth analysis of the novel Th2-type cytokines in salmonid fish reveals distinct patterns of expression and modulation but overlapping bioactivities. Oncotarget 7:10917–10946

Wang T, Secombes CJ (2015) The evolution of IL-4 and IL-13 and their receptor subunits. Cytokine 75:8–13

Wells TN, Graber P, Proudfoot AE, Arod CY, Jordan SR, Lambert MH, Hassel AM, Milburn MV (1994) The three-dimensional structure of human interleukin-5 at 2.4-angstroms resolution: implication for the structures of other cytokines. Ann N Y Acad Sci 725:118–127

Wenchao A, Haishan L, Song N, Lei L, Huiming C (2013) Optimal method to stimulate cytokine production and its use in immunotoxicity assessment. Int J Environ Res Public Health 10:3834–3842

Wong ES, Young LJ, Papenfuss AT, Belov K (2006) In silico identification of opossum cytokine genes suggests the complexity of the marsupial immune system rivals that of eutherian mammals. Immunome Res 2:4

Wu TD, Watanabe CK (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21:1859–1875

Wynn TA (2003) IL-13 effector functions. Annu Rev Immunol 21:425–456

Yamaguchi T, Takizawa F, Fischer U, Dijkstra JM (2015) Along the axis between type 1 and type 2 immunity; principles conserved in evolution from fish to mammals. Biology (Basel) 4:814–859

Yang YC, Ciarletta AB, Temple PA, Chung MP, Kovacic S, Witek-Giannotti JS, Leary AC, Kriz R, Donahue RE, Wong GG, et al. (1986) Human IL-3 (multi-CSF): identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3. Cell 47:3–10

Young PR, Sylvester D (1989) Cloning of rabbit interleukin-1 beta: differential evolution of IL-1 alpha and IL-1 beta proteins. Protein Eng 2:545–551

Zapata AG, Varas A, Torroba M (1992) Seasonal variations in the immune system of lower vertebrates. Immunol Today 13:142–147

Zelus D, Robinson-Rechavi M, Delacre M, Auriault C, Laudet V (2000) Fast evolution of interleukin-2 in mammals and positive selection in ruminants. J Mol Evol 51:234–244

Zimmerman LM, Vogel LA, Bowden RM (2010) Understanding the vertebrate immune system: insights from the reptilian perspective. J Exp Biol 213:661–671

Zurawski G, de Vries JE (1994) Interleukin 13 elicits a subset of the activities of its close relative interleukin 4. Stem Cells 12:169–174