CHARLES UNIVERSITY

Faculty of Science

Study program: Biology

Branch of study: Immunology

Bc. Markéta Harazim

Coevolution of cytokines from the interleukin 10 family and their receptors

Koevoluce cytokinů z rodiny interleukinu 10 a jejich receptorů

DIPLOMA THESIS

Supervisor: doc. Ing. Bohdan Schneider, CSc.

Praha, 2017

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 25.4.2017

Markéta Harazim

Acknowledgement

Firstly, I would like to express my gratitude to Bohdan Schneider, the supervisor of my thesis, for his professional advice, guidance and useful remarks. I would also like to thank all the members of Laboratory of Biomolecular Recognition of Institute of Biotechnology of CAS, above all to Hana Pařízková and Jiří Zahradník for their useful remarks on my thesis and inspiring discussion.

I thank Alena Černíková for teaching statistics to lost cases in mathematics so patiently and for help with understanding the methods used in the thesis and Natália Martínková for her helpful advice and for making computers obey a little more.

Last, but not least, I thank my parents, my brother, Darka and Jack for their everlasting support and understanding.

ABSTRAKT

Rodina interleukinu 10 (FIL-10, family of interleukin 10) je důležitá skupina cytokinů regulující imunitní odpověď různého charakteru, od protizánětlivé odpovědi interleukinu (IL) 10, přes reakci regulující imunitní odpověď epitelech podrodiny IL-19, IL-20 a IL-24 k IL-22 a IL-26, které ovliňují imunitní odpověď při infekcích. Celá rodina je příbuzná interferonům (IFN), z nichž několik (interferony λ a v této studii také IFNγ) je zařazováno do FIL-10 z důvodu funkčních a strukturních podobností s členy FIL-10. Interleukiny této rodiny používají k signalizaci několik podjednotek receptorů, které v různých kombinacích a při vazbě různých interleukinů vyvolávají rozdílnou imunitní odpověď. Proteiny FIL-10 jsou produkovány již v evolučně velmi starých organismech jako paryby, proto studie předpokládala, že koevoluce probíhající mezi cytokiny této rodiny a receptory je detekovatelná v sekvencích genů a následně jimi kódovaných proteinů. Pomocí statistických a strukturně biologických metod studie popisuje evoluční vztahy v rámci FIL-10 a ve skupině jejich receptorů, Zajímavé je rozštěpení IFNγ do sekvenčně nezávislých skupin ryb a ostatních. Zaznamenali jsme koevoluci mezi většinou studovaných interleukinů a jejich receptorů, s výjimkou některých ligandů IL10RB, nejšíře používaného receptorového řetězce v rámci studované rodiny, především v interakci IL-26 a IL10RB a interakcích proteinů s podobnou funkcí divergujících později v evoluci , IL-19 a IL-20.

KLÍČOVÁ SLOVA

rodina interleukinu 10, interferon gamma, koevoluce, ligand-receptor, Bayesiánské metody

ABSTRACT

Interleukin 10 family (FIL-10) is an important family of cytokines triggering immune response of different outcome, from antiinflammatory factor interleukin (IL) 10 through epithelia related subfamily of IL-19, IL-20 and IL-24, to IL-22 and IL-26 with role in infection immunity. The family is closely related to interferons (IFNs), several of which (IFNλs, and in this study also IFNγ) are commonly placed into FIL-10 for its functional and structural similarities with FIL-10 proteins. FIL-10 interleukins share several receptor subunits, which in different combinations of receptors and interleukin bound result in different immune response. As the family proteins are expressed in as evolutionary old taxa as cartilaginous fish, we presumed a coevolution in the protein family and the corresponding receptors would be detectable in the sequences of genes and subsequently proteins of FIL-10. Using statistical and structure biological methods, evolutionary relations within the group of FIL-10 and group of their receptors were resolved, with notable division of IFNγ into independent groups of fish and the later vertebrates. Coevolution of the ligand receptor combination in FIL-10 was detected in most cases, with exception of some interactions of IL10RB, the most widely used receptor subunit in the family, most notably IL-26 – IL10RB interaction and interactions of proteins with similar function diverging later in evolution, IL-20 and IL-19.

KEYWORDS

interleukin 10 family, interferon gamma, coevolution, ligand-receptor, Bayesian methods

Contents

Abbreviations

aa	Amino acid, amino acids
Abbrev.	Abbreviation
AIC	Akaike information criterion
Aln.	alignment
Avg.	average
BIC	Bayesian information criterion
BLAST	Basic local alignment search tool
FIL-10	Interleukin 10 family
IFN	interferon
IFNGR1	Interferon gamma receptor 1
IFNGR2	Interferon gamma receptor 2
IFNLR1	Interferon lambda receptor 1
IL	interleukin
IL10RA	Interleukin 10 receptor 1, IL-10 receptor α
IL10RB	Interleukin 10 receptor 2, IL-10 receptor β
IL20RA	Interleukin 20 receptor 1, IL-10 receptor α
IL20RB	Interleukin 20 receptor 2, IL-10 receptor β
IL22BP	Interleukin 22 binding protein
IL22RA1	Interleukin 22 receptor α 1
IL22RA2	Interleukin 22 receptor α 2, also IL22BP
mDCs	Myeloid dendritic cells
ML	maximum likelihood
MSA	Multiple sequence alignment
NJ	Neighbour joining
NK cell	Natural killer
PDB	Protein Data Bank

pDCs	Plasmacytoid dendritic cells
TLR	Toll-like receptor
UID	Unique identifier
UPGMA	Unweighted pair group method with arithmetic mean
WGD	Whole genome duplication
WSP	weighted sums of pairs

1 Introduction

1.1 Interleukin 10 family

Interleukins, in general, trigger immune response of different types by activating signalling pathways. Interleukin 10 is a representative of an important and large family of interleukins with mainly antiinflammatory function, but the family also includes members with antiviral function. Interleukin 10 homologues have been found in as evolutionary old organisms as cartilaginous fish demonstrating thus a high level of conservation of their function and structure. Nevertheless, complexity of immune response is increasing in evolution, thus in lower vertebrates we can only find a few representatives of interleukin 10 family.

Interleukin 10 family proteins in general have a 5 exon, 4 intron genomic pattern. Conserved six-helical structure often forms homodimers. It has been shown that IL-10 has important structural similarities with interferon (IFN) γ (Zdanov et al. 1995), for which IFN γ has been also included in this study.

Interleukins of the family of interleukin 10, called hereafter FIL-10, signal through binding to cellular membrane receptors, whose combination implies their different functions. As with many signalling pathways in immune system, FIL-10 proteins use in many cases signal transfer through variants of Janus kinase – Signal Transducer and Activator of Transcription – (JaK/STAT) pathway. Regulation of transcription in such cases is yet not well described, although it is well known that effects of signal transduction of JaK/STAT pathways may be either of pro-inflammatory or antiinflammatory character.

Proteins of this family have very diverse biological effects, dependent on the type of interleukin as well as target tissue. The effects include immunosuppressive IL-10 and skin and mucosal immunity related IL-20 and IL-24 or proteins involved in antiviral response. Receptor subunits are often shared by several members of FIL-10 (Figure 1), which implies different signalling functions in diverse tissues and ligand-receptor pairs.

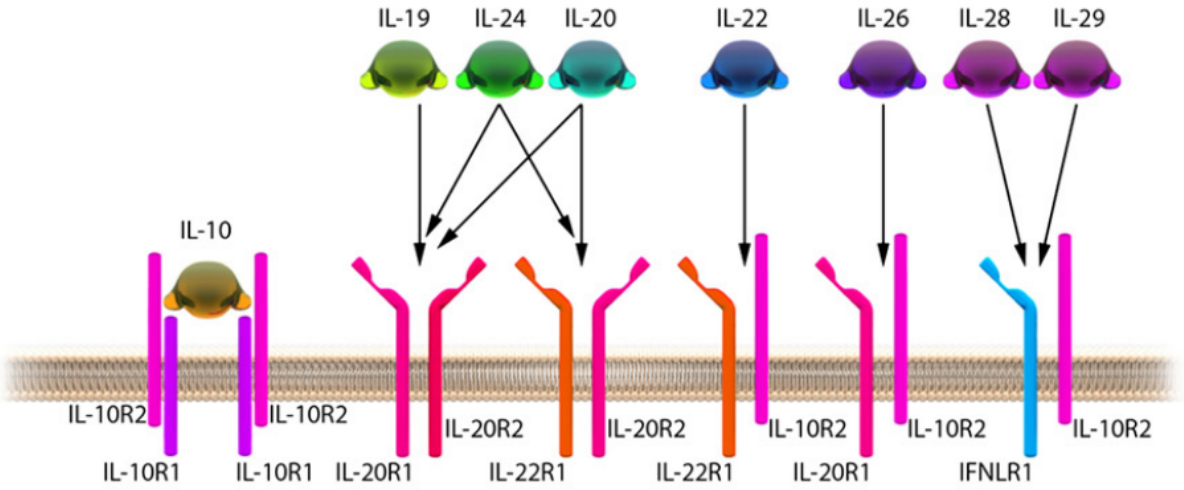

Figure 1. Combinations of interleukin 10 family proteins and their receptors. Subunits marked as 1 are referred to as A (e.g. IL10R1 = IL10RA), "2" subunits are marked B in the thesis. Image from Akdis et al. (2011) (edited).

1.1.1 Interleukin 10

Interleukin 10 (IL-10) is homodimeric protein in length of 178 amino acids in each subunit (human IL-10). IL-10 is an important factor in antiinflammatory response and immune suppression.

IL-10 is produced by multiple immune cell types, major producers of IL-10 are monocytes, regulatory T and regulatory B cells (Del Prete et al. 1993; Holan et al. 2014), it is also secreted by macrophages and myeloid dendritic cells (mDCs), however there is no evidence of its expression in plasmacytoid dendritic cells (pDCs) (Boonstra et al. 2006).

IL-10 is a pleiotropic signalling molecule. It is involved in activation of JAK/STAT pathway, it is capable of blocking NF-κB transcription factor activity and is regulates differentiation of several cell types. One of its main activities is suppression of IFNγ production (Dandrea et al. 1993).

Its signalling through JaK/STAT pathway is not yet well described, since many cytokines share the similar signalling pathway to promote opposite effects (Jones et al. 2016). IL-10 inhibits production of pro-inflammatory IFNγ and IL-2, IL-5, IL-6 and IL-12. By regulation of MHC II expression in monocytes, IL-10 inhibits antigen presentation of microbial peptides (Commins et al. 2008). Treg cells, present throughout

the body, produce IL-10, down-regulating immune response by Th1 and Th17 (Wang et al. 2016).

Maintenance of serum levels of IL-10 is crucial in pathology of several diseases. While deficiency inhibits tumour growth (Wang et al. 2016), impaired regulation leads to several diseases and symptoms. Antiinflammatory function implies role in pathological immune response, as is described primarily in respiratory tract and intestine.

IL-10 has crucial role in maintenance of allergen tolerance. Constant high levels of IL-10 are produced on mucosal tissues mainly in upper and lower airways. Allergen tolerance established by this mechanism is impaired in respiration allergy and asthma patients (Palomares et al. 2010).

Regulation of intestinal epithelial immunity is dependent on IL-10 as well. Mutant variants of both IL-10 gene and its receptor may cause severe colitis (Shah et al. 2012; Kole and Maloy 2014). Recombinant IL-10 has been used in experimental biological treatment of Crohn's disease and colitis (Braat et al. 2006; Marlow 2013).

IL-10 binds to receptor consisting of two IL10RA and two IL10RB chains (Kotenko et al. 1997).

1.1.2 Interleukin 19, interleukin 20 and interleukin 24 subfamily

Interleukins 19, 20 and 24 form protein subfamily, characterized primarily by its monomer helical structure and expression in epithelial tissues. Interleukin 19, or IL-19 is an IL-10 orthologue of 177 amino acids in length (human variant) (Gallagher et al. 2000). Interleukin 20 (IL-20) is a protein in length of 176 amino acids (Blumberg et al. 2001). IL-24, firstly described as melanoma differentiation associated 7 (mda7) is a 206 amino acid protein (Jiang et al. 1996; Caudell et al. 2002). 7-helical monomer structure is typical for the subfamily (Chang et al. 2003).

IL-19, IL-20 and IL-24 modulate inflammatory response in favour of Th2 type (Liao et al. 2004; Wu et al. 2014) and are also further expressed by Th2 cells. Primarily produced by monocytes induced by IL-4 in combination with LPS stimulation (Gallagher et al. 2000) members of this sub-family and their receptors are not only expressed in activated immune cells, but also in similar amounts in keratinocytes (Kunz et al. 2006).

Furthermore, IL-19 is produced in central nervous system by microglia, for which it also serves as autocrine negative regulator (Horiuchi et al. 2015).

As mentioned above, subfamily of IL-19, IL-20 and IL-22 is expressed in high levels in skin and connective tissues. Proteins of the sub-family have positive effect on wound healing in skin and mucosal tissues, IL-19 effects in skin cell proliferation and wound healing by up-regulation expression of keratinocyte growth factor (Sun et al. 2013).

Similarly to IL-10, IL-19 serum levels and concentration in lungs are increased in asthma patients (Liao et al. 2004). Connection between kidney disease or injury and increased levels of IL-19 has been recently pointed out, but details of IL-19 function in kidney have not been described (Jennings et al. 2015). IL-19 is known to be involved in pathology of systemic lupus erythematodes (Lin et al. 2016) and together with other members of the sub-family, also in psoriasis (Otkjaer et al. 2005; Wang et al. 2012).

IL-24, apart from its immunity effects in skin, has specific tumor-suppresor activity. Through IL-20 receptors, independently of JaK/STAT pathway, IL-24 induces effectively apoptosis specifically to tumour cells (Zheng et al. 2006) by inducing autophagy, which in later stages switches to apoptosis (Yang et al. 2010; Bhutia et al. 2010).

The subfamily signals through different combinations of shared receptor subunits (some members are also shared with closely related IL-22 and IL-26, see Figure 1). IL-19 signals through IL20RA/IL20RB heterodimer. IL-20 binds to IL20RA/IL20RB heterodimer, nevertheless it is also able to signal through IL22RA/IL20RB heterodimer (Commins et al. 2008). IL-24 the uses same two receptor pairs as IL-20 (Wang et al. 2002). IL-20 and IL-24 N-terminus forms a β-hairpin structure distincting them in binding to receptor subunits from IL-19. Affinity to receptor chains is defined by their secondary structure (Logsdon et al. 2012).

1.1.3 Interleukin 22

IL-22 is 6 antiparallel helices protein with 179 amino acids. IL-22 is produced by activated immune cells, NK-22 (Cella et al. 2009), Th17 cells (Liang et al. 2006) and Th22 cells (express IL-22 without IL-17) (Duhen et al. 2009). IL-22 molecules secreted

by NK-cells in lymphoid associated mucosal tissues (tonsils, Peyer's patches) provide innate protection against bacterial and viral infection (Cella et al. 2009).

As in its relatives, it was described that IL-22 facilitates wound healing in mucosal tissue by increase of fibroblast activity (McGee et al. 2013). Dysregulation in IL-22 signalling is observed in development of psoriasis and atopic dermatitis symptoms (Ma et al. 2008) and promotes hyperplasia of epidermis (Zheng et al. 2007).

Interleukin 22 binds to IL22RA/IL10RB heterocomplex (Kotenko et al. 2001). Soluble Interleukin 22 Binding Protein (IL22BP) competitively inhibits IL-22 activity (Xu et al. 2001).

1.1.4 Interleukin 26

Interleukin 26 or IL-26 is 171 amino acids long (human) and has similar 6 helices structure to IL-19,20 and 24 family, however unlike them, forms a homomer. IL-26 is conserved throughout mammals, interestingly is lacking in mice and rats, it is present also in amphibian, bird (Donnelly et al. 2010) and fish (Igawa et al. 2006) models. Due to its absence in the most important model organisms – mice and rats, information about its *in vivo* function is quite limited.

IL-26 is produced by Th17 alongside with the other members of IL-10 family (Wilson et al. 2007). IL-26 main function seems to be in antimicrobial defence and Th17 antimicrobial activity is hugely dependent on IL-26 production. IL-26 activity in infection is based on its ability to form pores on extracellular bacteria membranes. IL-26 recognizes bacterial DNA released in lysis and forms insoluble complexes, afterwards presented, independently of IL-26 receptor, to pDCs to stimulate production of type I interferon (IFNα) (Meller et al. 2015).

Besides of its activity in innate infection, IL-26 signals through IL10R2 and IL20R1 combination of receptor subunits activating JaK/STAT pathway in epithelia. Pathway activation by IL-26 may be inhibited by heparin (Hor et al. 2004).

1.1.5 Interferon λ family

Interferon lambda (IFNλ) group is a group of 3 proteins in human consisting of IL 29 – Interleukin 29 / IFNλ1 – Interferon lambda 1, IL-28A or IFNλ2 and IL-28B (IFNλ3). Known as IL-28 and IL-29, IFNλs are commonly seen as IL-10 family members, due to their (monomeric) structure, in particular to its resemblance of IL-22 structure and function – antiinfection immunity (Gad et al. 2009). Members of this subgroup are further referred to as IFNλs, not ILs (IFNλ1, IFNλ2, IFNλ3) in this thesis. Similarity with IFN type I is presumably only of functional, not genetic character, as gene structure of IFNλ matches IL-10 5 exon – 4 intron pattern (Kotenko et al. 2003).

IFNλ group is categorized as type III interferons. Interferon-typical antiviral activity lies in blocking viral replication. IFNλs is known primarily for its activity against hepatitis viruses – it is responsible for inducing immune response against Hepatitis C (HCV) virus in liver (Marukian et al. 2011) predominantly by inhibition of both HBV and HCV replication (Robek et al. 2005). HCV clearance from organism is hugely dependent on IFNλ genotype (Sheahan et al. 2014).

Antiviral activity of IFNλ is not limited to hepatitis viruses. IFNλ is expressed together with IFN type I (IFNα) by wide variety of cell types after stimulation of toll-like receptor (TLR)3 and TLR9 (Ank et al. 2008). IFNλ is able to act against coronaviruses (Hamming et al. 2013) or norovirus (Baldridge et al. 2015; Nice et al. 2015) and possibly many other viruses.

IFNλ uses the complex of IFNRL1/IL10RA chains for immune signalling (Kotenko et al. 2003). IL10RA, being used by IL-10, IL-22 and IL-26 as well, is another similarity supporting IFNλ assignment into the family.

1.1.6 Interferon γ

Long known IFN type II group member was first discovered in 1960s and described to have antiviral activity. IFNγ occurs in homodimers with 6-helical structure (Ealick et al. 1991) similar to IL-10. The protein has 166 amino acids in length, the encoding gene has 4 exons and 3 introns.

Pro-inflammatory IFNγ is continuously produced by NK and NKT cells in innate immune system. NKT cells posses characteristics of both NK (many of inhibitory and activation receptors) and T cells (TCR with CD3 coreceptor, expression of CD4 coreceptor) and are activated after TCR recognition of antigen. IFNγ further stimulates NK cell activity. Production of IFNγ in innate system seems to be regulated by IL-12 and IL-18 (Schoenborn and Wilson 2007).

Th1 and cytotoxic T-cells produce IFNγ as a part of adaptive response to infection. Being the main chemokine product of Th1, its effects predefine character of Th1 response as cytotoxic activity targeted mainly against intracellular parasites (Schoenborn and Wilson 2007).

IFNγ is critical in protection from intracellular bacteria and viruses. Antiviral activity of IFNγ lies in induction of expression of antiviral enzymes, such as dsRNA adenosine deaminase (Patterson et al. 1995). Through JaK/STAT activation, IFNγ is capable of inducing apoptosis (Chin et al. 1997). Polymorphism in IFNγ and IL-26 gene cluster has been shown to impact susceptibility to rheumatoid arthritis (Vandenbroeck et al. 2003). Besides, IFNγ is involved in tumour immunity.

IFNγ signaling is induced by its binding to IFNγ receptor 1 (IFNGR1) and subsequent association with IFNγ receptor 2 (IFNGR2). The ternary complex then activates the JaK/STAT pathway (Akdis et al. 2011).

1.1.7 Receptors and signalling

Regulation of cytokine signaling is dependent on their interactions with the receptor chains (Figure 2). Their expression is different in different tissues. Most receptors involved in signaling of IL-10 family are expressed in immune cells, primarily T, B and NK cells. In addition, their expression has been also detected in skin, liver or pancreas (Wolk et al. 2002). IL20RB is expressed in keratynocytes (Wolk et al. 2004). IL20RA expression, unlike other receptors, has not been detected in NK cells, T, B cells or monocytes, but it is present in high levels in skin (Wolk et al. 2002). The group of FIL-10 receptors includes IFNL1 or IL28RA interacting with SH2 domain of JaK proteins (Zhang et al. 2016b).

Several defects in signalling caused by mutations in receptor subunits are known. IL10RB mutation causes defect in JaK/STAT signalling connected with inflammatory bowel diseases such as Crohn's disease or colitis in adults (Glocker et al. 2009; Chaudhry et al. 2011) and specific mutation in IL-10RA causes very early onset inflammatory bowel disease in children (Shim and Seo 2014).

