Universidade de Aveiro Departamento de Biologia
2010

Márcio Adriano Guiomar de Oliveira

Receptores do factor estimulante de colónias de macrófagos do robalo

Sea bass macrophage colony stimulating factor receptors

Universidade de Aveiro Departamento de Biologia
2010

Márcio Adriano Guiomar de Oliveira

Receptores do factor estimulante de colónias de macrófagos do robalo

Sea bass macrophage colony stimulating factor receptors

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Molecular e Celular, realizada sob a orientação científica do Doutor Nuno Miguel Simões dos Santos, Investigador Auxiliar do Instituto de Biologia Molecular e Celular da Universidade do Porto e do Doutor António Carlos Matias Correia, Professor Catedrático do Departamento de Biologia da Universidade de Aveiro.

Apoio financeiro da Fundação para a Ciência e a Tecnologia e do Fundo Social Europeu

PTDC/CVT/69086/2006

o júri

presidente

Prof.ª. Doutora Maria de Lourdes Gomes Pereira
Professora Associada com Agregação do Departamento de Biologia da Universidade de Aveiro

Prof.ª. Doutora Ana Maria Luís Ramos Tomás
Professora Associada do Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Doutor Nuno Miguel Simões dos Santos
Investigador Auxiliar do Instituto de Biologia Molecular e Celular da Universidade do Porto

Prof. Doutor António Carlos Matias Correia
Professor Catedrático do Departamento de Biologia da Universidade de Aveiro

agradecimentos

Ao Doutor Nuno dos Santos por me ter dado a oportunidade, e a liberdade, de realizar este trabalho. Obrigado pela orientação científica e pelo incentivo.

A todos os membros, actuais e passados, do nosso grupo de Immunologia e Vacinologia de Peixes que contribuíram para este trabalho.

Um especial agradecimento aos membros mais seniores do grupo, Ana do Vale e Carolina Costa-Ramos, pelo confronto de ideias e pela ajuda.

Um grande obrigado às minhas colegas Daniela, Marta, Rute, Marisa, Eunice e Raquel pela ajuda e amizade.

Aos vizinhos da Divisão de Biologia da Infecção e Imunologia que ajudaram a animar o dia-a-dia e a equilibrar a “guerra dos sexos” da vizinhança.

Aos meus pais, Celeste e Alfredo, a minha eterna gratidão.

Aos meus irmãos, Liliana e Cláudio, que são o meu apoio, um grande obrigado.

À Sara, com quem partilho a minha vida, as minhas angústias mas também os momentos de felicidade e de prazer, obrigado pela tua amizade e amor.

Partilho este pequeno sucesso com vocês.

palavras-chave Robalo, *Dicentrarchus labrax*, CSF1R, CSF1R1/CSF1Ra, CSF1R2/CSF1Rb

resumo

O receptor do factor estimulante de colónias de macrófagos, também conhecido como receptor do factor estimulante de colónias-1 (CSF1R), é um receptor de um factor de crescimento hematopoiético que é especificamente expresso em células do sistema fagocítico-mononuclear e desempenha um papel essencial no desenvolvimento e regulação destas células. O CSF1R já foi descrito em vários mamíferos e a sua biologia tem sido exaustivamente caracterizada nestes vertebrados mas o conhecimento sobre esta molécula em peixes é ainda muito reduzido.

Neste trabalho é descrita a sequenciação e caracterização de duas formas do CSF1R de robalo (*Dicentrarchus labrax* L.). Foram isolados dois cDNAs diferentes que codificam proteínas homólogas ao CSF1R de outros vertebrados. Um cDNA de 4535 bp, com uma *open reading frame* (ORF) de 2946 bp que codifica uma proteína de 981 aminoácidos homóloga a proteínas CSF1R1/CSF1Ra já descritas em peixes e um cDNA de 3229 bp, com uma ORF de 2817 bp que codifica uma proteína de 938 aminoácidos homóloga a proteínas CSF1R2/CSF1Rb já descritas em peixes. Ambas as proteínas caracterizadas conservam domínios e aminoácidos chave que são funcional e estruturalmente importantes nos CSF1Rs de mamíferos.

A identificação destes receptores irá permitir uma melhor compreensão da biologia das células do sistema fagocítico-mononuclear do robalo. Nomeadamente, poderão ser usados como marcadores específicos deste tipo de células e permitir estudos mais detalhados sobre a immunologia do robalo e dos teleósteos em geral.

keywords Sea bass, *Dicentrarchus labrax*, CSF1R, CSF1R1/CSF1Ra, CSF1R2/CSF1Rb

abstract

The macrophage colony stimulating factor receptor, also known as colony stimulating factor-1 receptor (CSF1R), is a hematopoietic growth factor receptor that is specific of the cells of the mononuclear phagocytic lineage and has a key role in its development and regulation. Several mammalian CSF1R proteins have been described and its biology has been extensively characterized but in fish, knowledge about this receptor is still scarce.

We have sequenced and characterized the sea bass (*Dicentrarchus labrax* L.) CSF1R molecules. Two different cDNAs coding for proteins homologous to vertebrate CSF1R proteins were isolated. A 4535 bp cDNA, with a 2946 bp open reading frame (ORF) that codes for a 981 amino acid protein, homologous to known bony fish CSF1R1/CSF1Ra proteins and a 3229 bp cDNA, with a 2817 bp ORF that codes for a 938 amino acid protein, homologous to known bony fish CSF1R2/CSF1Rb molecule. Both receptors display a high degree of conservation of domains and key amino acids known to be important functional and structural features of the mammalian counterparts.