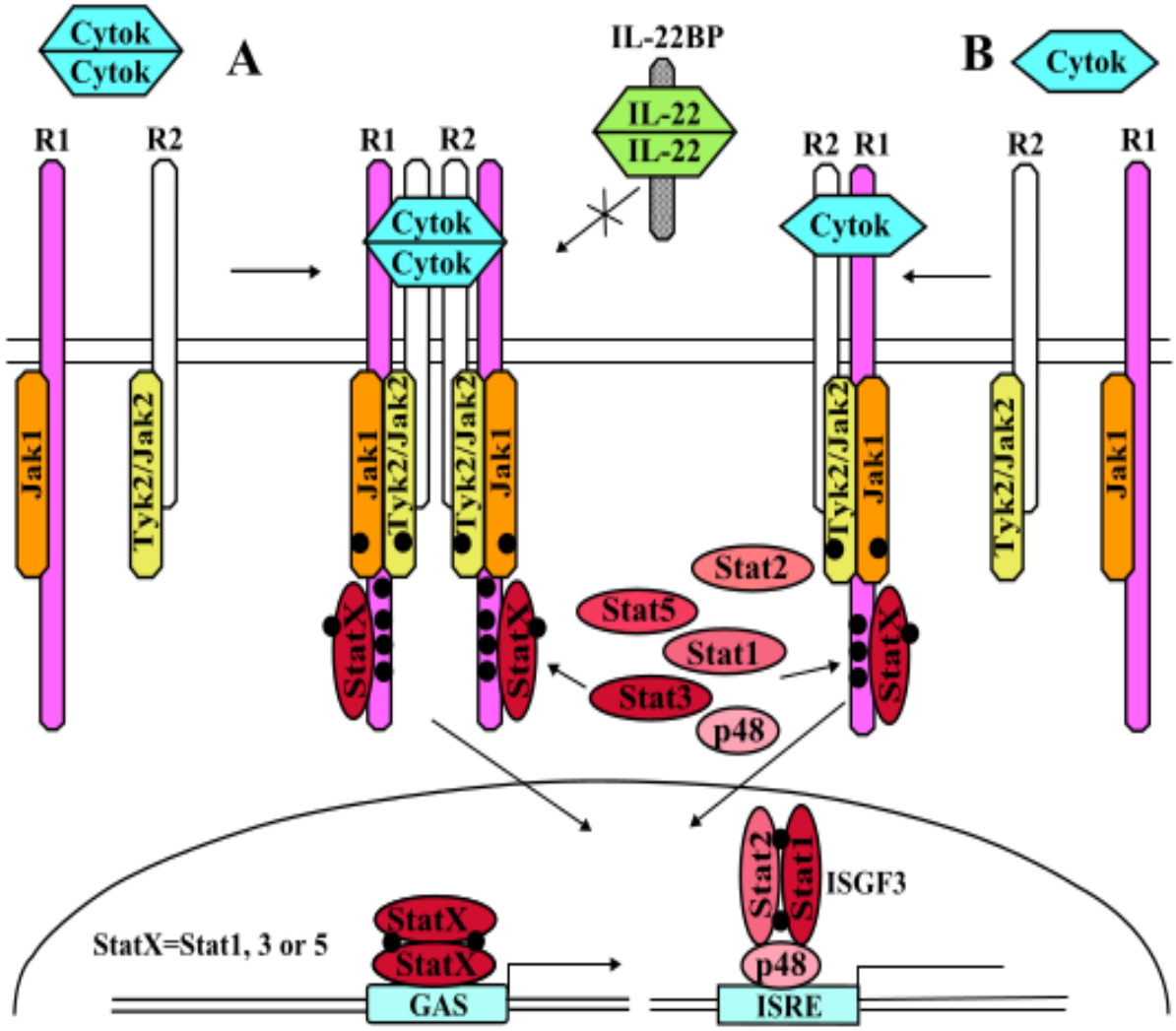

Figure 2. Receptor binding and signalling in monomeric and dimeric proteins of IL-10 family.
R1-type chain proteins: IL-10R1, IL-20R1, IL-22R1 and IFN-gR1. R2-type chain proteins IL-10R2, IL-20R2 and IFN-gR2. Image from Kotenko and Langer (2004).

Interleukin 22 binding protein

Interleukin 22 receptor A2, (IL22RA2) or interleukin 22 binding protein (IL22BP) is a soluble antagonist of IL-22 sharing amino acid homology with IL22RA1. It is highly expressed in placenta and spleen as well as digestive tract and immune system (Xu et al. 2001), particularly in mDCs (Martin et al. 2014).

IL22BB binds IL-22, but is unable of inducing JaK/STAT signalling (Kotenko et al. 2001), functioning thus as IL-22 neutraliser. IL22BP is involved in regulation of pathological IL-22 response (Martin et al. 2014) and tumour-genesis in intestine (Huber et al. 2012).

Activity of IL22BP is affected by tissue damage sensing. Inflammasome down-regulates production of IL22BP. Disrupted ratio of IL-22 and its antagonist supports tumour development after previous damage by chronic inflammations of colon (Huber et al. 2012).

Interferon γ receptors

Interferon γ signals through 2 receptor subunits distinct from IL-10 receptors, interferon gamma receptor 1 (IFNGR1) and interferon gamma receptor 2 (IFNGR2). Interferon gamma receptor genetic variants are involved in various diseases.

IFNGR1 is likely involved in intestinal and colorectal carcinoma formation (Wang et al. 2015; Zhang et al. 2016a). The IFN γ receptor promotes antiviral activity against hepatitis viruses (Lam et al. 2014), against bacterial tuberculosis (Sahiratmadja et al. 2007) and against Dengue Virus in connection with interferons type I (Prestwood et al. 2012) and . As a result of viral coevolution with the host, some viruses developed mechanisms to avoid immune response by interferon γ pathway. For instance, herpesviruses are able to suspend expression of IFNGR1 (Li et al. 2007).

1.1.8 Gene clustering in Interleukin 10 family

In vertebrate genomes, interleukin genes commonly occur in clusters. This fact may be consequence of gene duplications in evolution and it effects regulation of gene expression.

Human IL-10, IL-19, IL-20 and IL-24 genes are located on chromosome 1q31-32 (Kim et al. 1992; Blumberg et al. 2001). Human IFNγ, IL-26 and IL-22 genes are located on chromosome 12q15 (Donnelly et al. 2010). IL10RB and IFNGR2 are encoded on chromosome 21q33. IL20RA, IL22BP and IGNFR1 are clustered on human chromosome 6q23 (data from UCSC and NCBI databases). Vertebrate homologues of the mentioned groups are clustered in genome as well.

1.2 Coevolution of ligand-receptor pairs

Protein coevolution (correlated evolution) is a fundamental principle of evolution and occurs in every organism or group of organisms. By means of coevolution of protein and receptor pairs, signalling pathways are preserved despite changes in their genetic information and consequently the protein structures. The correlation of evolution of many ligand-receptor pairs in immune signalling molecules has been previously pointed out (Goh et al. 2000).

Coevolution may be studied on different levels, from inter-organism level to inter-residual relations. Current methods of protein coevolution analysis are based on comparison of distance matrices (Figure 3). For inter-protein evolution, distance matrices calculated from phylogenies are used, whereas inter-amino acid studies require a MSA for calculation of the distance matrix (de Juan et al. 2013).

Characterization of phylogenetic history of a protein family may elucidate foundation of interleukin diversity as known in human, as well as origins of combinations of particular ligand-receptor pairs, since complete understanding of signal transduction from particular interleukin 10 family members is still incomplete.

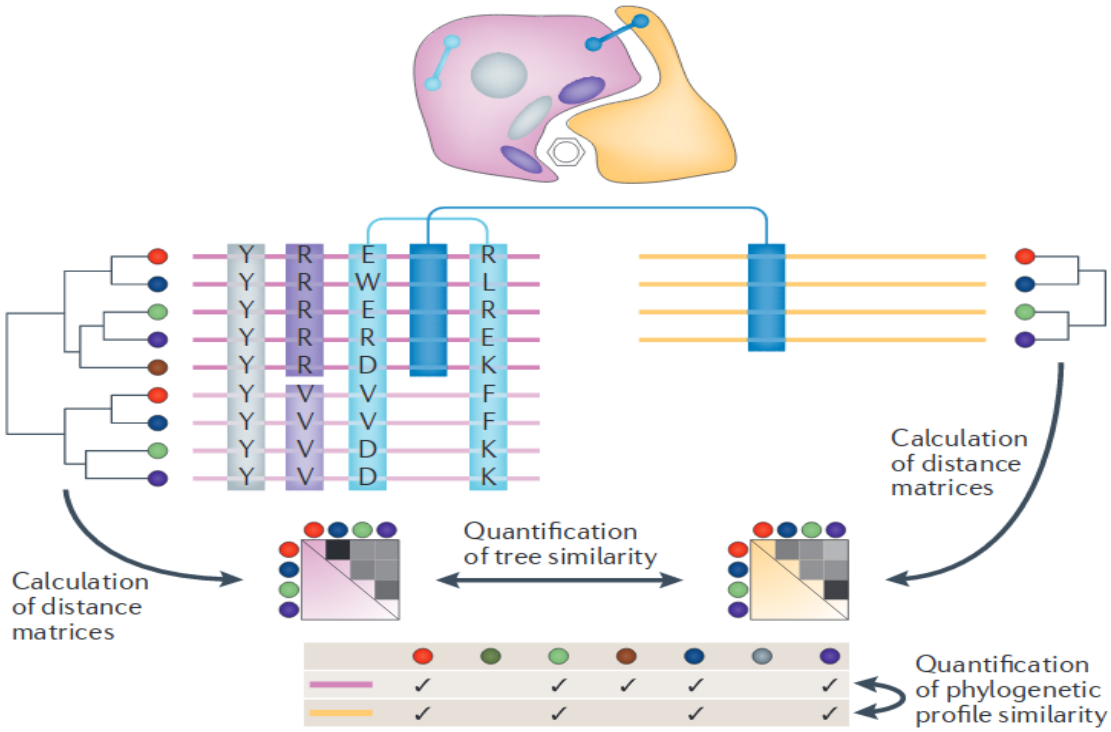

Figure 3.Analysis of protein receptor coevolution. Distance matrix is calculated from a phylogeny or multiple sequence alignment and analysed for pairwise correlation in protein evolution. Image from de Juan et al. (2013).

1.3 Introduction to phylogenetic reconstruction

Characterization of phylogenetic history of interleukin 10 family is valuable for understanding of the divergence of signalling pathways in the family and for general understanding of immunity evolution. Evolutionary trees of the family and their receptors may elucidate foundation of the interleukin diversity as known in human, as well as origins of ligand–receptor combination and their specificity.

Phylogenetic tree is graphical representation of relationships within a group of organisms or proteins. Unrooted trees only present relationships within the family, may be rooted using outgroup or molecular clock to show the common ancestor.

Phylogenetic trees may be constructed by several methods using different statistical approaches. Distance based methods use distance matrix of sequences in alignment. The methods include neighbour joining (NJ) trees, finding least related taxa and recalculating their tree nodes and branches, or UPGMA (Unweighted pair group method with arithmetic mean) trees, using hierarchical clustering based on pairwise similarity

matrix creating rooted trees. Methods using optimality criteria include maximum parsimony, maximum likelihood (ML), or Bayesian. Methods implement evolutionary models to establish the phylogeny. Maximum parsimony trees use the principle of Occam's razor. As the name implies, the lowest possible number of evolutionary events is conducted to create the tree (Mount 2008).

Maximum likelihood (Whelan and Goldman 2001) methods are based on calculation of likelihood of parameters. Likelihood corresponds to probability of obtaining recorded data under given model. The method is quite time consuming and dependent on computational resources. The calculation of likelihood is rather straightforward, yet the estimation of parameters used in the phylogeny is quite demanding.

Bayesian estimation of phylogeny infers trees based on prior probability. The method incorporates Markov Chain Monte Carlo algorithms to produce distribution of posterior probability of phylogenies. The resulting trees are concatenated from sampling of calculations of trees with high posterior probability (Felsenstein 2004).

Probability based methods such as ML and Bayesian inference should be used for well supported trees. Nonetheless, the overall phylogeny quality depends not only on the selected method, but also largely on sequence alignment quality.

2 Aims

Coevolution of interleukins of the group of IL-10 and corresponding receptor pairs is important for understanding complexity of immune signalling pathway of the family. The thesis aims at describing evolutionary history of FIL-10 and the protein group serving as receptors of FIL-10 proteins.

The particular aims of the thesis are:

- to construct phylogenies of both interleukin 10 family proteins and corresponding receptor protein groups
- using statistical methods, describe correlation between evolution of the interleukins and their receptors.
- describe protein conservation on structural level for selected representatives of the interleukin 10 family.

3 Materials and methods

The methods used in this thesis and the processes are illustrated in Figure 4.

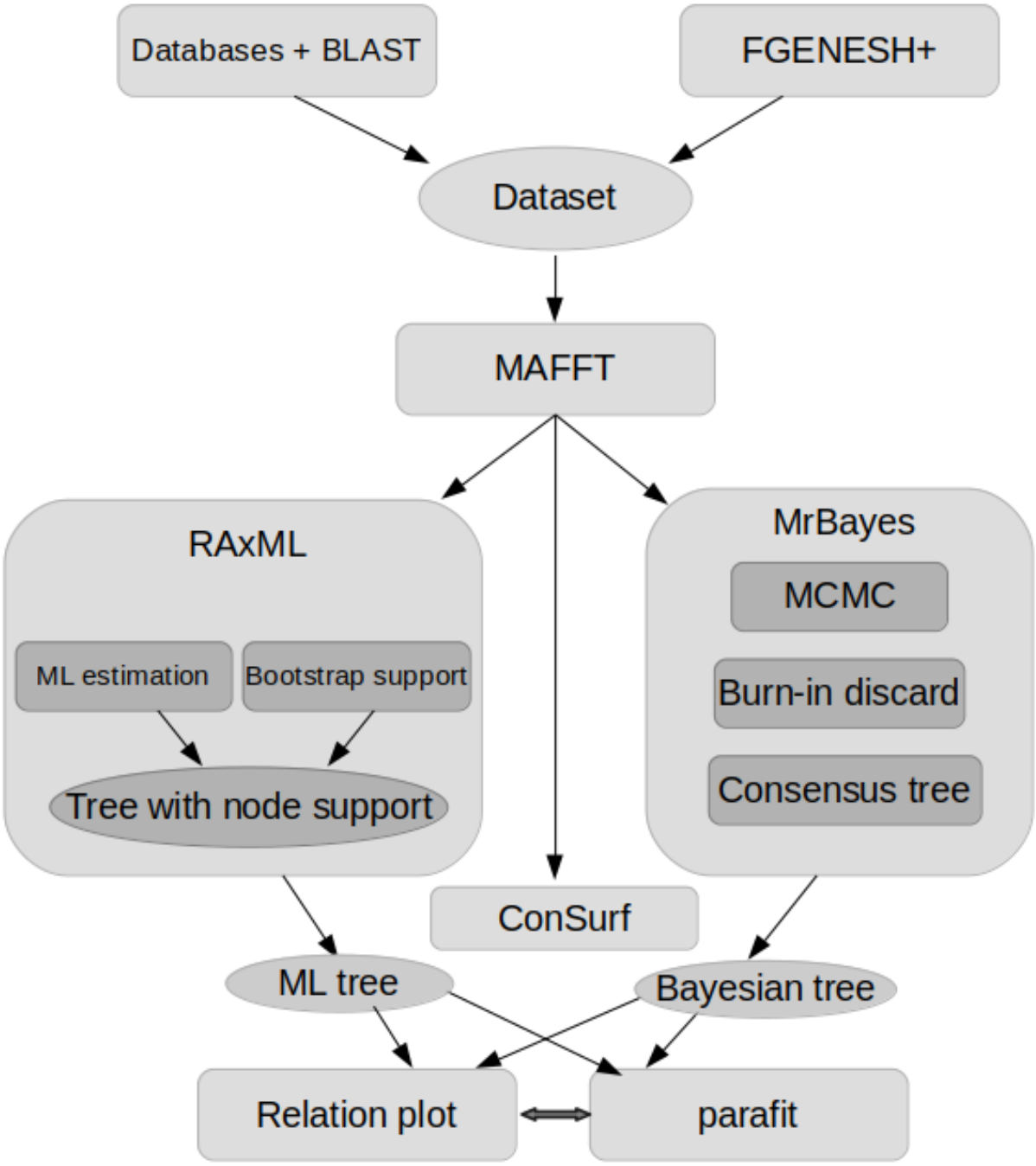

Figure 4. Workflow chart. Data (ellipses) obtained by methods (rectangular) on top are further analysed by RAxML software and MrBayes software. The coevolutionary analysis is preformed in Parafit tool of R and compared to relation plot of interleukins and corresponding receptors created by tools available in ape package in R. Based on the data, I performed conservation analysis in ConSurf.

3.1 Data collection

Interleukins of IL-10 family (FIL-10) need to be studied in species from distant taxonomic groups to determine their relationships. Species were selected based on their evolutionary distance, but also based on completeness and quality of genome sequence. Species from groups with expected changes in interleukin structure were preferred.

Fish species were selected with consideration of the past whole genome duplications (WGD). *Tetraodon nigrovidis* represents teleost fish with 1 WGD (Jaillon et al. 2004) and *Oncorhynchus mykiss* (rainbow trout) was selected as a representative of *Salmonidae* with two WGDs (Berthelot et al. 2014).

Table 1: Binomial and English names of selected species. The species is mentioned as a representative of mentioned taxonomic group.

Binomial name	Taxonomic group	English name
Calorhinchus milii	*Chondrichthyes*	The Australian ghost shark
Danio rerio	*Actinopterygii – Cypriniformes*	The Zebrafish
Oncorhynchus mykiss	*Actinopterygii – Salmoniformes*	The Rainbow trout
Tetraodon nigroviridis	*Actiopterygii – Tetraodontiformes*	The Green spotted puffer
Xenopus tropicalis	*Amphibia*	The Western clawed frog
Alligator mississippiensis	*Crocodylia*	American alligator
Gallus gallus	*Aves*	The red junglefowl/domestic chicken
Ornithorhynchus anatinus	*Mammalia – Prototheria*	The Platypus
Mus musculus	*Mammalia – Rodentia*	The house mouse
Homo sapiens	*Mammalia – Primates*	Modern human

3.1.1 Database search and BLAST

The NCBI protein database at https://www.ncbi.nlm.nih.gov/protein/ and UniProt protein database at http://www.uniprot.org/ were queried by the binomial name of selected model species (Table 1) and name of selected protein.

Sequences of non-annotated proteins were searched useing Basic local alignment search tool (BLAST) at https://blast.ncbi.nlm.nih.gov/Blast.cgi. Protein databases were queried with the evolutionary closest homologous protein sequence available from database search using BLASTp, delta-BLAST and tBLASTn algorithms.

Desired protein sequences were exported from databases in FASTA format.

3.1.2 FGENESH+

FGENESH+ (Solovyev 2001) is a prediction tool finding exon sequences in given genomic DNA. In this study, FGENESH+ was used for prediction of exon and protein sequences in unannotated genomes, particularly in evolutionary older species, where genomes of non-model species are not annotated completely or correctly.

FGENESH+ uses hidden Markov model and similar protein sequence to predict protein homologues in eukaryotes. Required inputs for the program are: a homologous protein sequence from evolutionary close species, species of prediction and genomic sequence of the species, that based on clustering of proteins should include coding sequence of the predicted protein. FGENESH+ is limited in ability to process genomic sequences longer than ~200 000 bp, therefore information about clustering of genes is necessary for successful prediction.

3.1.3 Data formatting and processing

Data from databases were downloaded in FASTA format and alignments were preformed in FASTA format. For further analysis, data was converted to NEWICK format and PHYLIP formats using ReadSeq tool (Gilbert 2002). Complete data used for the analysis is presented in Tables 2, 3 and 4, with unique identifiers (UID) of the sequences in databases.

UGENE (Okonechnikov et al. 2012) and Jalview (Waterhouse et al. 2009) were used for data visualization and control The dataset was aligned and checked for error and duplicate sequences. For further visualisation of trees, R package ape (Paradis et al. 2004) or FigTree at http://tree.bio.ed.ac.uk/ software was used.

3.1.4 Final dataset

Table 2. **Interleukin sequences from databases.** Table mentions binomial species name, protein name as stated in the database, database from which sequence has been obtained and sequence UID.

Alligator mississippiensis	interleukin-10	NCBI	1011565932	Mus musculus	Interleukin-19	UniProt	Q8CJ70
Alligator mississippiensis	interleukin-20	NCBI	1011565931	Mus musculus	Interleukin-20	UniProt	Q9JKV9
Alligator mississippiensis	interleukin-22	NCBI	1011613344	Mus musculus	Interleukin-22	UniProt	Q9JJY9
Alligator mississippiensis	Interferon lambda-3	UniProt	A0A151N9Y8	Mus musculus	Interleukin-22b	UniProt	Q9JJY8
Alligator mississippiensis	interferon gamma	NCBI	1011613342	Mus musculus	Interleukin-24	UniProt	Q925S4
Callorhinchus milii	interleukin-10	NCBI	632963185	Mus musculus	Interferon lambda-2	UniProt	Q4VK74
Callorhinchus milii	interleukin-20-like	NCBI	632963320	Mus musculus	Interferon lambda-3	UniProt	Q8CGK6
Callorhinchus milii	interleukin-22	NCBI	632973502	Mus musculus	interferon gamma	NCBI	237845681
Callorhinchus milii	interferon gamma	NCBI	632973500	Oncorhynchus mykiss	interleukin-10	NCBI	47678893
Danio rerio	Interleukin 10	NCBI	190337256	Oncorhynchus mykiss	Interleukin 20	NCBI	311771762
Danio rerio	interleukin-20	NCBI	130508100	Oncorhynchus mykiss	interleukin 22	NCBI	242098052
Danio rerio	interleukin 22	NCBI	66472836	Oncorhynchus mykiss	interferon-lambda	NCBI	209972108
Danio rerio	Interleukin 26	NCBI	190339888	Oncorhynchus mykiss	interferon gamma	NCBI	56291619
Danio rerio	interferon-lambda	UniProt	A0FJI5	Ornithorhynchus anatinus	interleukin-10	NCBI	620945898
Danio rerio	interferon gamma	NCBI	40363745	Ornithorhynchus anatinus	interleukin-19	UniProt	F6SIJ2
Gallus gallus	interleukin-10	NCBI	51173888	Ornithorhynchus anatinus	interleukin-20	NCBI	345320613
Gallus gallus	interleukin-20	NCBI	118102427	Ornithorhynchus anatinus	interleukin-22	NCBI	149632275
Gallus gallus	interleukin-22	NCBI	571255083	Ornithorhynchus anatinus	interleukin-26	NCBI	149632273
Gallus gallus	interleukin 28A	NCBI	184186388	Ornithorhynchus anatinus	interferon lambda-3	NCBI	345314447
Gallus gallus	Interferon lambda-3	UniProt	B4ER10	Ornithorhynchus anatinus	interferon gamma	NCBI	620958883
Gallus gallus	interferon-gamma	NCBI	27549285	Tetraodon nigroviridis	interleukin 10	NCBI	29125864
Homo sapiens	Interleukin-10	UniProt	P22301	Tetraodon nigroviridis	interleukin-20	NCBI	31747223
Homo sapiens	Interleukin-19	UniProt	Q9UHD0	Tetraodon nigroviridis	interleukin-24	NCBI	31747227
Homo sapiens	Interleukin-20	UniProt	Q9NYY1	Tetraodon nigroviridis	Interferon gamma	NCBI	28475279
Homo sapiens	Interleukin-22	UniProt	Q9GZX6	Xenopus tropicalis	interleukin-10	NCBI	284813591
Homo sapiens	Interleukin-24	UniProt	Q13007	Xenopus tropicalis	interleukin-20	NCBI	847096564
Homo sapiens	Interleukin-26	UniProt	Q9NPH9	Xenopus tropicalis	interleukin-22	NCBI	213983241
Homo sapiens	Interferon lambda-1	UniProt	Q8IU54	Xenopus tropicalis	interleukin-26	NCBI	212549595
Homo sapiens	Interferon lambda-2	UniProt	Q8IZJ0	Xenopus tropicalis	interferon lambda5	NCBI	256860238
Homo sapiens	Interferon lambda-3	UniProt	Q8IZI9	Xenopus tropicalis	interferon lambda2	NCBI	256860234
Homo sapiens	Interferon lambda-4	UniProt	K9M1U5	Xenopus tropicalis	interferon lambda3	NCBI	847153596
Homo sapiens	Interferon gamma	UniProt	P01579	Xenopus tropicalis	interferon lambda1	NCBI	256860232
Mus musculus	Interleukin-10	UniProt	P18893	Xenopus tropicalis	interferon gamma	NCBI	301618299

Table 3. Receptor sequences from databases. Table mentions binomial species name, protein name as stated in the database, database from which sequence has been obtained and sequence unique identifier UID.

Alligator mississippiensis	Interleukin-10 receptor alpha	UniProt	A0A151NFH0	Mus musculus	interferon gamma receptor 1	NCBI	148671505
Alligator mississippiensis	interferon gamma receptor 1	NCBI	1011562648	Mus musculus	interferon gamma receptor 2	NCBI	148671877
Alligator mississippiensis	interferon gamma receptor 2	NCBI	1011567279	Mus musculus	Protein Il20rb	UniProt	E9Q9A6
Alligator mississippiensis	interferon lambda receptor 1	NCBI	1011571573	Mus musculus	Interleukin-10 receptor beta	UniProt	Q61190
Alligator mississippiensis	interleukin-20 receptor beta	UniProt	A0A151LYK3	Mus musculus	Interleukin-10 receptor alpha	UniProt	Q61727
Alligator mississippiensis	Interleukin-20 receptor alpha	UniProt	A0A151M3P1	Mus musculus	Interleukin-20 receptor alpha	UniProt	Q6PHB0
Alligator mississippiensis	Interleukin-10 receptor beta	UniProt	A0A151ME26	Mus musculus	interleukin-22 receptor alpha-2	UniProt	Q80XF5
Alligator mississippiensis	interleukin-22 receptor alpha-1	UniProt	A0A151MML2	Mus musculus	interleukin-22 receptor alpha-1	UniProt	Q80XZ4
Callorhinchus milii	interleukin-20 receptor beta	NCBI	632934306	Mus musculus	Interferon lambda receptor 1	UniProt	Q8CGK5
Callorhinchus milii	interferon gamma receptor 1	NCBI	632951240	Oncorhynchus mykiss	interferon-gamma receptor 2	NCBI	166406457
Callorhinchus milii	Interleukin-22 receptor alpha-1	UniProt	V9KQD1	Oncorhynchus mykiss	interferon-gamma receptor alpha	NCBI	185132696
Callorhinchus milii	Interleukin-10 receptor alpha-like	UniProt	V9KUB8	Oncorhynchus mykiss	interleukin-20 receptor alpha	NCBI	185133176
Callorhinchus milii	Interleukin-10 receptor beta	UniProt	V9KWB3	Oncorhynchus mykiss	Interleukin-10 receptor beta	NCBI	526252816
Callorhinchus milii	Interleukin-20 receptor alpha	UniProt	V9LAR4	Oncorhynchus mykiss	Interleukin-10 receptor alpha	NCBI	642084689
Danio rerio	interleukin 10 receptor beta-like	NCBI	76563837	Oncorhynchus mykiss	interleukin-20 receptor beta	NCBI	642096911
Danio rerio	interleukin-22 receptor alpha-2	NCBI	113674671	Oncorhynchus mykiss	interferon lambda receptor 1	NCBI	642126604
Danio rerio	interleukin 10 receptor alpha	NCBI	117606403	Oncorhynchus mykiss	IL-22 binding protein	UniProt	K0J8Z9
Danio rerio	Interferon gamma receptor 1	NCBI	190338988	Ornithorhynchus anatinus	interferon gamma receptor 1	NCBI	620942285
Danio rerio	interleukin-20 receptor beta	NCBI	300490528	Ornithorhynchus anatinus	interleukin-20 receptor beta	NCBI	620955148
Danio rerio	interferon lambda receptor 1	NCBI	308275358	Ornithorhynchus anatinus	interferon lambda receptor 1	NCBI	620962825
Gallus gallus	interferon gamma receptor 2	NCBI	56711284	Ornithorhynchus anatinus	Interleukin-10 receptor beta	NCBI	620974654
Gallus gallus	interleukin 10 receptor 1	NCBI	83999156	Ornithorhynchus anatinus	interferon gamma receptor 2	NCBI	1018961860
Gallus gallus	Interleukin-10 receptor beta	NCBI	84618077	Ornithorhynchus anatinus	Interleukin-10 receptor alpha	UniProt	F6SGX7
Gallus gallus	interferon gamma receptor 1	NCBI	158420743	Ornithorhynchus anatinus	interleukin-22 receptor alpha-1	UniProt	F6UJT9
Gallus gallus	interleukin-22 receptor alpha-1	UniProt	E1BRV0	Ornithorhynchus anatinus	Interleukin-20 receptor alpha	UniProt	F6VJN7
Gallus gallus	Uncharacterized protein	UniProt	E1BW22	Ornithorhynchus anatinus	interleukin-22 receptor alpha-2	UniProt	F6VJR3
Gallus gallus	Interleukin-20 receptor alpha	UniProt	F1NYV0	Tetraodon nigroviridis	helical cytokine receptor CRFB8	NCBI	28475293
Gallus gallus	interleukin-22 receptor alpha-2	UniProt	F1NYV1	Tetraodon nigroviridis	interferon gamma receptor alpha	NCBI	337729935
Gallus gallus	Interferon lambda receptor 1	UniProt	K9JA28	Tetraodon nigroviridis	interferon lambda receptor 1	UniProt	H3C6M4
Homo sapiens	Interferon gamma receptor 1	UniProt	P15260	Tetraodon nigroviridis	interleukin-20 receptor beta	UniProt	H3CAT4
Homo sapiens	Interferon gamma receptor 2	UniProt	P38484	Tetraodon nigroviridis	interleukin-22 receptor alpha-2	UniProt	H3DHY6
Homo sapiens	Interleukin-10 receptor beta	UniProt	Q08334	Tetraodon nigroviridis	Interleukin-20 receptor alpha	UniProt	Q7ZT35
Homo sapiens	Interleukin-10 receptor alpha	UniProt	Q13651	Xenopus tropicalis	interferon gamma receptor 1	NCBI	195540123
Homo sapiens	Interleukin-20 receptor beta	UniProt	Q6UXL0	Xenopus tropicalis	Interleukin-10 receptor beta	NCBI	284521656
Homo sapiens	Interferon lambda receptor 1	UniProt	Q8IU57	Xenopus tropicalis	Interferon lambda receptor 1	NCBI	284795282
Homo sapiens	interleukin-22 receptor alpha-1	UniProt	Q8N6P7	Xenopus tropicalis	interferon gamma receptor 2-like	NCBI	512824041
Homo sapiens	interleukin-22 receptor alpha-2	UniProt	Q969J5	Xenopus tropicalis	interleukin-20 receptor beta	NCBI	512866588
Homo sapiens	Interleukin-20 receptor alpha	UniProt	Q9UHF4	Xenopus tropicalis	interleukin-22 receptor alpha-1	NCBI	847101222
				Xenopus tropicalis	Interleukin-10 receptor alpha	UniProt	F7C5L9

Table 4. FGENESH+ predicted protein sequences with UIDs of homologue proteins used for prediction. Predicted sequences are presented in Supplement 1. Only the marked sequence is used in following analysis.

Protein	Homologue ID, species	Length, exons	Score
Anolis carolinensis IL-20	1011565931 Alligator mississippiensis	174 aa, 5	1282.128223
Gallus gallus IL-26	**558155504 Pelodiscus sinensis**	**178 aa, 4**	**774.493945**
Pelodiscus sinensis IFNLR1	1011571573 Alligator mississippiensis	587 aa, 7	1370.3634

3.2 Multiple sequence alignment (MSA)

Critical step in phylogenetic reconstruction is multiple sequence alignment. Quality of the tree is hugely dependent on the quality of alignment, therefore high-quality alignment with well-defined positional homology is a necessary prerequisite. Several different tools for alignment of higher number of divergent sequences are commonly used.

MAFFT, MUSCLE, ClustalW, Clustal OMEGA and T-Coffee, represent some of widely used tools for MSA. As the alignment quality is essential for phylogenetic inference, error sequences and overall quality of the alignment needs to be manually checked after aligning sequences. In this study, I used MAFFT tool to align obtained sequences.

3.2.1 MAFFT

MAFFT (multiple alignment by fast Furier transform) uses identification of homologous regions by fast Furier transform (FFT). MAFFT is claimed to be quicker than T-Coffee or ClustalW (Katoh and Standley 2013) while preserving high accuracy. FFT identifies homologous regions by grouping of amino acids in sequence by chemical and physical characteristics (Katoh et al. 2002). Afterwards, similarly to many other alignment tools, MAFFT refines the alignment to create better results. MAFFT uses 2 types of scores to determine the quality of alignment – WSP (weighted sums of pairs) and consistency score (also called *importance value* as described by Katoh et al. (2005)).

MAFFT is used in desktop version 7.307 (Katoh and Standley 2013). The `--globalpair` option for alignment of sequences of similar length is used. Pairwise alignments are computed with Needleman-Wunsch algorithm for global alignment (Needleman and Wunsch 1970). Maximum number of iterative refinement cycles is set to recommended 1000 (`--maxiterate 1000`). MAFFT uses BLOSUM62 scoring matrix for amino acid sequences by default.

After aligning the sequences, I used trimAl tool (Capella-Gutierrez et al. 2009) on web interface on Phylemon2 web server at http://phylemon2.bioinfo.cipf.es/ with gappyout option to remove columns unsuitable for further analysis.

3.3 Amino acid evolution model selection

Inference of phylogeny requires selection of best-fitting model of amino acid evolution for particular data. The sequence alignment was analysed by ProtTest 3 software (Abascal et al. 2005; Darriba et al. 2011).

Selection was based on Akaike information criterion (AIC)

$$AIC = 2p - 2\ln(L)$$

and Bayesian information criterion (BIC)

$$BIC = 2 \cdot \ln(L) + p \cdot \ln(n)$$

where p corresponds to number of free parameters and L is maximum value of likelihood function. With higher penalization for number of parameters, BIC is less likely to propose an overfitted model.

Based on both criteria, JTT + Γ + F model was selected for both interleukin and receptor alignments. JTT is a protein evolution model based on substitutional matrix proposed by (Jones et al. 1992) and is used with Γ distribution of parameters with shape parameter α estimated within the analysis. F option assumes empirical amino acids frequencies.

3.4 Maximum Likelihood inference of phylogeny

Maximum likelihood is one of the basic methods of frequentist statistics (as opposed to Bayesian statistical framework) and it is widely used in bioinformatics. ML method aims to estimate the parameters of the model, when given the data. The method estimates the parameters, so that the likelihood of the data coming from the distribution defined by the values of parameters, reaches maximal possible values.

3.4.1 RAxML

For maximum likelihood phylogeny I used RAxML (Randomised Axelerated maximum likelihood) software (Stamatakis 2014). RAxML produces maximum parsimony trees, followed by calculation of likelihood of each tree by evaluating the tree parameters. Nodes are supported by bootstrap values from standard bootstrapping algorithm or rapid bootstrapping algorithm (Stamatakis et al. 2008), developed to lower computational demands of maximum likelihood phylogenetic inference.

ML estimations were calculated by RAxML MPI version 8.2.4. As the input, RAxML requires alignment in PHYLIP format. Protein evolution model is set according to ProtTest 3 results, gamma distribution shape parameter is calculated by RAxML:

```
raxmlHPC -f a -s alignment_file.phy -n il -m PROTGAMMAJTTX -p 12345 -x 12345
-#1000 > log_file
```

-s input alignment

-n output files

-m model used for estimation

PROT – protein model

GAMMA – Γ distribution of parameters

X – empirical frequencies

-p random seed for parsimony inference (important to reproduce results, required by RAxML)

-x random seed for bootstrapping

-# number of bootstrap replicates

-f a execute rapid bootstrapping in one step with ML search

Above stated command executes rapid bootstrap analysis (Stamatakis 2014) with 1000 replicates and afterwards performs throrough search for best ML scoring tree. The output of the run is a tree with bootstrap values of the nodes, obtained by one command. Nodes with bootstrap values < 50 are not to be considered reliable and are therefore collapsed using TreeGraph 2 (Stöver and Müller 2010).

3.5 Bayesian inference of phylogeny

Bayesian statistical methods are widely used not only in biological applications, but also in physics and other fields. Bayesian inference of phylogenies uses likelihood function to calculate posterior probability distribution of phylogenies (P(A|B)), implementing model of evolution. Following the Bayes' theorem:

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)}$$

where P(A) represents posterior probability of the tree and P(B) likelihood of the data the trees are calculated. Nodal support of phylogeny is set by posterior probability of the node in the phylogenetic tree.

3.5.1 MrBayes

MrBayes is a software using above described principles to infer phylogenies. MrBayes uses Markov chain Monte Carlo methods (MCMC) to sample from the posterior distribution. In default, MrBayes uses Metropolis coupling MCMC to run number of heated chains and a cold chain over adjusted distribution with less peaks and bigger steps, allowing crossing the "valleys" in probability, therefore it is less likely for the analysis to end in a local maximum, rather than finding global maximum probability.

MrBayes is used in version 3.2.6 (Ronquist et al. 2012; Ronquist and Huelsenbeck 2003). MrBayes works either in interactive mode, setting parameters one by one, or by inputting NEXUS format alignment in file with MrBayes command block for interleukins:

```
Begin MrBayes;
lset rates=gamma ngammacat=4;
prset statefreqpr=fixed(empirical) aamodelpr=fixed(jones) shapepr=fixed(3.074);

mcmcp ngen=5000000 nruns=2 nchains=4 printfreq=250 samplefreq=250 burnin=1000;
mcmc;
sump;
sumt;
end;
```

where lset and prset specify the model used for the analysis estimated by ProtTest 3, ngen parameter sets number of generations of MCMC estimations, nruns number of runs and nchains number of heated chains. Printfreq sets frequency of printed results, samplefreq frequency of trees, that will participate in final consensus tree. Burnin command sets number of initial samples (not probability estimations) that are discarded at the beginning of analysis for their lower posterior probability. In estimation of receptor phylogeny, shapepr parameter was set as calculated previously by ProtTest 3.

3.6 Comparison of phylogenetic approaches

I compared maximum likelihood and Bayesian estimation of phylogenies visually, based on knowledge of relations of proteins. Afterwards, I calculated distance of tree topologies using dist.topo tool of package *ape* (Paradis et al. 2004) in R to quantify differences of phylogenies. The dist.topo calculates two types of scores, "PH85", defining difference of internal nodes and "score", derived from previous but incorporating internal branch lengths into the analysis – score calculated as square root of the sum of the squared differences of the (internal) branch lengths, defining different tip topologies.

3.7 Ligand–receptor coevolution analysis

For further analysis only the best result tree was selected from the previous phylogenetic analyses. Analysis of coevolution, or correlated evolution, where topology of the tree and branch lengths should be correlated in related proteins, was preformed in R package *ape* (Paradis et al. 2004).

3.7.1 Parafit

Analysis of evolutionary relations is performed with Parafit tool (Legendre et al. 2002) of the ape package. Parafit was originally developed for testing of host–parasite coevolutionary relations, however is applicable to any related coevolving genes, proteins or organisms.

Parafit is able to calculate a global test of coevolution, indicating signs of relations within the 2 given trees or to test individual links between host and parasite, or in this case, ligand and receptor. For testing of individual links Parafit requires an input of the two distance matrices created from unrooted phylogenies of interacting proteins, which I calculated in R. Additionally, for individual protein link coevolution testing, a matrix of relations between proteins based on theoretical knowledge of interleukin–receptor interaction is needed. The input matrices are multiplied to create one matrix for the analysis.

Test statistics is calculated as a difference of sums of squares of values in the main diagonal of the combined matrix and a matrix without the particular interaction to determine the importance of the ligand–receptor relation (F1.stat) and in the second case, as difference standardised by the trace of non–permutated matrix (F2.stat). Permutation tests are based on random shuffling of values in rows of a relation matrix.

Results of Parafit search for coevolution were visualised by cophylo tool of *phytools* package in R (Revell 2012) in comparison with links of ligand–receptor pairs with no significant signs of coevolution.

3.8 ConSurf

ConSurf is a software for estimation of evolutionary conserved regions and sites in proteins based on secondary or tertiary structure (Glaser et al. 2003; Landau et al. 2005). Current version is available at http://consurf.tau.ac.il (Ashkenazy et al. 2016). Based on knowledge of conserved sites under slower evolution, biologically important and active regions may be predicted. Conserved regions are common in the protein core to maintain the structure, higher conservation is also expected in binding and active sites.

ConSurf calculates relative conservation of amino acid sites based on either sequence and BLAST of databases, or MSA. Optionally, known protein structure may be input to ConSurf. When used with structure file, ConSurf maps the conservation scheme to the structure.

I used ConSurf to find conservation of amino acid sites in IL–10 and IFNγ. In both cases, I used an input of the MSA of interleukins created as described above, Bayesian phylogeny of interleukins and a structure file retrieved from RSCB Protein Data Bank (PDB) at http://www.rcsb.org/pdb, describing structure of either IL–10 (PDB ID: 1Y6K (Yoon et al. 2005)) or IFNγ (PDB ID: 1FG9 (Thiel et al. 2000)) and queried by the protein sequence in *Homo sapiens* to establish the conservation.

4 Results

4.1 Phylogenetic inference

4.1.1 Maximum likelihood

ML search for optimal phylogeny (Figure 5) was performed using RAxML software. Bootstrap values for nodal support were obtained by rapid bootstrapping algorithm implemented in RAxML. Support values under 70 are marked by line weight, bootstrap values under 50 are generally not considered reliable and therefore were collapsed in the result tree. Figure 5 shows unrooted trees with grouping of related proteins with several polytomies in the tree created by RAxML.

Resulting trees from ML search for optimal tree have highly unresolved relations between the proteins and therefore only show grouping of particular interleukins and receptors and relations among the proteins is in many cases unclear. In the interleukin tree, grouping of IFN group is shown with support between 50 and 70. In the tree of receptors, group of IL22RA1, IL20RA and IL22BP is formed, IL10RB and IL20RB form a group, the rest of relations remains unresolved.

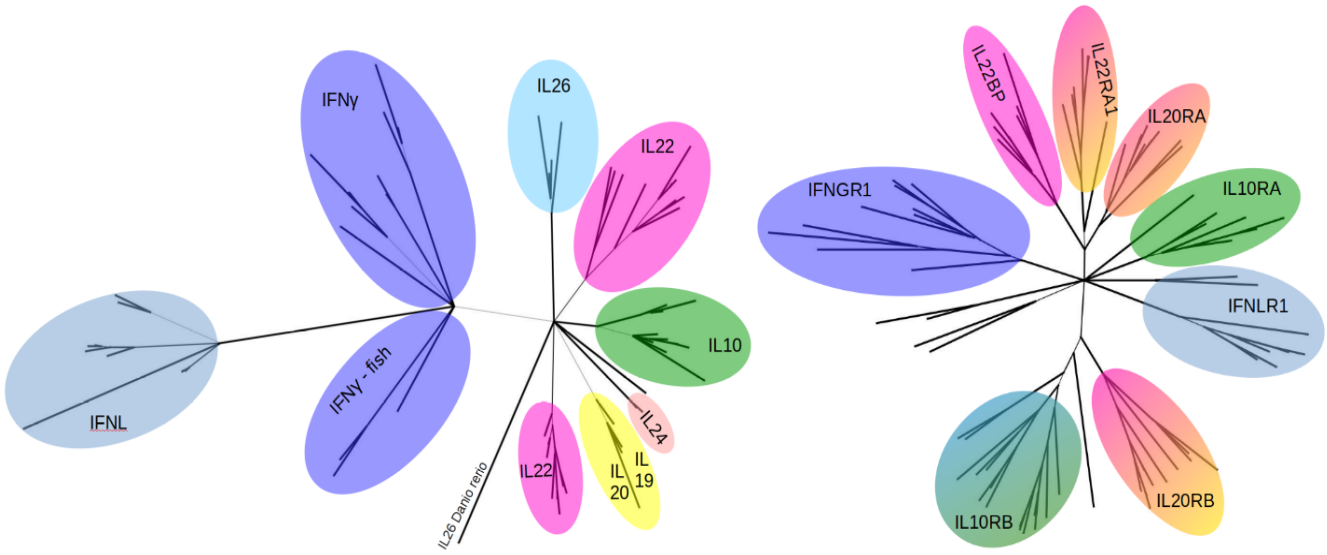

Figure 5. Unrooted maximum likelihood tree of FIL-10 interleukins (left) and their receptors. Line weight is scaled to bootstrap values. Nodes with bootstrap support < 50 were collapsed. Colours and colour gradient of receptors shows relation to interleukins. Results were visualised by FigTree and edited.