The identification of these receptors will allow for a better understanding of the biology of sea bass mononuclear phagocytic cells. They can be used as specific markers of sea bass monocytes and macrophages and enable more detailed studies on sea bass immunity.

Table of Contents

Abbreviations

3R	third genome duplication
AIP56	apoptosis inducing protein of 56 kDa
APSE-2	bacteriophage 2 of A*cyrthosiphon pisum* secondary endosymbiont
C2	constant-2
Cbl	ubiquitin-protein ligase Casitas B lineage
CD11b	leukocyte integrin α subunit
CD16	Fc gamma receptor III
cDNA	coding DNA
c-fms	cellular homolog of the McDonough feline sarcoma virus oncogene (same as CSF1R)
CK	chemokine
CSF-1	colony stimulating factor 1 (same as M-CSF)
csf1r(a/b)	colony stimulating factor 1 receptor (a or b)
CSF1R	colony stimulating factor 1 receptor
CXCL8	α-chemokine interleukin-8
E. coli	*Escherichia coli*
ECPs	extracellular products
FMIP	FMS-interacting protein
FSGD	fish-specific genome duplication
Grb2	growth factor receptor-bound protein 2
HLA	Human leukocyte antigen
IFNγ	interferon gamma
Ig-like	immunoglobulin-like
IL	interleukin
γIP-10	interferon-inducible protein-10
kDa	kilodalton
KIT	stem cell factor receptor
kit(a/b)	mast/stem cell growth factor receptor (a or b)
kitl(a/b)	kit ligand (a or b)
LPS	lipopolysaccharide
M-CSF	macrophage colony stimulating factor (same as CSF-1)
MMC	melanomacrophage centres
MONA	monocyte adaptor
OD	optical density
PBS	phosphate buffered saline
PDGF	platelet-derived growth factor
pdgfrβ	platelet derived growth factor receptor beta
Pen-Strep	penicillin-streptomycin solution
Phdp	*Photobacterium damselae* subsp. *piscicida*
PI3K	phosphatidylinositol 3-kinase
PLCγ	phospholipase C-gamma
RACE	rapid amplification of cDNA ends

ROS	reactive oxygen species
RTK	receptor tyrosine kinase
sbPBS	sea bass phosphate buffered saline
SFK	Src family kinase
SOCS1	supressor of cytokine signaling-1
$t_{1/2}$	half-life time
TACE	tumour necrosis factor alpha converting enzyme
TNF-(α/β)	tumour necrosis factor (alpha or beta)
UTR	untranslated region

Introduction

European seabass (*Dicentrachus labrax* L.)

Sea bass, also known as european seabass, is a ray-finned fish, member of the Actinopterygii class, Perciformes order and Moronidae family. Adult sea bass are gregarious fish that inhabit coastal waters down to a depth of about 100 m, with temperatures ranging 5-28 ºC and salinities from $3^0/_{00}$ to maximum salinity of sea water (in some Mediterranean zones can be as high as $38^0/_{00}$). Their high tolerance to salinity changes allows them to frequent inshore brackish waters or even fresh water courses. Voracious predators, when young they feed mostly on shrimps, crabs, mollusks and other invertebrates, and as adults their diet shifts to be increasingly composed by fish (FAO 2004-2010).

Sea bass inhabit mainly the North Sea and the seas surrounding the British Isles. It can also be found in the Baltic and Norwegian seas, in the northeastern Atlantic Ocean from Iceland to Senegal and in the Mediterranean and Black seas (FAO 2004-2010; Froese and Pauly 2010).

Sea bass aquaculture

The capture of wild sea bass has increased from 300 tonnes in 1950 to more than 8500 tonnes in 2008, though currently, the vast majority of consumed sea bass is from aquaculture (FAO 2004-2010).

Until the early 1970's, sea bass aquaculture was relatively insignificant, it had been mainly performed in evaporation pans and marshes, integrated with traditional marine salt production. Since the mid 1970's, with the development of modern hatcheries, the culture of sea bass has increased significantly having reached a maximum of 66000 tonnes in the year 2008 with an annual production of more than 43000 tonnes since 2000 (FAO 2004-2010).

In the Mediterranean region, sea bass is one of the most prized fishes and consumer demand has caused and increase in sea bass culture; currently, sea bass culture is performed in Greece, Italy, Spain, Croatia, Egypt Morocco, Algeria, Tunisia, Israel, Cyprus, Malta, Slovenia, Bosnia and Herzegovina, France, Germany and Portugal (FAO 2004-2010).

Main sea bass diseases

Fish are affected by a diverse range of pathogens such as bacteria, virus and parasites. Diseases can have devastating effects in reared fish, especially when it comes to intensive production in which fish stress levels are high.

Cultured sea bass is affected by bacterial infections such as vibriosis (caused by *Vibrio anguillarum* and other *Vibrio* species), myxobacteriosis (caused by *Tenacibaculum maritimum*) and photobacteriosis (caused by *Photobacterium damselae* subsp. *piscicida*), viral infections like the one caused by Nodavirus and several parasitic infections caused by protozoans (ciliates and myxosporidia), dinoflagellates and other eukaryotic organisms such as trematodes, nematodes and crustaceans (FAO 2004-2010).

Fish photobacteriosis – a major bacterial disease

Photobacteriosis, formerly known as pasteurellosis, is an important bacterial disease that affects several fish species, namely sea bass, sea bream, yellowtail, striped bass, white perch and others (reviewed in (Magariños, Toranzo et al. 1996)). Acute photobacteriosis is associated with rapid and high mortality and generally has few external symptoms. Skin of the dorsal and lateral surfaces may become darker and discrete hemorrhagic events may occur in the head and gills. Internally, organs like the liver, kidney and spleen are heavily affected and display signs of necrosis and bacterial accumulation. In some species, it may occur a sub-acute/chronic infection characterized by the formation of granulomatous white tubercles, ranging from 0.5 to 3.5 mm in diameter, containing bacteria and phagocytes undergoing secondary necrosis (reviewed in (Magariños, Toranzo et al. 1996)).

The etiological agent of fish photobacteriosis, *Photobacterium damselae* subsp. *piscicida* (*Phdp*; formerly known as *Pasteurella piscicida*), is a gram-negative bacteria characterized by a straight rod-like shape and absence of flagella (Gauthier, Lafay et al. 1995; Trüper and De' Clari 1997). Growth is observed at 25 and 35 ºC with NaCl concentrations ranging from 0.5 to 3% (Gauthier, Lafay et al. 1995).

Phdp is part of the *Vibrionaceae* family (Gauthier, Lafay et al. 1995), an extremely diverse phylogenetic group whose members are abundant in aquatic environments, either as free-swimming planktonic organisms or associated with other aquatic organisms.

It was suggested that infection by *Phdp* occurs orally, through the gills or through the gastro-intestinal tract (reviewed in (Magariños, Toranzo et al. 1996)). Several strains of *Phdp* bind readily to the intestines of gilthead sea bream, sea bass and turbot (Magariños, Romalde et

al. 1996). It was shown that *Phdp* was capable of adhere, invade and survive in *Sparus aurata* derived epithelial cells (Acosta, Vivas et al. 2009) and cyprinid epithelial cells (Lopez-Doriga, Barnes et al. 2000), indicating that infection can also occur through the skin.

The extracellular products (ECPs) of *Phdp* were shown to have extremely deleterious effects on fish (Magariños, Santos et al. 1992; Noya, Magariños et al. 1995; Bakopoulos, Hanif et al. 2004). Sea bass injected with *Phdp* ECPs showed inflammatory and necrotic lesions in the spleen, liver, head kidney, intestine and heart just 2 days after injection (Bakopoulos, Hanif et al. 2004). The spleens of gilthead sea breams injected with ECPs displayed major necrosis at 3 days post-injection (Noya, Magariños et al. 1995).

Phdp ECPs were shown to be extremely toxic and sufficient to display the full *Phdp* pathology. It was verified that the ECPs' cytotoxicity was not correlated with the proteolytic activities of putative secreted enzymes (Magariños, Santos et al. 1992). In fact, the key virulence factor of *Phdp* was shown to be an exotoxin, the apoptosis inducing protein of 56 kDa (AIP56) (do Vale, Silva et al. 2005).

AIP56 is encoded by a high-copy plasmid that is present in all *Phdp* virulent strains. Blast analysis of the protein revealed a close proximity with a hypothetical protein of *Vibrio campbellii*, another major aquatic pathogen. The N-terminal portion of AIP56, which includes a zinc-metalloprotease signature, revealed similarity with type III secreted effectors C from enteric pathogenic bacteria and the C-terminal portion of AIP56 revealed similarity with a hypothetical protein of the temperate lambda-like bacteriophage APSE-2. This homology analysis suggested a modular structure that was further supported by limited proteolysis experiments that revealed two digestion fragments similar to the ones predicted by homology analysis. Cleavage of AIP56 occurs between two cysteine residues that were shown to form a disulfide bridge. These observations supported the hypothesis that AIP56 is an AB-toxin and preliminary data suggested that the C-terminal subunit is responsible for the toxin binding and entry into cells and the N-terminal subunit is responsible for the pathogenic activity.

AIP56 alone was shown to provoke known hallmarks of pasteurellosis such as systemic phagocyte destruction (reviewed in (Silva, dos Santos et al. 2010)).

Phdp immunosuppressive effect – targeting the defenses

At the cellular level, it was verified that *Phdp* and/or its ECPs could induce the apoptotic death of sea bass neutrophils and macrophages both *in vivo*, under natural and experimental infections, and *ex vivo* (do Vale, Marques et al. 2003; do Vale, Silva et al. 2005; do Vale,

Costa-Ramos et al. 2007). A full range of apoptosis hallmarks were observed in professional phagocytes of fish that were injected either with *Phdp* or with *Phdp* ECPs, namely DNA and nuclear fragmentation, chromatin condensation, cell shrinkage, membrane blebbing and shedding of apopotic bodies (do Vale, Marques et al. 2003). Further *ex vivo* studies showed that the exotoxin AIP56 induced the same apoptotic phenotype in phagocytes. It was also shown that the apoptosis induced by AIP56 involves the caspases 3, 8 and 9, loss of mitochondrial membrane potential, translocation of cytochrome c and increased production of reactive oxygen species (ROS) (reviewed in (Silva, dos Santos et al. 2010)).