4.1.2 Bayesian

Bayesian inference of phylogeny is calculated in MrBayes software. The result tree is concatenated from tree sampling from the distribution defined by command block given to MrBayes. Nodal support is provided by node posterior probability values. Probability values of all nodes in both phylogenies were higher than 55, however, MrBayes itself creates polytomies in the topology.

Figure 6 shows resolved relations of interleukins in FIL-10 on the left and the receptor relations are shown on the right. The interleukin tree shows grouping of IFNγ into 2 groups, one containing only fish species. The grouping is supported by 93% posterior probability. The following node shows lower posterior probability of 60%. The receptor tree shows relation of IFNGR2 to FIL-10 receptors, IFNLR1 is divided into 2 groups, one containing mammal, bird and amphibian species, another with only fish species.

The result trees are visualised in detail with species specific tip labels as phylograms of FIL-10 (Figure 7) and receptors individual FIL-10 interleukins use for signalling (Figure 8).

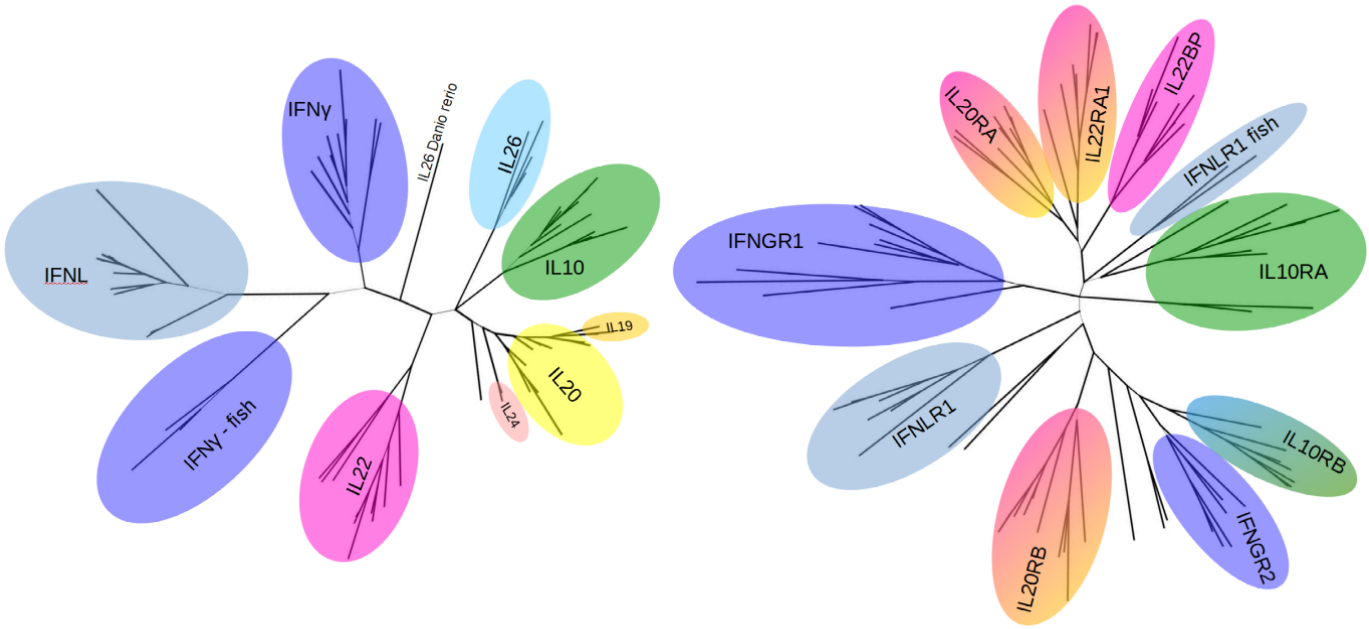

Figure 6. Unrooted Bayesian tree of FIL-10 interleukins (left) and their receptors. Line weight is scaled to posterior probability values. Colours and colour gradient of receptors show relation to interleukins. Results were visualised by FigTree and edited.

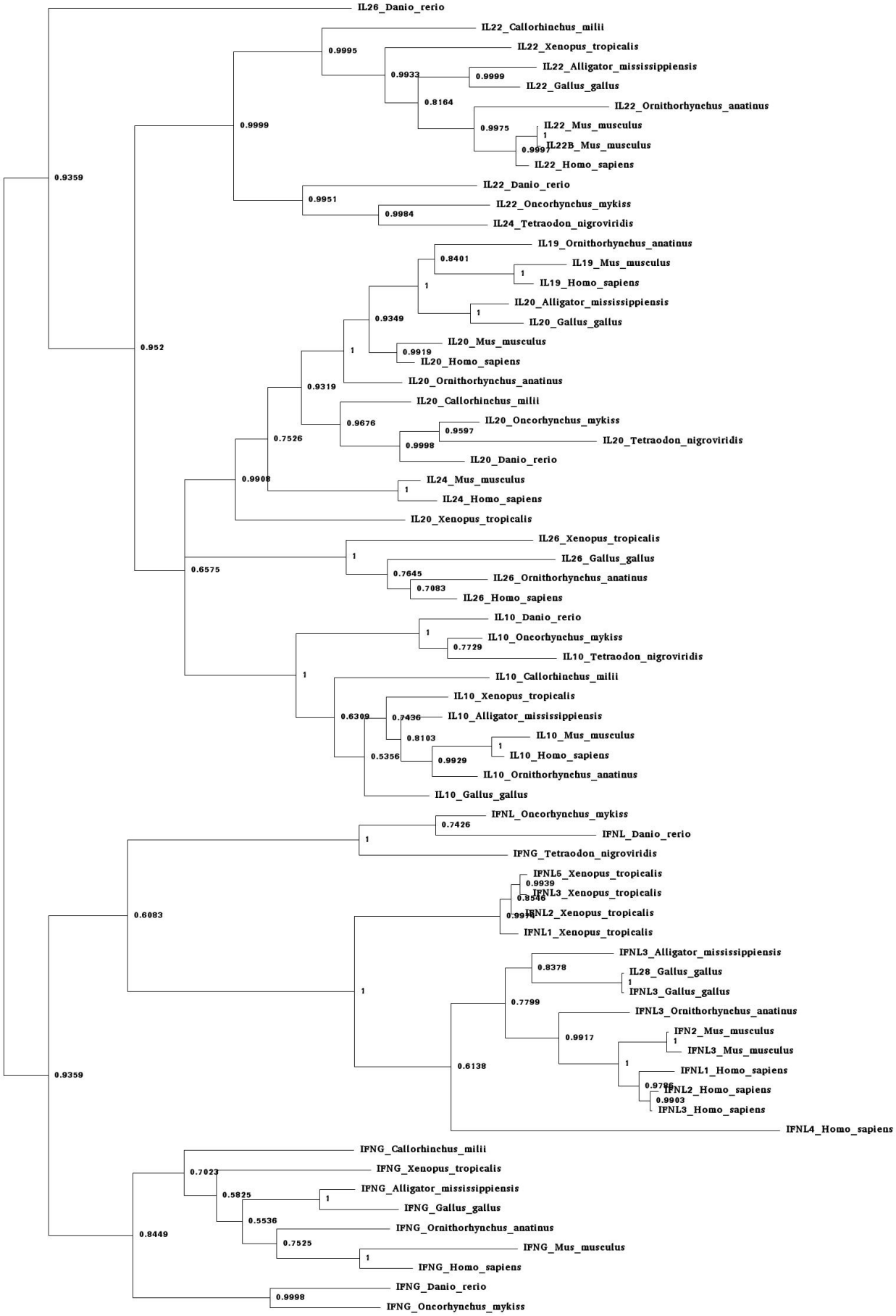

Figure 7. MrBayes phylogeny of interleukins of FIL-10. Node labels mark posterior probabilities. Result trees were visualised by FigTree.

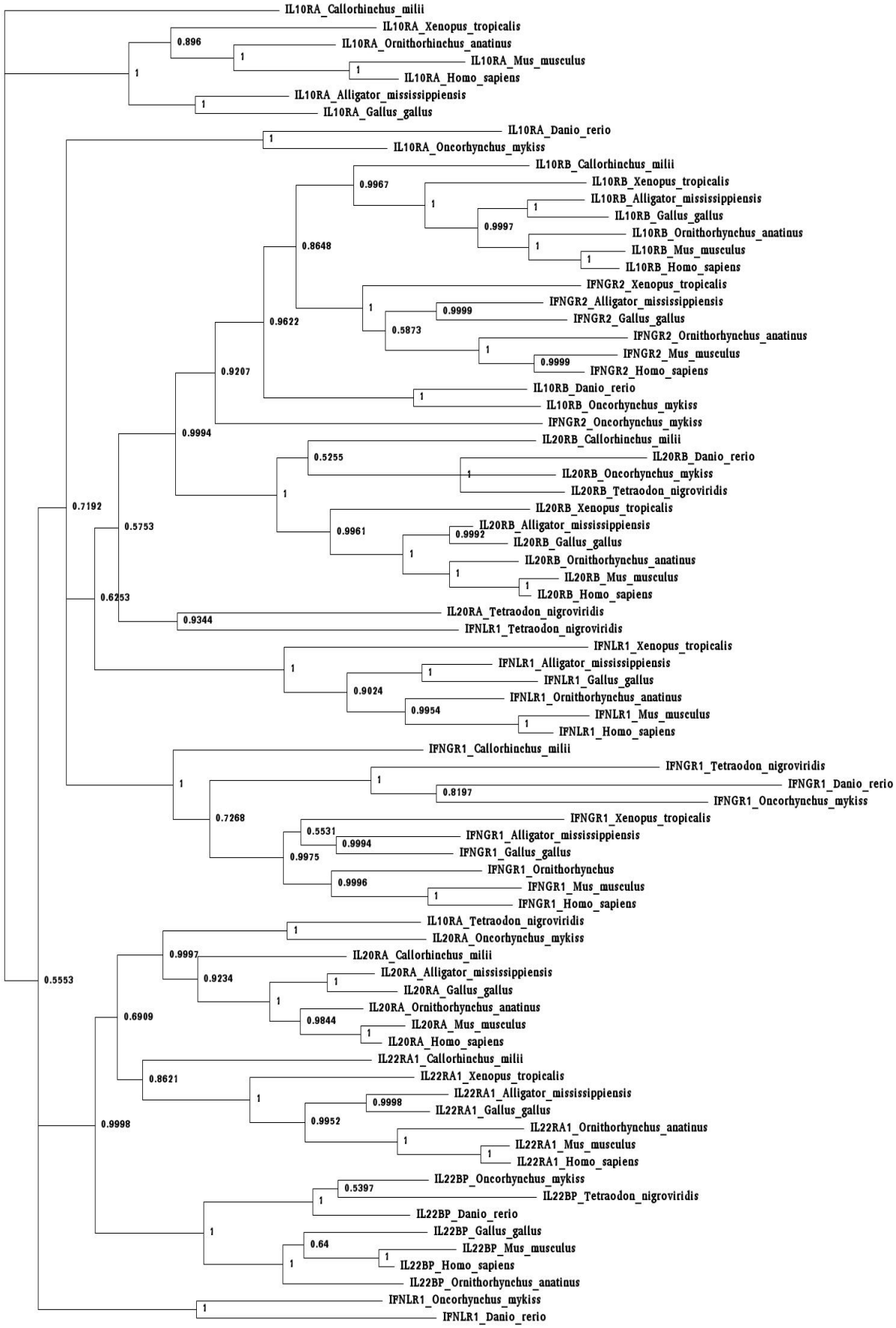

Figure 8. MrBayes phylogeny of receptors of FIL-10. Node labels represent posterior probabilities of the nodes. Result trees were visualised by FigTree.

4.1.3 Comparison of phylogenetic approaches

After inferring the phylogenies by two distinct methods, only one resulting phylogeny is to be used in further analysis. I tested similarity of the tree topologies created by RAxML and MrBayes with R *ape* package tool dist.topo.

In phylogenies of FIL-10, "PH85" method showing difference of number of internal branches was 17, and "score" method incorporating branch lengths into the calculation was 3.1996. FIL-10 receptors showed differences in internal branch numbers to be 9 and "score" value impaired by branch length to be 0.5480. Based on visual comparison and analysis of differences of topologies, Bayesian result trees were selected for further analysis.

4.2 Interleukin-receptor coevolution analysis

I calculated 2 types of coevolutionary statistics using Parafit, test of global coevolution within the two trees and tests of individual links between receptors and ligands. Global test of coevolution shown significant coevolution within the two trees with p-value = 0.001.

In the test of individual interactions two types of statistics were calculated. Significant interleukin receptor coevolution was found in most given relations (Figure 9) using both statistical tests (F1.stat and F2.stat) (Table 5) with p-value lower than 0.01. No significant signs of coevolution between the interleukin and receptor were found in IL-26 - IL10RB relation in every species tested, from fish to human proteins included in the analysis. IFN λ and both its receptors (IFNLR1 and IL10RA) were not significantly coevolved in *Mus musculus* and *Gallus gallus*. IL-19 and its receptors (IL20RA, IL20RB) showed no significant coevolution in none of the tested mammals (IL is present only in mammals). IL-20 shows no signs of significant coevolution with its receptors (IL20RA and IL20RB) in *Callorhinchus milii*, *Danio rerio* and *Gallus gallus*.

Table 5. **Parafit results of coevolution of FIL−10 and corresponding receptors.** Interactions with no detected significant coevolution was found are marked by grey. Types of used statistics are described in Chapter 3.7.1.

Interleukin	Receptor	F2..stat	p.F1	F2..stat	p.F2
IFNL2_Mus_musculus	IFNLR1_Mus_musculus	79	0.398	0.0023	0.108
IFNL2_Mus_musculus	IL10RA_Mus_musculus	78	0.337	0.0023	0.057
IFNG_Alligator_mississippiensis	IFNGR1_Alligator_mississippiensis	529	0.006	0.0153	0.001
IFNG_Alligator_mississippiensis	IFNGR2_Alligator_mississippiensis	1026	0.003	0.0296	0.001
IFNG_Callorhinchus_milii	IFNGR1_Callorhinchus_milii	637	0.003	0.0184	0.001
IFNG_Danio_rerio	IFNGR1_Danio_rerio	727	0.002	0.0210	0.001
IFNG_Gallus_gallus	IFNGR1_Gallus_gallus	698	0.002	0.0201	0.001
IFNG_Gallus_gallus	IFNGR2_Gallus_gallus	1112	0.002	0.0321	0.001
IFNG_Homo_sapiens	IFNGR1_Homo_sapiens	899	0.001	0.0259	0.001
IFNG_Homo_sapiens	IFNGR2_Homo_sapiens	1155	0.001	0.0333	0.001
IFNG_Mus_musculus	IFNGR1_Mus_musculus	789	0.003	0.0228	0.001
IFNG_Mus_musculus	IFNGR1_Mus_musculus	956	0.001	0.0276	0.001
IFNG_Oncorhynchus_mykiss	IFNGR1_Oncorhynchus_mykiss	526	0.001	0.0152	0.001
IFNG_Oncorhynchus_mykiss	IFNGR2_Oncorhynchus_mykiss	953	0.001	0.0275	0.001
IFNG_Ornithorhynchus_anatinus	IFNGR1_Ornithorhynchus_mykiss	462	0.001	0.0133	0.001
IFNG_Ornithorhynchus_anatinus	IFNGR2_Ornithorhynchus_anatinus	953	0.001	0.0275	0.001
IFNG_Xenopus_tropicalis	IFNGR1_Xenopus_tropicalis	517	0.003	0.0149	0.001
IFNG_Xenopus_tropicalis	IFNGR2_Xenopus_tropicalis	702	0.002	0.0202	0.001
IFNL_Danio_rerio	IFNLR1_Danio_rerio	594	0.004	0.0171	0.001
IFNL_Danio_rerio	IL10RB_Danio_rerio	564	0.006	0.0163	0.001
IFNL_Oncorhynchus_mykiss	IFNLR1_Oncorhynchus_mykiss	593	0.008	0.0171	0.001
IFNL_Oncorhynchus_mykiss	IL10RB_Oncorhynchus_mykiss	480	0.014	0.0138	0.001
IFNL_Tetraodon_nigroviridis	IFNLR1_Tetraodon_nigroviridis	1268	0.001	0.0366	0.001
IFNL1_Homo_sapiens	IFNLR1_Homo_sapiens	1178	0.001	0.0340	0.001
IFNL1_Homo_sapiens	IL10RB_Homo_sapiens	938	0.002	0.0271	0.001
IFNL1_Xenopus_tropicalis	IFNLR1_Xenopus_tropicalis	1240	0.001	0.0358	0.001
IFNL1_Xenopus_tropicalis	IL10RB_Xenopus_tropicalis	1164	0.003	0.0336	0.001
IFNL2_Homo_sapiens	IFNLR1_Homo_sapiens	1075	0.001	0.0310	0.001
IFNL2_Homo_sapiens	IL10RB_Homo_sapiens	859	0.003	0.0248	0.001
IFNL2_Xenopus_tropicalis	IFNLR1_Xenopus_tropicalis	1215	0.001	0.0350	0.001
IFNL2_Xenopus_tropicalis	IL10RB_Xenopus_tropicalis	1146	0.003	0.0331	0.001
IFNL3_Alligator_mississippiensis	IFNLR1_Alligator_mississippiensis	421	0.027	0.0121	0.001
IFNL3_Alligator_mississippiensis	IL10RB_Alligator_mississippiensis	929	0.001	0.0268	0.001
IFNL3_Gallus_gallus	IFNLR1_Gallus_gallus	907	0.001	0.0262	0.001
IFNL3_Gallus_gallus	IL10RB_Gallus_gallus	599	0.004	0.0173	0.001
IFNL3_Homo_sapiens	IFNLR1_Homo_sapiens	862	0.001	0.0249	0.001
IFNL3_Homo_sapiens	IL10RB_Homo_sapiens	694	0.005	0.0200	0.001
IFNL3_Mus_musculus	IFNLR1_Mus_musculus	906	0.001	0.0261	0.001
IFNL3_Mus_musculus	IL10RB_Mus_musculus	509	0.004	0.0147	0.001
IFNL3_Ornithorhynchus_anatinus	IFNLR1_Ornithorhynchus_anatinus	1032	0.001	0.0298	0.001
IFNL3_Ornithorhynchus_anatinus	IL10RB_Ornithorhynchus_anatinus	837	0.003	0.0242	0.001
IFNL3_Xenopus_tropicalis	IFNLR1_Xenopus_tropicalis	1282	0.001	0.0370	0.001
IFNL3_Xenopus_tropicalis	IL10RB_Xenopus_tropicalis	1252	0.003	0.0361	0.001
IFNL4_Homo_sapiens	IFNLR1_Homo_sapiens	947	0.001	0.0273	0.001
IFNL4_Homo_sapiens	IL10RB_Homo_sapiens	743	0.005	0.0214	0.001

IFNL5_Xenopus_tropicalis	IFNLR1_Xenopus_tropicalis	895	0.001	0.0258	0.001
IFNL5_Xenopus_tropicalis	IL10RB_Xenopus_tropicalis	876	0.006	0.0253	0.001
IL10_Alligator_mississippiensis	IL10RA_Alligator_mississippiensis	753	0.001	0.0217	0.001
IL10_Alligator_mississippiensis	IL10RB_Alligator_mississippiensis	868	0.003	0.0250	0.001
IL10_Callorhinchus_milii	IL10RA_Callorhinchus_milii	742	0.002	0.0214	0.001
IL10_Callorhinchus_milii	IL10RB_Callorhinchus_milii	700	0.005	0.0202	0.001
IL10_Danio_rerio	IL10RA_Danio_rerio	1050	0.001	0.0303	0.001
IL10_Danio_rerio	IL10RB_Danio_rerio	826	0.002	0.0238	0.001
IL10_Gallus_gallus	IL10RA_Gallus_gallus	1074	0.001	0.0310	0.001
IL10_Gallus_gallus	IL10RB_Gallus_gallus	602	0.002	0.0174	0.001
IL10_Homo_sapiens	IL10RA_Homo_sapiens	1110	0.001	0.0320	0.001
IL10_Homo_sapiens	IL10RB_Homo_sapiens	615	0.008	0.0177	0.001
IL10_Mus_musculus	IL10RA_Mus_musculus	1042	0.001	0.0301	0.001
IL10_Mus_musculus	IL10RB_Mus_musculus	465	0.004	0.0134	0.001
IL10_Oncorhynchus_mykiss	IL10RA_Oncorhynchus_mykiss	844	0.004	0.0243	0.001
IL10_Oncorhynchus_mykiss	IL10RB_Oncorhynchus_mykiss	681	0.004	0.0197	0.001
IL10_Ornithorhynchus_anatinus	IL10RA_Ornithorhynchus_anatinus	730	0.005	0.0211	0.001
IL10_Ornithorhynchus_anatinus	IL10RB_Ornithorhynchus_anatinus	663	0.003	0.0191	0.001
IL10_Tetraodon_nigroviridis	IL10RA_Tetraodon_nigroviridis	799	0.002	0.0230	0.001
IL10_Xenopus_tropicalis	IL10RA_Xenopus_tropicalis	677	0.005	0.0195	0.001
IL10_Xenopus_tropicalis	IL10RB_Xenopus_tropicalis	798	0.008	0.0230	0.001
IL19_Homo_sapiens	IL20RA_Homo_sapiens	−714	0.998	−0.0206	1.000
IL19_Homo_sapiens	IL20RB_Homo_sapiens	−892	0.998	−0.0257	1.000
IL19_Mus_musculus	IL20RA_Mus_musculus	−756	0.993	−0.0218	1.000
IL19_Mus_musculus	IL20RB_Mus_musculus	−993	0.998	−0.0286	1.000
IL19_Ornithorhynchus_anatinus	IL20RA_Ornithorhynchus_anatinus	−856	1.000	−0.0247	1.000
IL19_Ornithorhynchus_anatinus	IL20RB_Ornithorhynchus_anatinus	−1046	1.000	−0.0302	1.000
IL20_Alligator_mississippiensis	IL20RA_Alligator_mississippiensis	697	0.008	0.0201	0.001
IL20_Alligator_mississippiensis	IL20RB_Alligator_mississippiensis	−850	0.999	−0.0245	1.000
IL20_Alligator_mississippiensis	IL22RA1_Alligator_mississippiensis	−793	1.000	−0.0229	1.000
IL20_Callorhinchus_milii	IL20RA_Callorhinchus_milii	506	0.010	0.0146	0.001
IL20_Callorhinchus_milii	IL20RB_Callorhinchus_milii	−826	0.999	−0.0238	1.000
IL20_Danio_rerio	IL22RA1_Danio_rerio	−801	0.999	−0.0231	1.000
IL20_Danio_rerio	IL20RB_Danio_rerio	−851	0.999	−0.0246	1.000
IL20_Gallus_gallus	IL20RA_Gallus_gallus	−683	1.000	−0.0197	1.000
IL20_Gallus_gallus	IL20RB_Gallus_gallus	1451	0.001	0.0418	0.001
IL20_Gallus_gallus	IL22RA1_Gallus_gallus	1149	0.001	0.0332	0.001
IL20_Homo_sapiens	IL20RA_Homo_sapiens	1391	0.001	0.0401	0.001
IL20_Homo_sapiens	IL20RB_Homo_sapiens	1631	0.001	0.0471	0.001
IL20_Homo_sapiens	IL22RA1_Homo_sapiens	1328	0.001	0.0383	0.001
IL20_Mus_musculus	IL20RA_Mus_musculus	1406	0.001	0.0406	0.001
IL20_Mus_musculus	IL20RB_Mus_musculus	1663	0.001	0.0480	0.001
IL20_Mus_musculus	IL22RA1_Mus_musculus	1196	0.001	0.0345	0.001
IL20_Oncorhynchus_mykiss	IL20RA_Oncorhynchus_mykiss	2195	0.001	0.0633	0.001
IL20_Oncorhynchus_mykiss	IL20RB_Oncorhynchus_mykiss	2529	0.001	0.0729	0.001
IL20_Ornithorhynchus_anatinus	IL20RA_Ornithorhynchus_anatinus	2250	0.001	0.0649	0.001
IL20_Ornithorhynchus_anatinus	IL20RB_Ornithorhynchus_anatinus	2422	0.001	0.0699	0.001