The targeted cells, neutrophils and macrophages, are key components of the innate immune response. They are the first line of defense against microbial pathogens, being able to identify and destroy myriad pathogens without the intervention of an adaptive immune response. Macrophages are also essential to make the transition between the innate and adaptive immune response. Mature macrophages, derived from blood circulating monocytes, reside in tissues throughout the body and are usually the first specialized cells to come in contact with bacteria and recognize them as invading organisms. After recognition of pathogens by macrophages, the short-lived neutrophils are recruited from the blood to sites of infection (Travers, Walport et al. 2008). Macrophages and neutrophils have the capacity to engulf bacteria and contain several hydrolytic enzymes in their lysosomes which are essential to degrade bacteria. Macrophages have the ability to produce nitric oxide and, along with neutrophils, produce ROS which are potent bactericides (reviewed in (Ellis 1999)). The direct targeting of phagocytes by AIP56 is not only an evading strategy but also a major pathological component of *Phdp* infections. Macrophages assist in the homeostasis of the apoptotic process by clearing dying cells and apoptotic bodies. Given that macrophages themselves are directly affected by AIP56, the apoptotic process is deregulated and usually progresses to secondary necrosis, increasing the release of cytotoxic components that in turn will have non-specific noxious effects and trigger an inflammatory response, leading to tissue damage (do Vale, Marques et al. 2003; do Vale, Costa-Ramos et al. 2007).

Fish immune system

Teleost fish comprise a wide array of immune defenses that is comparable to those of other vertebrates. Fish immune cellular components include macrophages, granulocytes, thrombocytes, mast cells, nonspecific cytotoxic cell, dendritic cells, B and T lymphocytes (reviewed in (Press and Evensen 1999)). Several cellular immune activities are described for

fish, including macrophage and neutrophil antibacterial reactions, non-specific cell-mediated cytotoxicity, B-cell and T-cell activities (reviewed in (Scapigliati, Romano et al. 2002)).

The fish humoral immune system, equivalent to those of mammals, comprises direct defense mechanisms and mediators of the cellular immune activities.

Fish have a complement system that is similar to the mammalian one; it can be activated by the classical pathway, the alternative pathway and the lectin pathway (reviewed in (Whyte 2007)). Complement activation leads to the lytic pathway, opsonization and increased phagocytosis. Associated with the alternative pathway, the haemolytic serum activity of teleosts is considered to be more active than in mammals. Lectins homologous to the mammalian ones were also isolated from several fish species and antimicrobial polypetides have been identified in the mucus, liver and gills of fish. Naturally occurring antibodies and specific antibody responses were also observed in fish (reviewed in (Whyte 2007)).

A network of humoral mediators composed by a variety of substances such as growth factors, cytokines and chemokines has begun to be unveiled. The cytokines TNF-α and TNF-β, IL-1β, IL-2, IL-4, IL-6, IL-10, IL-11, IL-12, IL-15, IL-18, IL-21, IL-22, IL-26, IFNγ and the chemokines CXCL8 (IL-8), γIP-10, CK-1 and CK-2 have already been characterized in fish (reviewed in (Whyte 2007)).

Although fish possess similar cellular and humoral agents to those of mammals, morphologically the immune system of fish is quite dissimilar to the mammalian system. The main anatomical structures that compose the fish immune system are the thymus, head kidney, spleen and mucosal associated lymphoid tissue (reviewed in (Scapigliati, Romano et al. 2002)). The extreme diversity of teleosts is also reflected in the morphological and histological characteristics of the referred organs that are often significantly different between species.

The head kidney (the anterior portion of the kidney that is located in the cranial region) is considered to be the teleost equivalent of the mammalian bone marrow (Meseguer, Lopez-Ruiz et al. 1995). Its morphological characteristics are variable from species to species, in some cases the head kidney is practically indistinguishable from the more posterior trunk kidney, in other species it has a more distinctive appearance consisting of two lobes that protrude from the trunk kidney (Bjerkas, Evensen et al. 2006). Histologically, the head kidney is practically devoid of excretory tissue and its mainly composed by hematopoietic and lymphoid tissue (Bjerkas, Evensen et al. 2006). Hematopoietic processes such as erythropoiesis, granulopoiesis, thrombopoiesis and, at less extent, lymphoplasmapoiesis

were shown to occur in the head kidney of several fish species (Zuasti and Ferrer 1988; Esteban, Meseguer et al. 1989; Meseguer, Esteban et al. 1990; Abdel-Aziz, Abdu et al. 2009). The head kidney also plays an important role in blood filtering by clearing debris and retaining bacteria. Macrophage aggregates, known as melanomacrophage centres (MMC), occur frequently throughout the head kidney and are thought to play a role in immunological memory by retaining antigens (reviewed in (Press and Evensen 1999)).

The thymus of fish is a paired organ located in the periphery of the upper region of the gill chamber, embedded in connective tissue and covered by the gill chamber epithelium. It is mainly composed by epithelial cells and thymocytes which, as in mammals, mature in the thymus (reviewed in (Press and Evensen 1999)). Hematopoietic processes such as erythropoiesis and granulopoiesis were observed in thymus (Aviles-Trigueros and Quesada 1995) and it was reported, for salmonid species, that the erythropoietic activity of the thymus was subjected to seasonal variations (Alvarez, Flano et al. 1994).

The spleen has different cell populations that include lymphocytes, macrophages and melanomacrophage accumulations. Erythrocyte destruction occurs at the spleen's melanomacrophage centres. Another major function of the spleen is to filter blood, retaining substances and immune complexes that are subsequently taken up by the spleen macrophages (reviewed in (Press and Evensen 1999)).

The thymus, head kidney and spleen are considered to be the major lymphoid organs of fish (reviewed in (Zapata, Diez et al. 2006)).

The mucosa-associated lymphoid tissue of the gut, skin and gills has a major role in the immune response since it is usually the first active immune organ to deal with invading pathogens. Although fish do not have organized structures like the mammalian Peyer's patches, the gut is populated with macrophages, lymphocytes, mast cells, granulocytes and plasma cells. The skin and gills are populated by plasma cells and other leukocytes (reviewed in (Press and Evensen 1999)).

Macrophage development (mediated by CSF-1/CSF1R signaling)

Macrophages, which are found throughout the body, are terminally differentiated cells of the mononuclear phagocyte system that also includes monoblasts, promonocytes, monocytes, dendritic cells and osteoclasts (reviewed in (Wiktor-Jedrzejczak and Gordon 1996)). All mononuclear phagocytic cells are known to express the hematopoietic growth factor receptor CSF1R (colony stimulating factor 1 receptor or macrophage colony stimulating factor

receptor) and homeostasis of this cell lineage is regulated by its high affinity ligand, the macrophage colony-stimulating factor CSF-1 (also known as M-CSF; reviewed in (Auffray, Sieweke et al. 2009)).

CSF-1 is a homodimeric glycoprotein that promotes the survival, proliferation and differentiation of cells of the mononuclear phagocytic lineage (reviewed in (Pixley and Stanley 2004) and (Sherr, Roussel et al. 1988)). This hematopoietic growth factor is synthesized mainly by fibroblast, endothelial cells and by monocytes and macrophages themselves (reviewed in (Sherr 1990)).

The first step in the differentiation of mononuclear phagocytes is not solely dependent on CSF-1 signaling. This growth factor acts together with the KIT ligand and interleukin-3 (IL-3) to drive the differentiation of the multipotent progenitor cells to mononuclear phagocyte progenitor cells. The subsequent differentiation processes of multipotent progenitor cells into monocytes and macrophages can be driven solely by CSF-1 (reviewed in (Pixley and Stanley 2004))

The proliferative rate induced by CSF-1 was shown to be different for diverse cells of the mononuclear phagocytic system. Murine bone derived macrophages exhibited a higher proliferation response to CSF-1 than that of peritoneal exsudate macrophages, degrading CSF-1 12 times more slowly than the peritoneal ones (Guilbert and Stanley 1986).

Four different CSF-1 proteins, all with CSF1R mediated stimulatory effects, can be produced from the five different mRNA transcripts of CSF-1 that were reported (reviewed in (Douglass, Driggers et al. 2008)). Alternative splicing of the RNA transcripts and proteolytic processing of the resulting proteins yields two rapidly secreted protein isoforms and two membrane bound isoforms that appear in the extracellular region and are more slowly released. In fact, the secreted CSF-1 pool represents 95-99% of the total CSF-1 protein that is produced and only 1-5% of CSF-1 remains bound to the membrane and may be further cleaved and released (reviewed in (Douglass, Driggers et al. 2008)).

The two secreted isoforms, circulate in the blood or associate with the extracellular matrix, and perform distant endocrine processes such as recruitment of macrophages and monocytes to sites of infection by virtue of the chemotactic effect CSF-1 have on macrophages (Webb, Pollard et al. 1996). The circulating M-CSF in healthy individuals was measured to be around 4.9 ng mL^{-1} (Gilbert, Praloran et al. 1989) or 540 U mL^{-1} (Hanamura, Motoyoshi et al. 1988).

The membrane bound isoforms perform mainly paracrine and, in some cases, autocrine functions (reviewed in (Pixley and Stanley 2004)). It was suggested that these membrane bound CSF-1 isoforms can undergo an active ectodomain shedding process, a mechanisms by which cells could individually control and enhance the chemotactic and chemokinetic effects on other cells that express CSF1R receptors (reviewed in (Douglass, Driggers et al. 2008)). The ectodomain shedding of the membrane bound CSF-1 is performed by TNF-α converting enzyme (TACE) (Horiuchi, Miyamoto et al. 2007).

CSF-1 is thought to be important for priming macrophages to perform different functions, namely to increase its phagocytic ability and to increase secretion of other cytokines, namely TNF-α and IL-1 (reviewed in (Sherr, Roussel et al. 1988) and (Sherr 1990)).

CSF-1 stimulates cell adhesion and phagocytic activity by increasing the expression of the leukocyte integrin α subunit (CD11b) and the Fcγ receptor III (CD16). CSF-1 alone also stimulates the inflammatory response by directly increasing the production of the cytokines IL-6, TNF-α and, together with LPS, IL-18; TNF-α further stimulates the production of interferon-gamma (IFNγ; which has an important role in macrophage activation) and IL-18 further stimulates the production of IL-1β (a pro-inflammatory cytokine). The production of IL-8, a chemokine that functions as chemoattractant of neutrophils, basophils and T-cells, is also stimulated by CSF-1. Antigen presentation is stimulated by increasing the expression of HLA-I and HLA-II (reviewed in (Chitu and Stanley 2006)).