IL20_Ornithorhynchus_anatinus	IL22RA1_Ornithorhynchus_anatinus	415	0.001	0.0120	0.001
IL20_Tetraodon_nigroviridis	IL20RA_Tetraodon_nigroviridis	2288	0.001	0.0660	0.001
IL20_Tetraodon_nigroviridis	IL20RB_Tetraodon_nigroviridis	1784	0.001	0.0514	0.001
IL20_Xenopus_tropicalis	IL20RB_Xenopus_tropicalis	1806	0.001	0.0521	0.001
IL22_Alligator_mississippiensis	IL20RB_Alligator_mississippiensis	2590	0.001	0.0747	0.001
IL22_Alligator_mississippiensis	IL22RA1_Alligator_mississippiensis	2059	0.001	0.0594	0.001
IL22_Callorhinchus_milii	IL20RB_Callorhinchus_milii	2605	0.001	0.0751	0.001
IL22_Callorhinchus_milii	IL22RA1_Callorhinchus_milii	2175	0.001	0.0627	0.001
IL22_Danio_rerio	IL20RB_Danio_rerio	2552	0.001	0.0736	0.001
IL22_Danio_rerio	IL22BP_Danio_rerio	1739	0.001	0.0502	0.001
IL22_Gallus_gallus	IL20RB_Gallus_gallus	2573	0.001	0.0742	0.001
IL22_Gallus_gallus	IL22BP_Gallus_gallus	2097	0.001	0.0605	0.001
IL22_Gallus_gallus	IL22RA1_Gallus_gallus	2318	0.001	0.0669	0.001
IL22_Homo_sapiens	IL20RB_Homo_sapiens	2757	0.001	0.0795	0.001
IL22_Homo_sapiens	IL22BP_Homo_sapiens	2672	0.001	0.0771	0.001
IL22_Homo_sapiens	IL22RA1_Homo_sapiens	2600	0.001	0.0750	0.001
IL22_Mus_musculus	IL20RB_Mus_musculus	3163	0.001	0.0912	0.001
IL22_Mus_musculus	IL22BP_Mus_musculus	2434	0.001	0.0702	0.001
IL22_Mus_musculus	IL22RA1_Mus_musculus	2541	0.001	0.0733	0.001
IL22_Oncorhynchus_mykiss	IL20RB_Oncorhynchus_mykiss	3112	0.001	0.0898	0.001
IL22_Oncorhynchus_mykiss	IL22BP_Oncorhynchus_mykiss	2431	0.001	0.0701	0.001
IL22_Ornithorhynchus_anatinus	IL20RB_Ornithorhynchus_anatinus	2943	0.001	0.0849	0.001
IL22_Ornithorhynchus_anatinus	IL22BP_Ornithorhynchus_anatinus	2260	0.001	0.0652	0.001
IL22_Ornithorhynchus_anatinus	IL22RA1_Ornithorhynchus_anatinus	532	0.001	0.0154	0.001
IL22_Xenopus_tropicalis	IL22RA1_Xenopus_tropicalis	483	0.001	0.0139	0.001
IL22B_Mus_musculus	IL20RB_Mus_musculus	3241	0.001	0.0935	0.001
IL22B_Mus_musculus	IL22RA1_Mus_musculus	2519	0.001	0.0726	0.001
IL24_Homo_sapiens	IL20RA_Homo_sapiens	758	0.007	0.0219	0.001
IL24_Homo_sapiens	IL20RB_Homo_sapiens	920	0.006	0.0265	0.001
IL24_Homo_sapiens	IL22RA1_Homo_sapiens	560	0.019	0.0161	0.001
IL24_Mus_musculus	IL20RA_Mus_musculus	850	0.009	0.0245	0.001
IL24_Mus_musculus	IL20RB_Mus_musculus	1016	0.008	0.0293	0.001
IL24_Mus_musculus	IL22RA1_Mus_musculus	537	0.012	0.0155	0.001
IL24_Tetraodon_nigroviridis	IL20RA_Tetraodon_nigroviridis	874	0.013	0.0252	0.001
IL24_Tetraodon_nigroviridis	IL20RB_Tetraodon_nigroviridis	712	0.005	0.0205	0.001
IL26_Danio_rerio	IL10RB_Danio_rerio	−493	0.999	−0.0142	1.000
IL26_Danio_rerio	IL20RB_Danio_rerio	912	0.006	0.0263	0.001
IL26_Gallus_gallus	IL10RB_Gallus_gallus	−234	0.978	−0.0068	1.000
IL26_Gallus_gallus	IL20RB_Gallus_gallus	915	0.007	0.0264	0.001
IL26_Homo_sapiens	IL10RB_Homo_sapiens	−299	0.982	−0.0086	1.000
IL26_Homo_sapiens	IL20RB_Homo_sapiens	1029	0.004	0.0297	0.001
IL26_Ornithorhynchus_anatinus	IL10RB_Ornithorhynchus_anatinus	−417	0.986	−0.0120	1.000
IL26_Ornithorhynchus_anatinus	IL20RB_Ornithorhynchus_anatinus	976	0.004	0.0282	0.001
IL26_Xenopus_tropicalis	IL10RB_Xenopus_tropicalis	−436	0.984	−0.0126	1.000
IL26_Xenopus_tropicalis	IL20RB_Xenopus_tropicalis	631	0.013	0.0182	0.001
IL28_Gallus_gallus	IFNLR1_Gallus_gallus	−435	0.993	−0.0125	1.000
IL28_Gallus_gallus	IL10RB_Gallus_gallus	−205	0.967	−0.0059	1.000

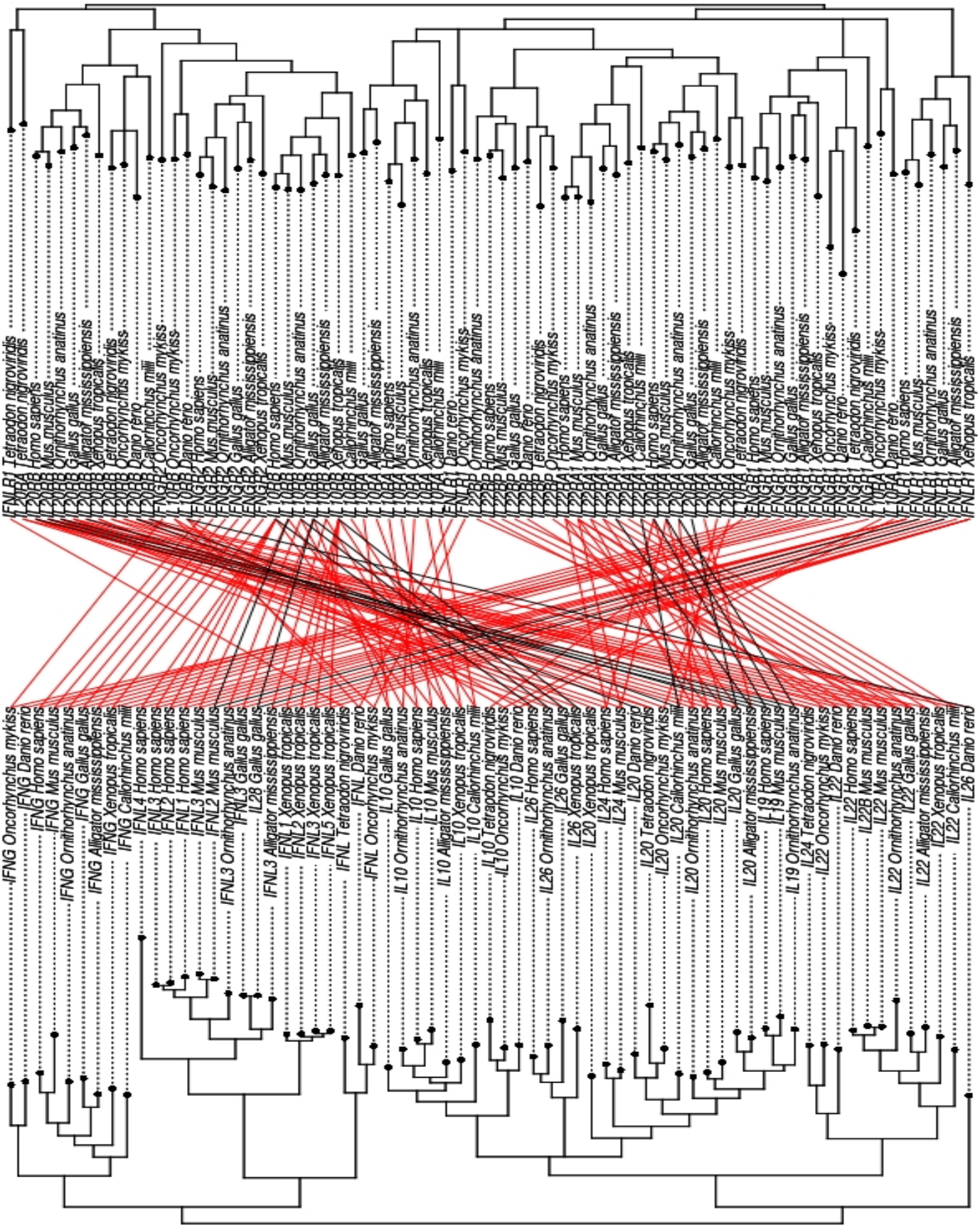

Figure 9. Graphical representation of coevolutionary relations and Parafit analysis. Red lines represent links between proteins with significant coevolution, black lines represent tested links with no significant signs of coevolution.

4.3 Protein conservation

Protein conservation was analysed based on amino-acid residue conservation in the dataset. The analysis was performed for interleukin 10 (in the structure in complex with IL10RA, PDB ID: 1Y6K (Yoon et al. 2005) and IFNγ in the multiplex with IFNGR1, PDB ID: 1FG9 (Thiel et al. 2000).

Figure 10 shows conservation of amino acid residues in IL-10, analysed by ConSurf (cyan represents the most variable regions, while magenta coloured regions are highest conserved) conservation of IFNγ amino acid residues and comparison of the two to the hydrophobicity conservation. Highly hydrophobic region marked by green is not overlapping with high conservation of amino acid marked by yellow, second highly conserved hydrophobicity region is partially overlapped. The regions of highest conservation of amino acid residues are identical for IL-10 and IFNγ.

Analysis of amino acid conservation is subsequently mapped to known structure of the protein. Conservation of amino acid residues in the structure of IL-10 is shown in Figure 11, with interaction with one IL10RA chain in monomer and homodimer. The conserved regions are visible in the core of the protein.

Amino acid residue conservation of IFNγ is mapped to the structure in Figure 12. The mapping shows interaction of one IFNγ molecule with IFNGR1 subunits and IFNγ dimer interaction with the receptor. The conserved regions are found in the centre of the protein interacting with second IFNγ subunit and facing interaction interface with IFNGR1 chain.

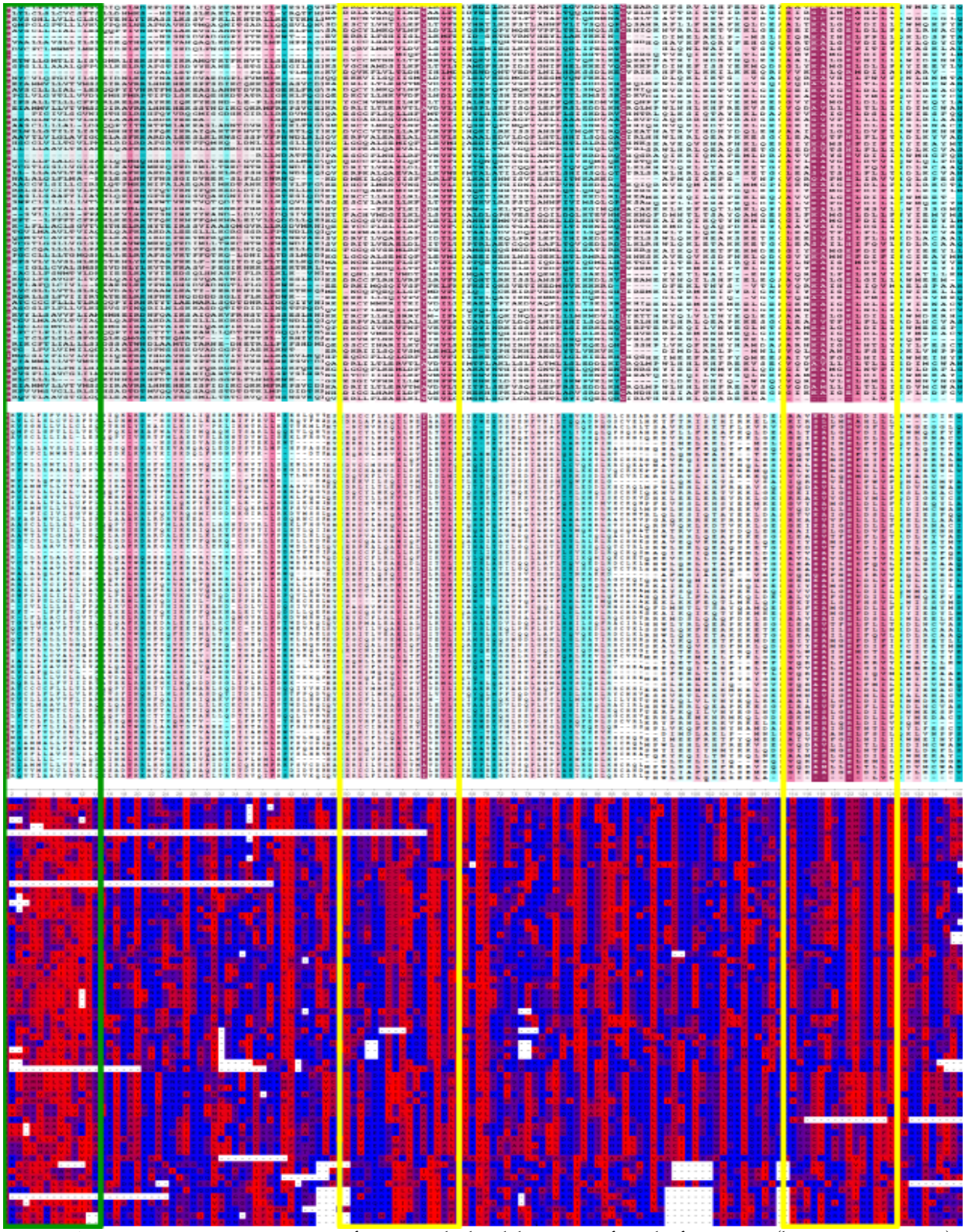

Figure 10. TOP: Conservation scores of IL−10 calculated by ConSurf, scale from cyan (lowest conservation) to magenta (highly conserved sites). MIDDLE: Conservation scores of IFN γ calculated by ConSurf, scale from cyan (lowest conservation) to magenta (highly conserved sites). BOTTOM: Visualisation of receptor protein MSA created by MAFFT and automatically trimmed using trimAl. Colour scheme marks hydrophobicity of amino acids (hydrophilic in blue, hydrophobic in red) Image obtained in UGENE software.

Green rectangle marks region of highly conserved hydrophobicity region, yellow high conservation of sites analysed by ConSurf. Similarity of the overall architecture of IL−10 and IFN γ is obvious.

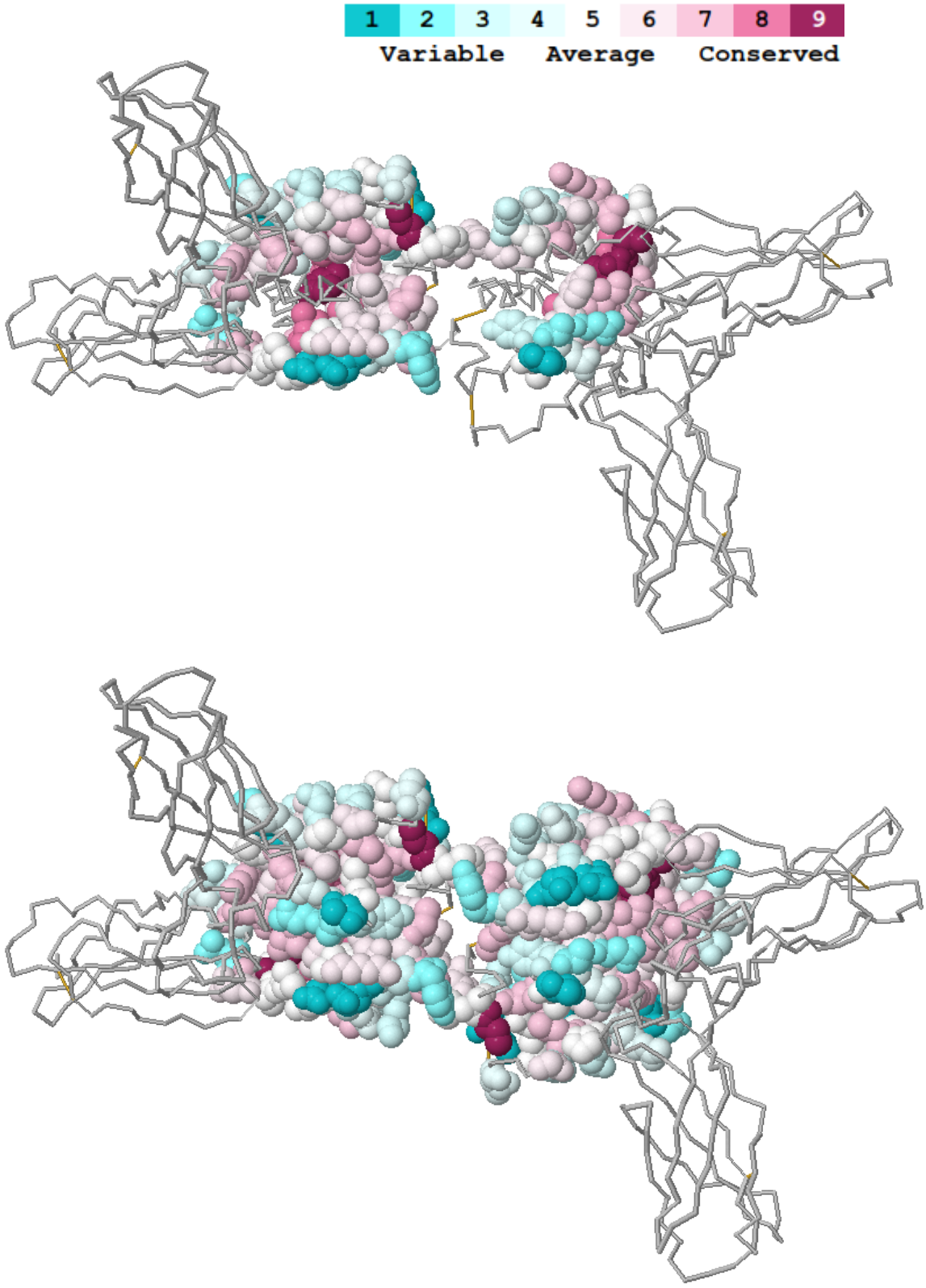

Figure 11. ConSurf scheme of amino acid residue conservation in IL-10 monomer (top) and homodimer (bottom) in complex with IL10RA (PDB:1Y6K) based on whole protein family. Both receptor chains are drawn as grey lines, IL-10 amino acids as spheres coloured by ConSurf.