CSF-1 is thought to stimulate mainly the production of tissue resident macrophages involved in scavenger functions and, at less extent, inflammatory macrophages (Cecchini, Dominguez et al. 1994). The CSF-1 stimulus induces major morphological changes in macrophages by remodeling the actin cytoskeleton, promoting cell adhesion, spreading, polarization, motility and, ultimately, phagocytosis (reviewed in (Pixley and Stanley 2004)). There are evidences that monocytes are capable of controlling their own survival, growth and differentiation through regulation of CSF-1 and CSF1R production.

It was suggested that some effects mediated by CSF1R were produced by another ligand (Dai, Ryan et al. 2002) and, in fact, a new CSF1R ligand whose action seems to be independent of CSF-1, the interleukin-34 (IL-34), was recently discovered in a high throughput screening of ligand-receptor interactions (Lin, Lee et al. 2008). IL-34 does not have sequence homology with CSF-1 and it binds to CSF1R in a different site (Chihara, Suzu et al. 2010). Its biological effects seem to be similar to those of CSF-1 but it is significantly less potent and has a different expression pattern (Wei, Nandi et al. 2010).

It was observed that the CSF-1 induced proliferation of murine macrophages was potentiated by TNF-α. TNF-α alone has no effect in the proliferation of murine macrophages, but together with CSF-1 it increases more than 2-fold the ability of CSF-1 to stimulate cell proliferation (Branch, Turner et al. 1989). This synergistic effect is probably due to an increase in CSF-1 secretion upon TNF-α stimuli, an effect previously observed in human monocytes (Oster, Lindemann et al. 1987).

CSF-1 receptor (CSF1R)

The majority of the biological effects of all CSF-1 isoforms are mediated through its interaction with the CSF-1 receptor (CSF1R), a class III receptor tyrosine kinase (RTK), member of the CSF1/PDGF (Fig. 1) receptor family of tyrosine kinases.

Each monomer that composes the CSF1R homodimeric structure is single-pass type I transmembrane protein. The size of the amino acid chain is slightly variable among different species. The human protein is 972 amino acids long, comprises an extracellular sequence of 512 amino acids that includes a predicted signal peptide of 19 amino acids, a hydrophobic transmembrane region of 25 amino acids and an intracellular sequence of 435 amino acids that contains the kinase domain. The extracellular domain is heavily glycosylated with eleven predicted N-linked oligosaccharide chains that represent 45 kDa in the total 150 kDa of the human mature CSF1R (reviewed in (Sherr 1990)).

The extracellular segment of 512 amino acids is structurally composed of five immunoglobulin-like (Ig-like) C2-set domains (Fig. 2), four of them stabilized by disulfide bonds (reviewed in (Sherr 1990)).

Though the modular arrangement of the extracellular sequences and the overall structure are somewhat conserved in all class III RTKs, the structures that mediate ligand binding of the different class III RTKs are not the same and the dimerization and activation process is slightly dissimilar.

The first three N-terminal Ig-like domains of CSF1R (D1, D2 and D3) were initially identified as the CSF-1 binding domains (Wang, Myles et al. 1993). More recently, structural studies revealed that, contrary to the binding mechanisms of the closely related KIT receptor, the first N-terminal Ig-like domain (D1) of CSF1R does not bind to CSF-1 (Chen, Liu et al. 2008). Nevertheless, it was observed that the module formed by the D2 Ig-like domain, the main responsible for the receptor-ligand interactions, and the D1 N-terminal Ig-like domain is structurally conserved among class III RTKs (Chen, Liu et al. 2008). The two juxtamembrane