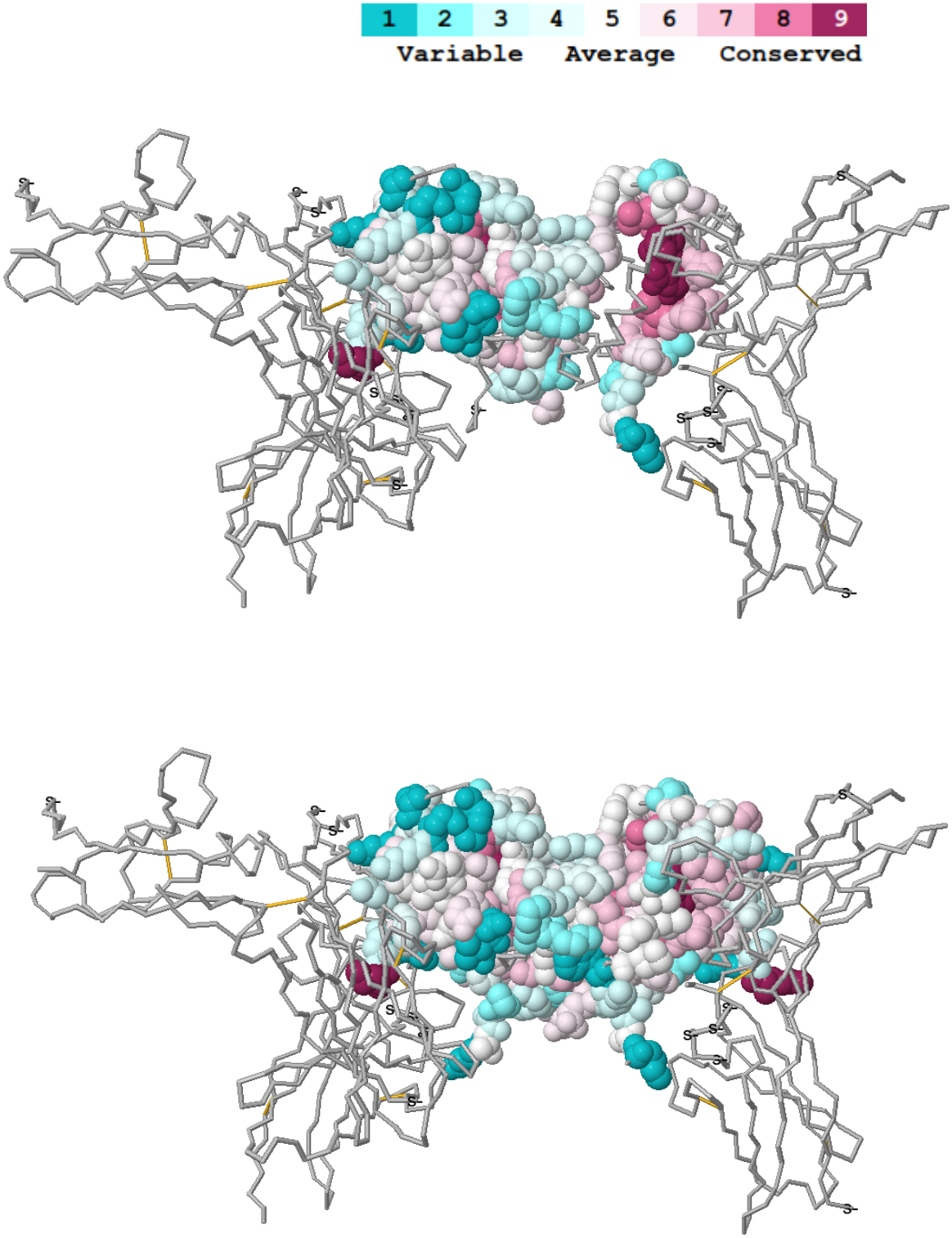

Figure 12. ConSurf scheme of amino acid residue conservation of IFN γ monomer (top) and homodimer (bottom) in complex with 3 IFNGR1 chains (PDB:1FG9), based on the whole protein family. -S marks incomplete side chains in the structure. All receptor chains are drawn as grey lines, IFN γ amino acids as spheres coloured by ConSurf.

5 Discussion

Interleukin 10 family represents an important family of immune signalling molecules directing immune response in divergent manner. Phylogenetic relations of the proteins with different, in some cases even opposite effect is not yet well described, similarly to FIL-10 evolutionary relation to IFNs with mainly antiviral activity. FIL-10 with included closely related IFNγ and IFNλs, as analysed in this study, uses combinations of shared receptor subunits with different or even opposite signalling outcome. Therefore, evolutionary relations are expected in the case of related proteins. Coevolution in sense of one to one substitution or convergent evolution of interleukins and receptors is detectable in evolving genome and subsequently in proteins and is analysable by statistical methods.

5.1 Phylogenies by ML and Bayesian approaches

Evolution of sequences may be analysed by several approaches. I analysed proteins by maximum-likelihood approach and Bayesian approach. From results of maximum likelihood analysis, relations of the proteins within the family were hardly definable, showing only grouping of orthologous proteins. Bayesian analysis of the same dataset showed relation of the orthologues in groups as well as relation between the proteins.

As the dataset consists of considerably high number of proteins of divergent species, length and variability of analysable regions is crucial for the successful definition of the relations among the family members. Bayesian analysis showed results with higher definition of the relations. The difference of the output of the methods may be caused by mechanism of inference of phylogeny, as well as different requirements for the data diversity. The maximum likelihood analysis creates several phylogenies, where the best scoring ML tree is found and afterwards used to map bootstrap values onto it. The best scoring ML tree may not describe the biological evolution exactly, since it is highly dependent on selection of a model describing amino acid evolution appropriate for the given data, and evaluation of parameters in ML search for the best scoring tree. In the Bayesian approach, several trees from predefined random sampling from the posterior distribution of the trees are used to create a consensus tree. Consensus tree, created with sampling of a number of trees is to be less likely to propose a phylogeny

that is not well describing the evolution of the protein family, while second or third to the best scoring ML phylogeny may propose better result of tree phylogeny with subsequently higher bootstrap values.

The difference of the two approaches is compared visually (Figure 5 and 6). Differences in the topology are clear, but the difference was also quantified as the difference of number of internal nodes and difference with a correction as described in Chapter 4.1.3. Differences in the topologies were higher in the trees of interleukins, than in receptor inference. Considering the limitation of ML approach in resolution of the differences discussed above, the difference in topologies may be explained by higher variability of amino acid residues in analysed receptor sequences, in comparison to the interleukin sequences.

For the further analysis, I selected Bayesian phylogenies based on higher resolution of inter protein relations. In the Bayesian phylogeny, we can see grouping of functionally similar proteins. IL-19 and IL-20 are shown to be very closely related, with IL-19 diverged from IL-20 very late in evolution. IL-19 is currently, according to the information in sequence databases, known to be expressed only in mammalian species, and the resulting phylogeny supports the divergence from mammalian IL-20. The branch of the tree with IL-19 and IL-20 also contains IL-24, closely related with its function in skin and epithelia.

Group of interferons – IFNγ and IFNλ – form a rather distinct group from the rest of analysed proteins. The corresponding receptors IFNLGR1 and IFNLR1 seem to be rather unrelated to the rest of the receptors as well, however the distinction is not clear. Interestingly, IFNγ evolution seems to be split into two parts, with fish IFNγ as evolutionary older proteins forming one group and amphibian, bird and mammal IFNγ forming another. No such separation is present in IFNGR1 or IFNGR2, however proteins annotated as IFNLR1 form two groups of proteins. In this context, structural research of fish IFNγ is of interest as it may elucidate some of the differences distinguishing them from the other IFNγ.

IL-26 of *Danio rerio* represents a single protein unrelated closely to the other protein groups in the analysis. The gene for IL-26 in *Danio rerio* was described by genomic analysis (Igawa et al. 2006) however the protein is quite distinct from the other members in the phylogeny. *Danio rerio* underwent a WGD recently in evolution, therefore 2 copies of IL-26 are present in the genome. Thus, one of the genes

may have changed under evolutionary pressure to acquire different function, with the second copy maintaining the function of IL-26. However, the distinction of *Danio rerio* IL-26 is seemingly rather caused by incorrect database annotation and not evolutionary events.

IL22BP is a soluble protein, that binds IL-22 and competitively inhibits activation of JaK/STAT pathway by blocking interaction of IL-22 with IL22RA1 and IL10RB (Kotenko et al. 2001). In the phylogeny of receptors, IL22BP is diverged from subgroup of IL20RA and IL22RA1, which supports hypothesis, that IL22BP, an IL-22 antagonist, has evolved from IL22RA1, which is used by IL-22 for signalling, by loss or mutation in the membrane domain.

The resulting phylogenies describe relations within the groups of interleukins in one phylogeny and receptors in the second, however due to unique combinations of the shared receptor subunits, coevolution of interleukins and their receptors is not obvious from the phylogenies at the first sight. Therefore, coevolution between the ligands and receptors needs to be analysed from the phylogenies using numerical methods.

5.2 Coevolution of interleukins and their receptors

We expect coevolution of related proteins not only in selection of parasites and hosts, where coevolution is often mediated by positive selection pressures (Anderson and May 1982), but also in inter-protein interactions of various types. In the coevolution of ligand-receptor pairs, purifying selection is expected, together with interrelated changes in amino acid structure in active sites (Fraser et al. 2002), since preservation of active amino acid sites is necessary to maintain the signalling function. Changes in evolutionary relations between functionally connected proteins are detectable at the genetic level and subsequently the encoded protein structure. Coevolution was analysed using method originally developed for comparison of evolution in parasites and hosts, since the analytic method implements presumptions of the evolution applicable likewise to ligand-receptor coevolution (Legendre et al. 2002).

Previous research on IL-10 family evolution and evolution of receptors of the family showed possible relation among the interleukins of the family, however,

methods of analysis are not well specified (Kotenko and Langer 2004) or structure of the used data provides results with limited number of branches with acceptable bootstrap, values to identify the relations and the common ancestor (Lutfalla et al. 2003) and relation of receptors and proteins is assumed, but not tested. Insight into coevolution of interleukins from FIL-10, emergence of new interleukins in evolution and correlation of evolution of interleukin-receptor pairs is provided by Krause and Pestka (2005) who assume that both IL-10 and IFNγ and some other members of the family as well as their receptors diverged before formation of vertebrates. Thus, correlation of evolution of the signalling pairs should be strong in ancient species, and with newly emerged interleukins be decreasing.

In this study, correlated evolution of interleukins of FIL-10 is detected in most tested interactions. However, several inter-protein interactions were statistically insignificant in more than one species, or even in every tested case. The most versatile receptor in the family - IL10RB is shared by IL-10, IL-22, IL-26 and IFNλ of all subtypes. However, in case of interaction with IL-26 and IFNλ, no significant coevolution with IL10RB was detected. IL-26, where all tested relations with IL10RB were insignificant, and IFNλ, which also shares IL10RB for signalling, are according to the calculated phylogeny and in addition to consideration of functional aspects, more distant members of the protein family. Both proteins are crucial for antiviral and antimicrobial response of the organism, unlike the other members of the family involved mainly in skin and epidermis immune reactions. IFNλ uses IFNLR1 receptor subunit for signalling, unshared with the other members of the FIL-10. In the interaction with IFNLR1, significant coevolution is found, thus we may assume that the evolution of IFNλ is directed by reciprocal changes between IFNλ and IFNLR1. IL-26 signals through a shared subunit of IL10RB and subunit of IL20RB used by less interleukins for signalling, therefore evolution of IL-26 is likely more affected by IL20RB than IL10RB.

Correlation of IL-19 and both of its receptors evolution is insignificant in all tested cases, coevolution of IL-20 and its receptors is in some cases insignificant as well. IL-19 diverges from IL-20 in mammals, therefore the proteins are highly similar in their function as well as usage of the receptor subunits. Since the two use the same receptor subunits, despite the divergence, certain level of similarity needs to be maintained by ongoing processes of negative selection in the interleukins and receptor, rather than directed coevolution. The further divergence of IL-19 and IL-20 without influence

of coevolution with IL20RA and IL20RB is unlikely, since it would likely lead to loss of signalling function in the two mentioned interleukins.

Evolution of proteins with more interactions with other proteins are generally slower than interaction of more independent proteins (Fraser et al. 2002). Therefore, interleukins using IL10RB with more interactions with other proteins are generally expected to be more conserved than proteins with more independent signalling subunits. In the tree of interleukins, it is clear that subfamily of IL−19, IL−20 and IL−24 is more diverged from the other proteins in the family, which may be a result of divergence of independent receptors from the receptor group. IL10RB, the most promiscuous receptor chain is further diverged from the other members of the family, therefore its divergence may force changes in the related interleukins. Nevertheless, the question of evolutionary divergence time is hard to be answered from unrooted trees, definition of ancestral sequence of the proteins is necessary to define the evolutionary age.

5.3 Conserved amino acid sites and domains

Proteins encoded in one genome evolve under different rates, dependent on the functional network connecting it to other proteins. The same differentiation of evolutionary speed occurs in amino acids in single gene and further in genome, regions in related proteins evolve dependently on each other (Fraser et al. 2002). Coevolution of the protein families is however a complex process whose speed is dependent on regions that need to be conserved to maintain the folding, solubility, and in a broader sense function of coevolving proteins.

Conserved amino acid residues and whole regions are most expected in cores of proteins to maintain protein folding and thus binding interfaces and even more so active sites in the protein and therefore are quite easily detected by analysing number of amino acid substitutions (Pils et al. 2005). Regions of amino acid residue variability are commonly on the surface of the protein. In the analysis of conservation, both monomeric and dimeric forms of selected proteins are shown. Monomeric form is shown for illustration of position of conserved sites otherwise not visible in the core of the protein. Highly conserved regions are often detected at sites where the protein binds to its partners. In addition, highly conserved regions are expected in sites of homodimer interactions (Valdar and Thornton 2001), where conservation regions are visible in our

analysis in both analysed proteins. Dimeric form conservation reflects variability of surface proteins in both IL-10 and IFNγ.

Regions with the most conserved amino acid do not completely coincide with regions of hydrophilic or hydrophobic character. In this case, folding is preserved by substitutions of amino acids with similar hydrophobicity and steric characteristics (Ladunga and Smith 1997). Regions with highly preserved hydrophobicity are likely regions of membrane domains, where particular amino acid conservation is unnecessary. In analogy, hydrophilic regions with unpreserved amino acids may indicate surface amino acids not involved in critical (functional) interactions. Therefore, on protein surfaces or in intermembrane regions, the amino acid character, not the particular amino acid in the sequence is conserved.

Conservation of amino acid sites and its effect on structure and hence function requires further research, especially in proteins and species with interesting evolutionary history such as IFNγ in fish, and later divergence of IFNγ in other organisms, paying attention to the relation to orthologues in species with well described function and structure.

6 Conclusions

This study investigates evolutionary history of important group of immune regulators, interleukin 10 family proteins, and its relation to evolutionary history of their receptors and describes structural conservation throughout the interleukin 10 family proteins using statistical methods and data available at public databases.

The evolutionary relations of the FIL-10 and its related proteins from groups of interferons are not yet well understood, however the presented results help to understand the relation within the groups. The relations between sequences within the interleukin family show distinguishing of IL-19, IL-20 and IL-24 subfamily, interferons included in the analysis and closeness of IL-26 to antiviral interferons. The receptor phylogeny shows relation of IL22BP to IL22RA1 and subgrouping of the related interleukin receptors.

Evolution of interleukin-receptor pairs is in most cases correlated, with a few important exceptions: interaction of IL10RA, the most used receptor chain throughout the protein family with some, but by far not most, of its ligands, most notably IL-26 and interaction of the most diverged interleukins in the group - IL-19 and IL-20 and their receptors.

The evolution of amino acid sequences and their conservation or variability depends in a complex way on the protein structure and function. The most conserved regions are observed in cores of proteins and intra-monomer interfaces, as well as in functionally critical regions such as binding sites. Interestingly, high conservation of amino acid residues observed here is not directly overlapped to high conservation of hydrophobic regions expected in the protein interiors as we also showed here on example of three dimensional structures of IL-10 and IFN γ.

Relations between interleukins in the group may also provide a guide for search for interleukins of FIL-10 in taxonomic groups, especially in fish taxa, where no similar signalling molecules are currently known. Thorough analysis of available genomic data may provide basis for further experimental research of ancestors of the family.

7 References

ABASCAL, Federico, Rafael ZARDOYA and David POSADA, 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics (Oxford, England)* [online]. 1. 5., **21**(9), 2104–2105. ISSN 1367-4803. Available at: doi:10.1093/bioinformatics/bti263

AKDIS, Mübeccel, Simone BURGLER, Reto CRAMERI, Thomas EIWEGGER, Hiroyuki FUJITA, Enrique GOMEZ, Sven KLUNKER, Norbert MEYER, Liam O'MAHONY, Oscar PALOMARES, Claudio RHYNER, Nadia OUAKED, Nadia QUAKED, Anna SCHAFFARTZIK, Willem VAN DE VEEN, Sabine ZELLER, Maya ZIMMERMANN and Cezmi A. AKDIS, 2011. Interleukins, from 1 to 37, and interferon-γ: receptors, functions, and roles in diseases. *The Journal of Allergy and Clinical Immunology* [online]. 3., **127**(3), 701-721–70. ISSN 1097-6825. Available at: doi:10.1016/j.jaci.2010.11.050

ANDERSON, R. M. and R. M. MAY, 1982. Coevolution of hosts and parasites. *Parasitology* [online]. 10., **85**(02), 411. ISSN 0031-1820, 1469-8161. Available at: doi:10.1017/S0031182000055360

ANK, N., M. B. IVERSEN, C. BARTHOLDY, P. STAEHELI, R. HARTMANN, U. B. JENSEN, F. DAGNAES-HANSEN, A. R. THOMSEN, Z. CHEN, H. HAUGEN, K. KLUCHER and S. R. PALUDAN, 2008. An Important Role for Type III Interferon (IFN- /IL-28) in TLR-Induced Antiviral Activity. *The Journal of Immunology* [online]. 15. 2., **180**(4), 2474–2485. ISSN 0022-1767, 1550-6606. Available at: doi:10.4049/jimmunol.180.4.2474

ASHKENAZY, Haim, Shiran ABADI, Eric MARTZ, Ofer CHAY, Itay MAYROSE, Tal PUPKO and Nir BEN-TAL, 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* [online]. 8. 7., **44**(W1), W344–W350. ISSN 0305-1048, 1362-4962. Available at: doi:10.1093/nar/gkw408

BALDRIDGE, M. T., T. J. NICE, B. T. MCCUNE, C. C. YOKOYAMA, A. KAMBAL, M. WHEADON, M. S. DIAMOND, Y. IVANOVA, M. ARTYOMOV and H. W. VIRGIN, 2015. Commensal microbes and interferon- determine persistence of enteric murine norovirus infection. *Science* [online]. 16. 1., **347**(6219), 266–269. ISSN 0036-8075, 1095-9203. Available at: doi:10.1126/science.1258025

BERTHELOT, Camille, Frédéric BRUNET, Domitille CHALOPIN, Amélie JUANCHICH, Maria BERNARD, Benjamin NOËL, Pascal BENTO, Corinne DA SILVA, Karine LABADIE, Adriana ALBERTI, Jean-Marc AURY, Alexandra LOUIS, Patrice DEHAIS, Philippe BARDOU, Jérôme MONTFORT, Christophe KLOPP, Cédric CABAU, Christine GASPIN, Gary H. THORGAARD, Mekki BOUSSAHA, Edwige QUILLET, René GUYOMARD, Delphine GALIANA, Julien BOBE, Jean-Nicolas VOLFF, Carine GENÊT, Patrick WINCKER, Olivier JAILLON, Hugues Roest CROLLIUS and Yann GUIGUEN, 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications* [online]. 22. 4., **5**. ISSN 2041-1723. Available at: doi:10.1038/ncomms4657

BHUTIA, S. K., R. DASH, S. K. DAS, B. AZAB, Z. z. SU, S. G. LEE, S. GRANT, A. YACOUB, P. DENT, D. T. CURIEL, D. SARKAR and P. B. FISHER, 2010. Mechanism of Autophagy to Apoptosis Switch Triggered in Prostate Cancer Cells by Antitumor Cytokine Melanoma Differentiation-Associated Gene 7/Interleukin-24. *Cancer Research* [online]. 1. 5., **70**(9), 3667–3676. ISSN 0008-5472, 1538-7445. Available at: doi:10.1158/0008-5472.CAN-09-3647

BLUMBERG, Hal, Darrell CONKLIN, WenFeng XU, Angelika GROSSMANN, Ty BRENDER, Susan CAROLLO, Maribeth EAGAN, Don FOSTER, Betty A. HALDEMAN, Angie HAMMOND and OTHERS, 2001. Interleukin 20: discovery, receptor identification, and role in epidermal function. *Cell*. **104**(1), 9–19. Available at: doi:10.1016/S0092-8674(01)00187-8

BOONSTRA, A., R. RAJSBAUM, M. HOLMAN, R. MARQUES, C. ASSELIN-PATUREL, J. P. PEREIRA, E. E. M. BATES, S. AKIRA, P. VIEIRA, Y.-J. LIU, G. TRINCHIERI and A. O'GARRA, 2006. Macrophages and Myeloid Dendritic Cells, but Not Plasmacytoid Dendritic Cells, Produce IL-10 in Response to MyD88- and TRIF-Dependent TLR Signals, and TLR-Independent Signals. *The Journal of Immunology* [online]. 1. 12., **177**(11), 7551 - 7558. ISSN 0022-1767, 1550-6606. Available at: doi:10.4049/jimmunol.177.11.7551

BRAAT, H., P. ROTTIERS, D. W. HOMMES, N. HUYGHEBAERT, E. REMAUT, J. P. REMON, S. J. H. VAN DEVENTER, S. NEIRYNCK, M. P. PEPPELENBOSCH and L. STEIDLER, 2006. A phase I trial with Transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clinical Gastroenterology and Hepatology* [online]. 6., **4**(6), 754 - 759. ISSN 1542-3565. Available at: doi:10.1016/j.cgh.2006.03.028

CAPELLA-GUTIERREZ, S., J. M. SILLA-MARTINEZ and T. GABALDON, 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* [online]. 1. 8., **25**(15), 1972 - 1973. ISSN 1367-4803, 1460-2059. Available at: doi:10.1093/bioinformatics/btp348

CAUDELL, E. G., J. B. MUMM, N. POINDEXTER, S. EKMEKCIOGLU, A. M. MHASHILKAR, X. H. YANG, M. W. RETTER, P. HILL, S. CHADA and E. A. GRIMM, 2002. The Protein Product of the Tumor Suppressor Gene, Melanoma Differentiation-Associated Gene 7, Exhibits Immunostimulatory Activity and Is Designated IL-24. *The Journal of Immunology* [online]. 15. 6., **168**(12), 6041 - 6046. ISSN 0022-1767, 1550-6606. Available at: doi:10.4049/jimmunol.168.12.6041

CELLA, Marina, Anja FUCHS, William VERMI, Fabio FACCHETTI, Karel OTERO, Jochen K. M. LENNERZ, Jason M. DOHERTY, Jason C. MILLS and Marco COLONNA, 2009. A human natural killer cell subset provides an innate source of IL-22 for mucosal immunity. *Nature* [online]. 5. 2., **457**(7230), 722 - 725. ISSN 0028-0836, 1476-4687. Available at: doi:10.1038/nature07537

CHANG, C., E. MAGRACHEVA, S. KOZLOV, S. FONG, G. TOBIN, S. KOTENKO, A. WLODAWER and A. ZDANOV, 2003. Crystal Structure of Interleukin-19 Defines a New Subfamily of Helical Cytokines. *Journal of Biological Chemistry* [online]. 31. 1., **278**(5), 3308 - 3313. ISSN 0021-9258, 1083-351X. Available at: doi:10.1074/jbc.M208602200

CHAUDHRY, Ashutosh, Robert M. SAMSTEIN, Piper TREUTING, Yuqiong LIANG, Marina C. PILS, Jan-Michael HEINRICH, Robert S. JACK, F. Thomas WUNDERLICH, Jens C. BRÜNING, Werner MÜLLER and Alexander Y. RUDENSKY, 2011. Interleukin-10 Signaling in Regulatory T Cells Is Required for Suppression of Th17 Cell-Mediated Inflammation. *Immunity* [online]. 4., **34**(4), 566 - 578. ISSN 10747613. Available at: doi:10.1016/j.immuni.2011.03.018

CHIN, Yue E., Motoo KITAGAWA, Keisuke KUIDA, Richard A. FLAVELL and Xin-Yuan FU, 1997. Activation of the STAT signaling pathway can cause expression of caspase 1 and apoptosis. *Molecular and cellular biology*. **17**(9), 5328 - 5337.

COMMINS, Scott, John W. STEINKE and Larry BORISH, 2008. The extended IL-10 superfamily: IL-10, IL-19, IL-20, IL-22, IL-24, IL-26, IL-28, and IL-29. *Journal of Allergy and Clinical Immunology* [online]. 5., **121**(5), 1108 - 1111. ISSN 00916749. Available at: doi:10.1016/j.jaci.2008.02.026

DANDREA, A., M. ASTEAMEZAGA, Nm VALIANTE, Xj MA, M. KUBIN and G. TRINCHIERI, 1993. Interleukin-10 (il-10) Inhibits Human Lymphocyte Interferon Gamma-Production by Suppressing Natural-Killer-Cell Stimulatory Factor/Il-12 Synthesis in Accessory Cells. *Journal of Experimental Medicine* [online]. 1. 9., **178**(3), 1041 - 1048. ISSN 0022-1007. Available at: doi:10.1084/jem.178.3.1041

DARRIBA, D., G. L. TABOADA, R. DOALLO and D. POSADA, 2011. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* [online]. 15. 4., **27**(8), 1164–1165. ISSN 1367-4803, 1460-2059. Available at: doi:10.1093/bioinformatics/btr088

DE JUAN, David, Florencio PAZOS and Alfonso VALENCIA, 2013. Emerging methods in protein co-evolution. *Nature Reviews Genetics* [online]. 5. 3., **14**(4), 249–261. ISSN 1471-0056, 1471-0064. Available at: doi:10.1038/nrg3414

DEL PRETE, G., M. DE CARLI, F. ALMERIGOGNA, M. G. GIUDIZI, R. BIAGIOTTI and S. ROMAGNANI, 1993. Human IL-10 is produced by both type 1 helper (Th1) and type 2 helper (Th2) T cell clones and inhibits their antigen-specific proliferation and cytokine production. *Journal of Immunology (Baltimore, Md.: 1950)*. 15. 1., **150**(2), 353–360. ISSN 0022-1767.

DONNELLY, Raymond P., Faruk SHEIKH, Harold DICKENSHEETS, Ram SAVAN, Howard A. YOUNG and Mark R. WALTER, 2010. Interleukin-26: An IL-10-related cytokine produced by Th17 cells. *Cytokine & Growth Factor Reviews* [online]. 10., **21**(5), 393–401. ISSN 13596101. Available at: doi:10.1016/j.cytogfr.2010.09.001

DUHEN, Thomas, Rebekka GEIGER, David JARROSSAY, Antonio LANZAVECCHIA and Federica SALLUSTO, 2009. Production of interleukin 22 but not interleukin 17 by a subset of human skin-homing memory T cells. *Nature Immunology* [online]. 8., **10**(8), 857–863. ISSN 1529-2908, 1529-2916. Available at: doi:10.1038/ni.1767

EALICK, Steven E., William J. COOK and et AL, 1991. Three-Dimensional Structure of Recombinant Human Interferon-(Gamma). *Science; Washington*. 3. 5., **252**(5006), 698. ISSN 00368075.

FELSENSTEIN, Joseph, 2004. *Inferring Phylogenies*. B.m.: Sinauer. ISBN 978-0-87893-177-4.

FRASER, H. B., A. E. HIRSH, L. M. STEINMETZ, C. SCHARFE and M. W. FELDMAN, 2002. Evolutionary rate in the protein interaction network. *Science* [online]. 26. 4., **296**(5568), 750–752. ISSN 0036-8075. Available at: doi:10.1126/science.1068696

GAD, H. H., C. DELLGREN, O. J. HAMMING, S. VENDS, S. R. PALUDAN and R. HARTMANN, 2009. Interferon- Is Functionally an Interferon but Structurally Related to the Interleukin-10 Family. *Journal of Biological Chemistry* [online]. 31. 7., **284**(31), 20869–20875. ISSN 0021-9258, 1083-351X. Available at: doi:10.1074/jbc.M109.002923

GALLAGHER, G., H. DICKENSHEETS, J. ESKDALE, L. S. IZOTOVA, O. V. MIROCHNITCHENKO, J. D. PEAT, N. VAZQUEZ, S. PESTKA, R. P. DONNELLY and S. V. KOTENKO, 2000. Cloning, expression and initial characterisation of interleukin-19 (IL-19), a novel homologue of human interleukin-10 (IL-10). *Genes and immunity* [online]. **1**(7), 442. Available at: doi:10.1038/sj.gene.6363714

GILBERT, Don, 2002. Sequence File Format Conversion with Command-Line Readseq. In: *Current Protocols in Bioinformatics* [online]. B.m.: John Wiley & Sons, Inc. ISBN 978-0-471-25095-1. Available at: http://dx.doi.org/10.1002/0471250953.bia01es00

GLASER, Fabian, Tal PUPKO, Inbal PAZ, Rachel E. BELL, Dalit BECHOR-SHENTAL, Eric MARTZ and Nir BEN-TAL, 2003. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics*. **19**(1), 163–164.

GLOCKER, Erik-Oliver, Daniel KOTLARZ, Kaan BOZTUG, E. Michael GERTZ, Alejandro A. SCHÄFFER, Fatih NOYAN, Mario PERRO, Jana DIESTELHORST, Anna ALLROTH, Dhaarini MURUGAN and OTHERS, 2009. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *New England Journal of Medicine* [online]. **361**(21), 2033–2045. Available at: doi:10.1056/NEJMoa0907206

GOH, Chern-Sing, Andrew A. BOGAN, Marcin JOACHIMIAK, Dirk WALTHER and Fred E. COHEN, 2000. Co-evolution of proteins with their interaction partners. *Journal of Molecular Biology* [online]. 6., **299**(2), 283 – 293. ISSN 00222836. Available at: doi:10.1006/jmbi.2000.3732

HAMMING, Ole J., Ewa TERCZYŃSKA-DYLA, Gabrielle VIEYRES, Ronald DIJKMAN, Sanne E. JØRGENSEN, Hashaam AKHTAR, Piotr SIUPKA, Thomas PIETSCHMANN, Volker THIEL and Rune HARTMANN, 2013. Interferon lambda 4 signals via the IFNλ receptor to regulate antiviral activity against HCV and coronaviruses. *The EMBO Journal* [online]. 27. 11., **32**(23), 3055 – 3065. ISSN 0261-4189, 1460-2075. Available at: doi:10.1038/emboj.2013.232

HOLAN, Vladimir, Alena ZAJICOVA, Eliska JAVORKOVA, Peter TROSAN, Milada CHUDICKOVA, Michaela PAVLIKOVA and Magdalena KRULOVA, 2014. Distinct cytokines balance the development of regulatory T cells and interleukin-10-producing regulatory B cells. *Immunology* [online]. 4., **141**(4), 577 – 586. ISSN 00192805. Available at: doi:10.1111/imm.12219

HOR, S., H. PIRZER, L. DUMOUTIER, F. BAUER, S. WITTMANN, H. STICHT, J.-C. RENAULD, R. DE WAAL MALEFYT and H. FICKENSCHER, 2004. The T-cell Lymphokine Interleukin-26 Targets Epithelial Cells through the Interleukin-20 Receptor 1 and Interleukin-10 Receptor 2 Chains. *Journal of Biological Chemistry* [online]. 6. 8., **279**(32), 33343 – 33351. ISSN 0021-9258, 1083-351X. Available at: doi:10.1074/jbc.M405000200

HORIUCHI, Hiroshi, Bijay PARAJULI, Yue WANG, Yasu-Taka AZUMA, Tetsuya MIZUNO, Hideyuki TAKEUCHI and Akio SUZUMURA, 2015. Interleukin-19 Acts as a Negative Autocrine Regulator of Activated Microglia. *PLOS ONE* [online]. 20. 3., **10**(3), e0118640. ISSN 1932-6203. Available at: doi:10.1371/journal.pone.0118640

HUBER, Samuel, Nicola GAGLIANI, Lauren A. ZENEWICZ, Francis J. HUBER, Lidia BOSURGI, Bo HU, Matija HEDL, Wei ZHANG, William O'CONNOR, Andrew J. MURPHY, David M. VALENZUELA, George D. YANCOPOULOS, Carmen J. BOOTH, Judy H. CHO, Wenjun OUYANG, Clara ABRAHAM and Richard A. FLAVELL, 2012. IL-22BP is regulated by the inflammasome and modulates tumorigenesis in the intestine. *Nature* [online]. 17. 10., **491**(7423), 259 – 263. ISSN 0028-0836, 1476-4687. Available at: doi:10.1038/nature11535

IGAWA, Daisuke, Masahiro SAKAI and Ram SAVAN, 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. *Molecular Immunology* [online]. 3., **43**(7), 999 – 1009. ISSN 01615890. Available at: doi:10.1016/j.molimm.2005.05.009

JAILLON, O., J. M. AURY, F. BRUNET, J. L. PETIT, N. STANGE-THOMANN, E. MAUCELI, L. BOUNEAU, C. FISCHER, C. OZOUF-COSTAZ, A. BERNOT, S. NICAUD, D. JAFFE, S. FISHER, G. LUTFALLA, C. DOSSAT, B. SEGURENS, C. DASILVA, M. SALANOUBAT, M. LEVY, N. BOUDET, S. CASTELLANO, R. ANTHOUARD, C. JUBIN, V. CASTELLI, M. KATINKA, B. VACHERIE, C. BIEMONT, Z. SKALLI, L. CATTOLICO, J. POULAIN, V. DE BERARDINIS, C. CRUAUD, S. DUPRAT, P. BROTTIER, J. P. COUTANCEAU, J. GOUZY, G. PARRA, G. LARDIER, C. CHAPPLE, K. J. MCKERNAN, P. MCEWAN, S. BOSAK, M. KELLIS, J. N. VOLFF, R. GUIGO, M. C. ZODY, J. MESIROV, K. LINDBLAD-TOH, B. BIRREN, C. NUSBAUM, D. KAHN, M. ROBINSON-RECHAVI, V. LAUDET, V. SCHACHTER, F. QUETIER, W. SAURIN, C. SCARPELLI, P. WINCKER, E. S. LANDER, J. WEISSENBACH and H. R. CROLLIUS, 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* [online]. 21. 10., **431**(7011), 946 – 957. ISSN 0028-0836. Available at: doi:10.1038/nature03025

JENNINGS, Paul, Daniel CREAN, Lydia ASCHAUER, Alice LIMONCIEL, Konrad MOENKS, Georg KERN, Philip HEWITT, Karl LHOTTA, Arno LUKAS, Anja WILMES and Martin O. LEONARD, 2015. Interleukin-19 as a translational indicator of renal injury. *Archives of Toxicology* [online]. 1., **89**(1), 101 – 106. ISSN 0340-5761, 1432-0738. Available at: doi:10.1007/s00204-014-1237-3

JIANG, Hongping, Zao-Zhong SU, Jiao Jiao LIN, Neil I. GOLDSTEIN, C. S. YOUNG and Paul B. FISHER, 1996. The melanoma differentiation associated gene mda-7 suppresses cancer cell growth. *Proceedings of the National Academy of Sciences*. **93**(17), 9160－9165.

JONES, D. T., W. R. TAYLOR and J. M. THORNTON, 1992. The rapid generation of mutation data matrices from protein sequences. *Computer applications in the biosciences: CABIOS*. 6., **8**(3), 275－282. ISSN 0266-7061.

JONES, Lindsay L., Rajshekhar ALLI, Bofeng LI and Terrence L. GEIGER, 2016. Differential T Cell Cytokine Receptivity and Not Signal Quality Distinguishes IL-6 and IL-10 Signaling during Th17 Differentiation. *The Journal of Immunology* [online]. 1. 4., **196**(7), 2973－2985. ISSN 0022-1767, 1550-6606. Available at: doi:10.4049/jimmunol.1402953

KATOH, K. and D. M. STANDLEY, 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* [online]. 1. 4., **30**(4), 772－780. ISSN 0737-4038, 1537-1719. Available at: doi:10.1093/molbev/mst010

KATOH, Kazutaka, Kei-ichi KUMA, Hiroyuki TOH and Takashi MIYATA, 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research* [online]. **33**(2), 511－518. ISSN 0305-1048. Available at: doi:10.1093/nar/gki198

KATOH, Kazutaka, Kazuharu MISAWA, Kei-ichi KUMA and Takashi MIYATA, 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* [online]. 15. 7., **30**(14), 3059－3066. ISSN 1362-4962. Available at: doi:https://doi.org/10.1093/nar/gkf436

KIM, J. M., C. I. BRANNAN, N. G. COPELAND, N. A. JENKINS, T. A. KHAN and K. W. MOORE, 1992. Structure of the mouse IL-10 gene and chromosomal localization of the mouse and human genes. *Journal of Immunology (Baltimore, Md.: 1950)*. 1. 6., **148**(11), 3618－3623. ISSN 0022-1767.

KOLE, Abhisake and Kevin J. MALOY, 2014. Control of intestinal inflammation by interleukin-10. *Current Topics in Microbiology and Immunology* [online]. **380**, 19－38. ISSN 0070-217X. Available at: doi:10.1007/978-3-662-43492-5_2

KOTENKO, S. V., L. S. IZOTOVA, O. V. MIROCHNITCHENKO, E. ESTEROVA, H. DICKENSHEETS, R. P. DONNELLY and S. PESTKA, 2001. Identification, Cloning, and Characterization of a Novel Soluble Receptor That Binds IL-22 and Neutralizes Its Activity. *The Journal of Immunology* [online]. 15. 6., **166**(12), 7096－7103. ISSN 0022-1767, 1550-6606. Available at: doi:10.4049/jimmunol.166.12.7096

KOTENKO, Sergei V., Grant GALLAGHER, Vitaliy V. BAURIN, Anita LEWIS-ANTES, Meiling SHEN, Nital K. SHAH, Jerome A. LANGER, Faruk SHEIKH, Harold DICKENSHEETS and Raymond P. DONNELLY, 2003. IFN-λs mediate antiviral protection through a distinct class II cytokine receptor complex. *Nature Immunology* [online]. 1., **4**(1), 69－77. ISSN 15292908. Available at: doi:10.1038/ni875

KOTENKO, Sergei V and Jerome A LANGER, 2004. Full house: 12 receptors for 27 cytokines. *International Immunopharmacology* [online]. 5., **4**(5), 593－608. ISSN 15675769. Available at: doi:10.1016/j.intimp.2004.01.003

KOTENKO, Serguei V., Christopher D. KRAUSE, Lara S. IZOTOVA, Brian P. POLLACK, Wei WU and Sidney PESTKA, 1997. Identification and functional characterization of a second chain of the interleukin-10 receptor complex. *The EMBO journal*. **16**(19), 5894－5903.

KRAUSE, Christopher D. and Sidney PESTKA, 2005. Evolution of the Class 2 cytokines and receptors, and discovery of new friends and relatives. *Pharmacology & Therapeutics* [online]. 6., **106**(3), 299–346. ISSN 01637258. Available at: doi:10.1016/j.pharmthera.2004.12.002

KUNZ, Stefanie, Kerstin WOLK, Ellen WITTE, Katrin WITTE, Wolf-Dietrich DOECKE, Hans-Dieter VOLK, Wolfram STERRY, Khusru ASADULLAH and Robert SABAT, 2006. Interleukin (IL)-19, IL-20 and IL-24 are produced by and act on keratinocytes and are distinct from classical ILs. *Experimental Dermatology* [online]. 12., **15**(12), 991–1004. ISSN 0906-6705, 1600-0625. Available at: doi:10.1111/j.1600-0625.2006.00516.x

LADUNGA, I. and R. F. SMITH, 1997. Amino acid substitutions preserve protein folding by conserving steric and hydrophobicity properties. *Protein Engineering*. 3., **10**(3), 187–196. ISSN 0269-2139.

LAM, Yuk-Fai, Danny Ka-Ho WONG, Wai-Kay SETO, Kelvin Kai-Wang TO, Ivan Fan-Ngai HUNG, James FUNG, Ching-Lung LAI and Man-Fung YUEN, 2014. HLA-DP and γ-interferon receptor-2 gene variants and their association with viral hepatitis activity in chronic hepatitis B infection: Genetic variants and hepatitis B viral activity. *Journal of Gastroenterology and Hepatology* [online]. 3., **29**(3), 533–539. ISSN 08159319. Available at: doi:10.1111/jgh.12378

LANDAU, M., I. MAYROSE, Y. ROSENBERG, F. GLASER, E. MARTZ, T. PUPKO and N. BEN-TAL, 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Research* [online]. 1. 7., **33**(Web Server), W299–W302. ISSN 0305-1048, 1362-4962. Available at: doi:10.1093/nar/gki370

LEGENDRE, Pierre, Yves DESDEVISES and Eric BAZIN, 2002. A statistical test for host-parasite coevolution. *Systematic Biology* [online]. 4., **51**(2), 217–234. ISSN 1063-5157. Available at: doi:10.1080/10635150252899734

LI, Q., R. MEANS, S. LANG and J. U. JUNG, 2007. Downregulation of Gamma Interferon Receptor 1 by Kaposi's Sarcoma-Associated Herpesvirus K3 and K5. *Journal of Virology* [online]. 1. 3., **81**(5), 2117–2127. ISSN 0022-538X. Available at: doi:10.1128/JVI.01961-06

LIANG, Spencer C., Xiang-Yang TAN, Deborah P. LUXENBERG, Riyez KARIM, Kyriaki DUNUSSI-JOANNOPOULOS, Mary COLLINS and Lynette A. FOUSER, 2006. Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *The Journal of Experimental Medicine* [online]. 2. 10., **203**(10), 2271–2279. ISSN 0022-1007, 1540-9538. Available at: doi:10.1084/jem.20061308

LIAO, S.-C., Y.-C. CHENG, Y.-C. WANG, C.-W. WANG, S.-M. YANG, C.-K. YU, C.-C. SHIEH, K.-C. CHENG, M.-F. LEE, S.-R. CHIANG, J.-M. SHIEH and M.-S. CHANG, 2004. IL-19 Induced Th2 Cytokines and Was Up-Regulated in Asthma Patients. *The Journal of Immunology* [online]. 1. 12., **173**(11), 6712–6718. ISSN 0022-1767, 1550-6606. Available at: doi:10.4049/jimmunol.173.11.6712

LIN, J.R., H.H. QIN, Y. WANG, J. LIANG and J.H. XU, 2016. Analysis of interleukin 19 serum levels and single nucleotide polymorphisms in systemic lupus erythematosus. *Genetics and Molecular Research* [online]. **15**(2). ISSN 16765680. Available at: doi:10.4238/gmr.15028007

LOGSDON, N. J., A. DESHPANDE, B. D. HARRIS, K. R. RAJASHANKAR and M. R. WALTER, 2012. Structural basis for receptor sharing and activation by interleukin-20 receptor-2 (IL-20R2) binding cytokines. *Proceedings of the National Academy of Sciences* [online]. 31. 7., **109**(31), 12704–12709. ISSN 0027-8424, 1091-6490. Available at: doi:10.1073/pnas.1117551109

LUTFALLA, Georges, Hugues Roest CROLLIUS, Nicole STANGE-THOMANN, Olivier JAILLON, Knud MOGENSEN and Danièle MONNERON, 2003. Comparative genomic analysis reveals independent expansion of a lineage-specific gene family in vertebrates: the class II cytokine receptors

and their ligands in mammals and fish. *BMC genomics* [online]. **4**(1), 29. Available at: doi:10.1186/1471-2164-4-29

MA, Hak-Ling, Spencer LIANG, Jing LI, Lee NAPIERATA, Tom BROWN, Stephen BENOIT, Mayra SENICES, Davinder GILL, Kyriaki DUNUSSI-JOANNOPOULOS, Mary COLLINS, Cheryl NICKERSON-NUTTER, Lynette A. FOUSER and Deborah A. YOUNG, 2008. IL-22 is required for Th17 cell-mediated pathology in a mouse model of psoriasis-like skin inflammation. *Journal of Clinical Investigation* [online]. 17. 1. ISSN 0021-9738. Available at: doi:10.1172/JCI33263

MARLOW, Gareth J, 2013. Why interleukin-10 supplementation does not work in Crohn's disease patients. *World Journal of Gastroenterology* [online]. **19**(25), 3931. ISSN 1007-9327. Available at: doi:10.3748/wjg.v19.i25.3931

MARTIN, J Cj, G BÉRIOU, M HESLAN, C CHAUVIN, L UTRIAINEN, A AUMEUNIER, C L SCOTT, A MOWAT, V CEROVIC, S A HOUSTON, M LEBOEUF, F X HUBERT, C HÉMONT, M MERAD, S MILLING and R JOSIEN, 2014. Interleukin-22 binding protein (IL-22BP) is constitutively expressed by a subset of conventional dendritic cells and is strongly induced by retinoic acid. *Mucosal Immunology* [online]. 1., **7**(1), 101-113. ISSN 1933-0219, 1935-3456. Available at: doi:10.1038/mi.2013.28

MARUKIAN, Svetlana, Linda ANDRUS, Timothy P. SHEAHAN, Christopher T. JONES, Edgar D. CHARLES, Alexander PLOSS, Charles M. RICE and Lynn B. DUSTIN, 2011. Hepatitis C virus induces interferon-λ and interferon-stimulated genes in primary liver cultures. *Hepatology* [online]. 12., **54**(6), 1913-1923. ISSN 02709139. Available at: doi:10.1002/hep.24580

MCGEE, Heather M., Barbara A. SCHMIDT, Carmen J. BOOTH, George D. YANCOPOULOS, David M. VALENZUELA, Andrew J. MURPHY, Sean STEVENS, Richard A. FLAVELL and Valerie HORSLEY, 2013. IL-22 promotes fibroblast-mediated wound repair in the skin. *The Journal of Investigative Dermatology* [online]. 5., **133**(5), 1321-1329. ISSN 1523-1747. Available at: doi:10.1038/jid.2012.463

MELLER, Stephan, Jeremy DI DOMIZIO, Kui S VOO, Heike C FRIEDRICH, Georgios CHAMILOS, Dipyaman GANGULY, Curdin CONRAD, Josh GREGORIO, Didier LE ROY, Thierry ROGER, John E LADBURY, Bernhard HOMEY, Stanley WATOWICH, Robert L MODLIN, Dimitrios P KONTOYIANNIS, Yong-Jun LIU, Stefan T AROLD and Michel GILLIET, 2015. TH17 cells promote microbial killing and innate immune sensing of DNA via interleukin 26. *Nature Immunology* [online]. 13. 7., **16**(9), 970-979. ISSN 1529-2908, 1529-2916. Available at: doi:10.1038/ni.3211

MOUNT, David W., 2008. Maximum Parsimony Method for Phylogenetic Prediction. *Cold Spring Harbor Protocols* [online]. 1. 4., **2008**(4), pdb.top32. ISSN 1940-3402, 1559-6095. Available at: doi:10.1101/pdb.top32

NEEDLEMAN, Saul B. and Christian D. WUNSCH, 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology*. **48**(3), 443-453.