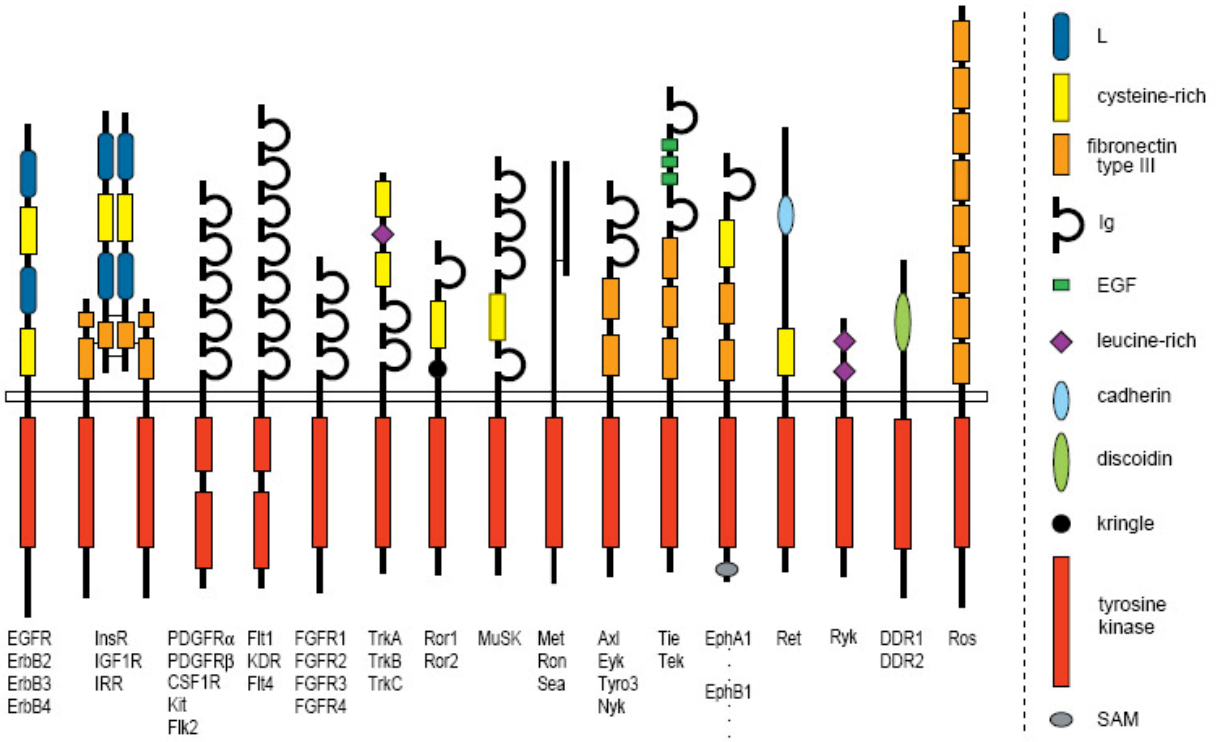


Figure 1. Diversity of receptor tyrosine kinases (RTKs). The modular structure of the extracellular domains (on top) is particularly evident in the PDGFR, VEGFR (Flt) and FGFR families. The members of the CSF1R and VEGFR share a typical kinase insert (intracellular portion on the bottom) that splits the catalytic kinase domain in two (adapted from (Hubbard and Till 2000)).

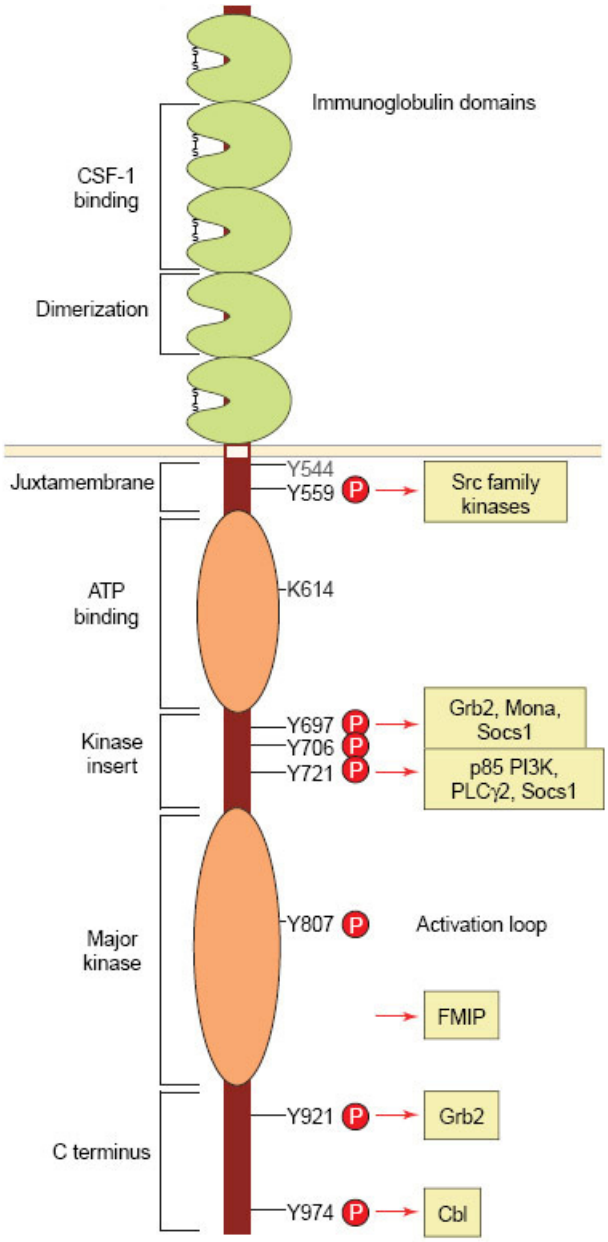
Ig-like domains (D4 and D5) are thought to be important for the efficient dimerization and activation of CSF1R upon CSF-1 binding (Wang, Myles et al. 1993; Chen, Liu et al. 2008).

The intracellular kinase domain is responsible for the enzymatic activity of the receptor and thus, for triggering the intracellular signaling cascades. The kinase domain is the receptor portion that is more widely conserved among homologues and orthologues (homologues of different species); despite this fact, an unique (and not so conserved) insert of 72 amino acids is present in the middle of the kinase domain. This kinase split is though to be important for the specific substrate recognition (reviewed in (Sherr 1990)).

In the absence of ligand, the CSF1R cell surface expression is maintained at approximately $5x10^4$ sites (murine macrophages). In the presence of ligand, CSF1R is rapidly internalized and degraded until almost complete depletion is reached in about 15 min. Despite of the fast recruitment of the intracellular CSF1R pool of approximately 10^5 molecules and *de novo* synthesis, cell surface expression of CSF1R remains low in the presence of ligand (Guilbert and Stanley 1986).

CSF-1/CSF1R signal transduction

CSF1R dimerization and subsequent activations is initiated by CSF-1 binding. A CSF-1 homodimer binds a single CSF1R molecule and induces conformational changes in the complex that allow the completion of the dimerized complex by recruitment of the second CSF1R molecule (Chen, Liu et al. 2008). The successful dimerization and subsequent internalization of the ligand-receptor complex is 6 times more frequent than ligand dissociation (Guilbert and Stanley 1986).