NICE, T. J., M. T. BALDRIDGE, B. T. MCCUNE, J. M. NORMAN, H. M. LAZEAR, M. ARTYOMOV, M. S. DIAMOND and H. W. VIRGIN, 2015. Interferon- cures persistent murine norovirus infection in the absence of adaptive immunity. *Science* [online]. 16. 1., **347**(6219), 269-273. ISSN 0036-8075, 1095-9203. Available at: doi:10.1126/science.1258100

OKONECHNIKOV, Konstantin, Olga GOLOSOVA and Mikhail FURSOV, 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* [online]. 15. 4., **28**(8), 1166-1167. Available at: doi:10.1093/bioinformatics/bts091

OTKJAER, K., K. KRAGBALLE, A. T. FUNDING, J. T. CLAUSEN, P. L. NOERBY, T. STEINICHE and L. IVERSEN, 2005. The dynamics of gene expression of interleukin-19 and interleukin-20 and their receptors in psoriasis. *British Journal of Dermatology* [online]. 11., **153**(5), 911–918. ISSN 0007-0963. Available at: doi:10.1111/j.1365-2133.2005.06800.x

PALOMARES, Oscar, Görkem YAMAN, Ahmet K. AZKUR, Tunc AKKOC, Mübeccel AKDIS and Cezmi A. AKDIS, 2010. Role of Treg in immune regulation of allergic diseases. *European Journal of Immunology* [online]. 10. 2., **40**(5), 1232–1240. ISSN 00142980. Available at: doi:10.1002/eji.200940045

PARADIS, Emmanuel, Julien CLAUDE and Korbinian STRIMMER, 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics (Oxford, England)* [online]. 22. 1., **20**(2), 289–290. ISSN 1367-4803. Available at: doi:10.1093/bioinformatics/btg412

PATTERSON, John B., Daniel C. THOMIS, Sherrie L. HANS and Charles E. SAMUEL, 1995. Mechanism of Interferon Action: Double-Stranded RNA-Specific Adenosine Deaminase from Human Cells Is Inducible by Alpha and Gamma Interferons. *Virology* [online]. 10. 7., **210**(2), 508–511. ISSN 0042-6822. Available at: doi:10.1006/viro.1995.1370

PILS, Birgit, Richard R. COPLEY and Jörg SCHULTZ, 2005. Variation in structural location and amino acid conservation of functional sites in protein domain families. *BMC Bioinformatics* [online]. **6**, 210. ISSN 1471-2105. Available at: doi:10.1186/1471-2105-6-210

PRESTWOOD, T. R., M. M. MORAR, R. M. ZELLWEGER, R. MILLER, M. M. MAY, L. E. YAUCH, S. M. LADA and S. SHRESTA, 2012. Gamma Interferon (IFN-) Receptor Restricts Systemic Dengue Virus Replication and Prevents Paralysis in IFN- / Receptor-Deficient Mice. *Journal of Virology* [online]. 1. 12., **86**(23), 12561–12570. ISSN 0022-538X. Available at: doi:10.1128/JVI.06743-11

REVELL, Liam J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* [online]. 1. 4., **3**(2), 217–223. ISSN 2041-210X. Available at: doi:10.1111/j.2041-210X.2011.00169.x

ROBEK, M. D., B. S. BOYD and F. V. CHISARI, 2005. Lambda Interferon Inhibits Hepatitis B and C Virus Replication. *Journal of Virology* [online]. 15. 3., **79**(6), 3851–3854. ISSN 0022-538X. Available at: doi:10.1128/JVI.79.6.3851-3854.2005

RONQUIST, F. and J. P. HUELSENBECK, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* [online]. 12. 8., **19**(12), 1572–1574. ISSN 1367-4803, 1460-2059. Available at: doi:10.1093/bioinformatics/btg180

RONQUIST, F., M. TESLENKO, P. VAN DER MARK, D. L. AYRES, A. DARLING, S. HOHNA, B. LARGET, L. LIU, M. A. SUCHARD and J. P. HUELSENBECK, 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology* [online]. 1. 5., **61**(3), 539–542. ISSN 1063-5157, 1076-836X. Available at: doi:10.1093/sysbio/sys029

SAHIRATMADJA, E., B. ALISJAHBANA, T. DE BOER, I. ADNAN, A. MAYA, H. DANUSANTOSO, R. H. H. NELWAN, S. MARZUKI, J. W. M. VAN DER MEER, R. VAN CREVEL, E. VAN DE VOSSE and T. H. M. OTTENHOFF, 2007. Dynamic Changes in Pro- and Anti-Inflammatory Cytokine Profiles and Gamma Interferon Receptor Signaling Integrity Correlate with Tuberculosis Disease Activity and Response to Curative Treatment. *Infection and Immunity* [online]. 1. 2., **75**(2), 820–829. ISSN 0019-9567. Available at: doi:10.1128/IAI.00602-06

SCHOENBORN, Jamie R. and Christopher B. WILSON, 2007. Regulation of Interferon-γ During Innate and Adaptive Immune Responses. In: *Advances in Immunology* [online]. B.m.: Elsevier, p. 41–101. ISBN 978-0-12-373709-0. Available at: doi:10.1016/S0065-2776(07)96002-2

SHAH, Neil, Jochen KAMMERMEIER, Mamoun ELAWAD and Erik-Oliver GLOCKER, 2012. Interleukin-10 and Interleukin-10-Receptor Defects in Inflammatory Bowel Disease. *Current Allergy and Asthma Reports* [online]. 10., **12**(5), 373–379. ISSN 1529-7322. Available at: doi:10.1007/s11882-012-0286-z

SHEAHAN, Timothy, Naoko IMANAKA, Svetlana MARUKIAN, Marcus DORNER, Peng LIU, Alexander PLOSS and Charles M. RICE, 2014. Interferon Lambda Alleles Predict Innate Antiviral Immune Responses and Hepatitis C Virus Permissiveness. *Cell Host & Microbe* [online]. 2., **15**(2), 190–202. ISSN 19313128. Available at: doi:10.1016/j.chom.2014.01.007

SHIM, Jung Ok and Jeong Kee SEO, 2014. Very early-onset inflammatory bowel disease (IBD) in infancy is a different disease entity from adult-onset IBD; one form of interleukin-10 receptor mutations. *Journal of Human Genetics* [online]. 6., **59**(6), 337–341. ISSN 1434-5161. Available at: doi:10.1038/jhg.2014.32

SOLOVYEV, V., 2001. Statistical approaches in eukaryotic gene prediction. *Handbook of statistical genetics*.

STAMATAKIS, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* [online]. 1. 5., **30**(9), 1312–1313. ISSN 1367-4803, 1460-2059. Available at: doi:10.1093/bioinformatics/btu033

STAMATAKIS, Alexandros, Paul HOOVER and Jacques ROUGEMONT, 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Systematic Biology* [online]. 10., **57**(5), 758–771. ISSN 1076-836X. Available at: doi:10.1080/10635150802429642

STÖVER, Ben C. and Kai F. MÜLLER, 2010. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* [online]. **11**, 7. ISSN 1471-2105. Available at: doi:10.1186/1471-2105-11-7

SUN, Ding-Ping, Ching-Hua YEH, Edmund SO, Li-Yun WANG, Tsui-Shan WEI, Ming-Shi CHANG and Chung-Hsi HSING, 2013. Interleukin (IL)-19 promoted skin wound healing by increasing fibroblast keratinocyte growth factor expression. *Cytokine* [online]. 6., **62**(3), 360–368. ISSN 10434666. Available at: doi:10.1016/j.cyto.2013.03.017

THIEL, D. J., M. H. LE DU, R. L. WALTER, A. D'ARCY, C. CHÈNE, M. FOUNTOULAKIS, G. GAROTTA, F. K. WINKLER and S. E. EALICK, 2000. Observation of an unexpected third receptor molecule in the crystal structure of human interferon-gamma receptor complex. *Structure (London, England: 1993)*. 15. 9., **8**(9), 927–936. ISSN 0969-2126.

VALDAR, William SJ and Janet M. THORNTON, 2001. Protein-protein interfaces: analysis of amino acid conservation in homodimers. *Proteins: Structure, Function, and Bioinformatics*. **42**(1), 108–124.

VANDENBROECK, K., S. CUNNINGHAM, A. GORIS, I. ALLOZA, S. HEGGARTY, C. GRAHAM, A. BELL and M. ROONEY, 2003. Polymorphisms in the interferon-γ/interleukin-26 gene region contribute to sex bias in susceptibility to rheumatoid arthritis: *IFNG/IL26* Polymorphisms and Sex Bias in Susceptibility to RA. *Arthritis & Rheumatism* [online]. 10., **48**(10), 2773–2778. ISSN 00043591. Available at: doi:10.1002/art.11236

WANG, F., N. SMITH, L. MAIER, W. XIA, C. HAMMERBERG, H. CHUBB, C. CHEN, M. RIBLETT, A. JOHNSTON, J.E. GUDJONSSON, Y. HELFRICH, S. KANG, G.J. FISHER and J.J. VOORHEES, 2012. Etanercept suppresses regenerative hyperplasia in psoriasis by acutely downregulating epidermal expression of interleukin (IL)-19, IL-20 and IL-24: Etanercept acutely suppresses the IL-20 cytokine subfamily. *British Journal of Dermatology* [online]. 7., **167**(1), 92–102. ISSN 00070963. Available at: doi:10.1111/j.1365-2133.2012.10961.x

WANG, Lu, Yan WANG, Zhiyu SONG, Jiahui CHU and Xianjun QU, 2015. Deficiency of Interferon-Gamma or Its Receptor Promotes Colorectal Cancer Development. *Journal of Interferon & Cytokine Research* [online]. 4., **35**(4), 273–280. ISSN 1079-9907, 1557-7465. Available at: doi:10.1089/jir.2014.0132

WANG, M., Z. J. TAN, R. ZHANG, S. V. KOTENKO and P. LIANG, 2002. Interleukin 24 (MDA-7/MOB-5) signals through two heterodimeric receptors, IL-22R1/IL-20R2 and IL-20R1/IL-20R2. *Journal of Biological Chemistry* [online]. 1. 3., **277**(9), 7341–7347. ISSN 0021-9258. Available at: doi:10.1074/jbc.M106043200

WANG, Shimin, Xiang GAO, Guobo SHEN, Wei WANG, Jingyu LI, Jingyi ZHAO, Yu-Quan WEI and Carl K. EDWARDS, 2016. Interleukin-10 deficiency impairs regulatory T cell-derived neuropilin-1 functions and promotes Th1 and Th17 immunity. *Scientific Reports* [online]. 14. 4., **6**, 24249. ISSN 2045-2322. Available at: doi:10.1038/srep24249

WATERHOUSE, A. M., J. B. PROCTER, D. M. A. MARTIN, M. CLAMP and G. J. BARTON, 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* [online]. 1. 5., **25**(9), 1189–1191. ISSN 1367-4803, 1460-2059. Available at: doi:10.1093/bioinformatics/btp033

WHELAN, Simon and Nick GOLDMAN, 2001. A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. *Molecular Biology and Evolution* [online]. 1. 5., **18**(5), 691–699. ISSN 0737-4038. Available at: doi:10.1093/oxfordjournals.molbev.a003851

WILSON, Nicholas J, Katia BONIFACE, Jason R CHAN, Brent S MCKENZIE, Wendy M BLUMENSCHEIN, Jeanine D MATTSON, Beth BASHAM, Kathleen SMITH, Taiying CHEN, Franck MOREL, Jean-Claude LECRON, Robert A KASTELEIN, Daniel J CUA, Terrill K MCCLANAHAN, Edward P BOWMAN and Rene DE WAAL MALEFYT, 2007. Development, cytokine profile and function of human interleukin 17–producing helper T cells. *Nature Immunology* [online]. 9., **8**(9), 950–957. ISSN 1529-2908. Available at: doi:10.1038/ni1497

WOLK, K., S. KUNZ, K. ASADULLAH and R. SABAT, 2002. Cutting Edge: Immune Cells as Sources and Targets of the IL-10 Family Members? *The Journal of Immunology* [online]. 1. 6., **168**(11), 5397–5402. ISSN 0022-1767, 1550-6606. Available at: doi:10.4049/jimmunol.168.11.5397

WOLK, Kerstin, Stefanie KUNZ, Ellen WITTE, Markus FRIEDRICH, Khusru ASADULLAH and Robert SABAT, 2004. IL-22 increases the innate immunity of tissues. *Immunity* [online]. **21**(2), 241–254. Available at: doi:http://dx.doi.org/10.1016/j.immuni.2004.07.007

WU, Jinxiang, Guicheng WANG, Junqing HAO, Wenbin GONG, Junfei WANG, Jiping ZHAO and Dong LIANG, 2014. The correlation between IL-20 and the Th2 immune response in human asthma. *Asian Pacific journal of allergy and immunology* [online]. **32**(4), 316. Available at: doi:110.12932/AP0447.32.4.2014

XU, W. F., S. R. PRESNELL, J. PARRISH-NOVAK, W. KINDSVOGEL, S. JASPERS, Z. CHEN, S. R. DILLON, Z. GAO, T. GILBERT, K. MADDEN, S. SCHLUTSMEYER, L. YAO, T. E. WHITMORE, Y. CHANDRASEKHER, F. J. GRANT, M. MAURER, L. JELINEK, H. STOREY, T. BRENDER, A. HAMMOND, S. TOPOUZIS, C. H. CLEGG and D. C. FOSTER, 2001. A soluble class II cytokine receptor, IL-22RA2, is a naturally occurring IL-22 antagonist. *Proceedings of the National Academy of Sciences of the United States of America* [online]. 14. 8., **98**(17), 9511–9516. ISSN 0027-8424. Available at: doi:10.1073/pnas.171303198

YANG, C., Y. TONG, W. NI, J. LIU, W. XU, L. LI, X. LIU, H. MENG and W. QIAN, 2010. Inhibition of autophagy induced by overexpression of mda-7/interleukin-24 strongly augments the

antileukemia activity in vitro and in vivo. *Cancer Gene Therapy* [online]. 2., **17**(2), 109–119. ISSN 0929-1903. Available at: doi:10.1038/cgt.2009.57

YOON, Sung Il, Brandi C. JONES, Naomi J. LOGSDON and Mark R. WALTER, 2005. Same Structure, Different Function. *Structure* [online]. 1. 4., **13**(4), 551–564. ISSN 0969-2126. Available at: doi:10.1016/j.str.2005.01.016

ZDANOV, A., C. SCHALK-HIHI, A. GUSTCHINA, M. TSANG, J. WEATHERBEE and A. WLODAWER, 1995. Crystal structure of interleukin-10 reveals the functional dimer with an unexpected topological similarity to interferon gamma. *Structure (London, England: 1993)*. 15. 6., **3**(6), 591–601. ISSN 0969-2126.

ZHANG, Caibo, Dong HOU, Haifeng WEI, Minnan ZHAO, Lin YANG, Qiao LIU, Xiyu ZHANG, Yaoqin GONG and Changshun SHAO, 2016a. Lack of interferon-γ receptor results in a microenvironment favorable for intestinal tumorigenesis. *Oncotarget* [online]. **7**(27), 42099. Available at: doi:10.18632/oncotarget.9867

ZHANG, Di, Alexander WLODAWER and Jacek LUBKOWSKI, 2016b. Crystal Structure of a Complex of the Intracellular Domain of Interferon λ Receptor 1 (IFNLR1) and the FERM/SH2 Domains of Human JAK1. *Journal of Molecular Biology* [online]. 11., **428**(23), 4651–4668. ISSN 00222836. Available at: doi:10.1016/j.jmb.2016.10.005

ZHENG, Mingzhong, Dora BOCANGEL, Blair DONESKE, Abner MHASHILKAR, Rajagopal RAMESH, Kelly K. HUNT, Suhendan EKMEKCIOGLU, R. Bryan SUTTON, Nancy POINDEXTER, Elizabeth A. GRIMM and Sunil CHADA, 2006. Human interleukin 24 (MDA-7/IL-24) protein kills breast cancer cells via the IL-20 receptor and is antagonized by IL-10. *Cancer Immunology, Immunotherapy* [online]. 27. 11., **56**(2), 205–215. ISSN 0340-7004, 1432-0851. Available at: doi:10.1007/s00262-006-0175-1

ZHENG, Yan, Dimitry M. DANILENKO, Patricia VALDEZ, Ian KASMAN, Jeffrey EASTHAM-ANDERSON, Jianfeng WU and Wenjun OUYANG, 2007. Interleukin-22, a TH17 cytokine, mediates IL-23-induced dermal inflammation and acanthosis. *Nature* [online]. 8. 2., **445**(7128), 648–651. ISSN 0028-0836, 1476-4687. Available at: doi:10.1038/nature05505

7.1 Online resources

FigTree software available from http://tree.bio.ed.ac.uk/

ConSurf available at http://consurf.tau.ac.il/

NCBI BLAST available at https://blast.ncbi.nlm.nih.gov/Blast.cgi

NCBI database at https://www.ncbi.nlm.nih.gov/protein/

Phylemon 2 webserver at http://phylemon2.bioinfo.cipf.es/

Protein Data Bank available at http://www.rcsb.org/pdb

UniPROT database at http://www.uniprot.org/

Supplementary material

FGENESH+ sequences

>FGENESH: Anolis carolinensis | IL20

MAFGAFSCLVLVAFLFAKTVVAEGRRLSLGQCELNSVSFRELRDNFDAIKENVQTQDI
RTDVILLKESVLREVPMSESCCLLRHLLRFYVESIFKHYEPTSNLLRRKTSTLANAFLSI
KAKLRECHNQNKCSCGEETNRRFKLVLDEYQKLDKTTAAIKSLGEMDVLFAWMEGF

>FGENESH: Gallus gallus | IL26

MKVYSIFRSGHLLVLLCLFTVEGKKSPTGKHTCRKGLLSQVTENLYTKASSLKSSVPKD
LIKNTRLLKKTTKMLFMTNCNVRDQLLSFYMKNVFSHLGMESEKLFVISAFRVLQE
NMNACLPCAPSTRLTSAVKNIKKTFLKVRVGGVGSCFGGVGGFTSNTFIFLTAWGEGG
LQGHQ

>FGENESH: Pelodiscus sinensis | IFNLR1

MSAGSRAVLVALCSFQQLLGSVALGQPGVPLPPPRNVKLLSKDFGVAVTWLPGEGSPP
DVLYSVRYQTLYHQSNWKQVRHCKNISHVTCNLTCGPDPYNKFSTRVKALAAGRQS
PWVESNSLEYHLDVHLAPPALAVSVAETTINVSATFPLASCVKSVFIGLKYDLDFWKAG
TGDKVPFHDRMKWENVTISTLALSGNYCLSARASYQAIQLKHSQFSRPLCMLLTPRA
KGWEFLITMAVPLLILLFFCTAPGTVLEELIERDLFICVVQPASAGRWRSDASRTARND
TSLVARNNASPVARNDASPVARNDTSPVARNDTSLTASLLSLSEEEDDDSGGRPYTEMP
LFLRRAPNCSGASMSQEGSHSGSELSGSHLAGGPVPDLAGLGFSRLVWRGGPAEEDAS
GFPDSEKSSSFSESSSVGEFSLSEAPCPVTCGGERQGWEADTGQEDPFLQVSVLAEGLK
GGSPAEEWGVPRRGPRKTDPQRHLHPDPSVCVARGVSEAADGFPLEEQLVRFQTVKL
ALDEGVASDSESLAGGAERDPPPLSAALSETGGAEAWGKGGGLWPARDPAWQCRGY
QHMRYMPRT