Successful ligand binding and complex noncovalent dimerization induces conformational changes that activate the intracellular kinase domain and various intracellular tyrosines residues become phosphorylated presumably by a transphosphorylation process. The phosphorylation is followed by the covalent stabilization of the CSF1R dimer by the formation of disulfide bridges (reviewed in (Pixley and Stanley 2004)).

The various phosphorylated tyrosines function as docking sites for different transducer and adaptor proteins, namely Src family kinases (SFKs), the monocyte adaptor (Mona), suppressor of cytokine signaling-1 (Socs1), phospolipase C-γ

Figure 2. CSF1R structural and functional key features. The five Ig-like domains on the extracellular portion of the receptor are depicted in the top and the intracellular kinase domain is at the bottom. Key tyrosine residues known to undergo phosphorylation upon receptor activation are highlighted in the intracellular domain and the docking proteins are indicated for each residues (adapted from (Pixley and Stanley 2004)).

(PLC-γ), growth factor receptor-bound protein 2 (Grb2), the regulatory p85 subunit of phosphatidylinositol 3-kinase (p85 PI3K), FMS-interacting protein (FMIP) and the ubiquitin-protein ligase Casitas B lineage (Cbl); these proteins are known to participate in several cellular processes such as the control of cell proliferation and the regulation/degradation of the receptor itself (reviewed in (Pixley and Stanley 2004)).

Eventually, the receptor becomes multiubiquitinated and is removed from cell surface via clathrin-dependent endocytosis. CSF-1 binding, dimerization, transphosphorylation, ubiquitation and clearance of the CSF-1:CSF1R complex occurs within 12–15 min (reviewed in (Douglass, Driggers et al. 2008)). A small portion (10-15%) of the internalized CSF-1 is rapidly degraded and the remaining undergo a slow ($t_{1/2}$ > 3.5 h) degradation process (Guilbert and Stanley 1986).

CSF1R associated pathologies

Several studies associate CSF1R overexpression with cancer, either as a result of gene copy, point mutations or translocations. Overexpression of CSF1R, or its ligand, and thus excessive signaling, has been reported for endometrial, ovarian, cervical, breast, pancreatic, colorectal and prostate cancer, Hodgkins lymphoma and leukemia (reviewed in (Douglass, Driggers et al. 2008)). This CSF-1/CSF1R excessive signaling, that stimulates cell proliferation, is performed in the tumor microenvironment and it was suggested that it may also be the result of an autocrine activation mechanism by the membrane bound CSF-1 isoform (Suzuki, Sekiguchi et al. 1996).

Abnormally elevated levels of circulating CSF-1 are also associated with several chronic inflammatory autoimmune disorders such as lupus and arthritis (reviewed in (Chitu and Stanley 2006)).

The *csf1r* gene

The human *csf1r* gene (also known as *c-fms*) is located on the long arm of chromosome 5, juxtaposed with its paralog gene *pdgfrβ* (platelet derived growth factor receptor beta). The first exon (noncoding) of the *csf1r* gene, located less than 0.5 kb downstream of the *pdgfrβ* polyadenilation signal, is separated from the remaining *csf1r* coding exons by a 26 kb intron (Roberts, Look et al. 1988). These two genes are thought to have derived from the same ancestral that was duplicated during a whole genome duplication event (Fig. 3), and although some divergence occurred, a high level of homology remains mainly in its intracellular kinase

domain (Roberts, Look et al. 1988). The human *csf1r* coding sequence consists of 21 small exons and heterogeneously sized introns, ranging from 6.3 kb to less than 0.1 kb in complexity.

Evidence suggests that the *csf1r* gene transcription is initiated at two independent promoters that function in a tissue-specific manner. Transcription initiation in a promoter immediately upstream of exon 2 yields a shorter transcript that is expressed in the mononuclear phagocytic cells. A longer transcript, specific of placental cells, is produced when transcription is initiated upstream of the first exon (Visvader and Verma 1989).

Fish CSF-1/CSF1R

The lineage that gave rise to teleost fish is thought to have undergone a third genome duplication (3R), also known as the fish-specific genome duplication (FSGD). In fact, additional copies of *csf1r* (Williams, Brenner et al. 2002) and *csf-1* (Wang, Hanington et al. 2008) genes were identified for some fish species.

The tandemly linked *csf1r* and *pdgfrβ* genes of *Takifugu rubripes* were duplicated in the FSGD (Fig. 3) (How, Venkatesh et al. 1996; Williams, Brenner et al. 2002; Braasch, Salzburger et al. 2006). It was suggested that since the duplication event, and as a result of asymmetrical co-evolution, the *pdgfrβb-csf1rb* locus genes acquired novel functions (Braasch, Salzburger et al. 2006). This hypothesis still lacks supporting evidence but studies performed in zebrafish concerning the more distant *csf1r* paralogs, the *kita* and *kitb* receptors and their respective ligands, *kitla* and *kitlb*, illustrate two typical outcomes of the evolution process that occur after a genome duplication event. The *kita* and *kitla* coevolved in a way that preserved the receptor-ligand bond; on the other hand, the *kitb* and *kitlb* receptor-ligand relationship was observed to be disrupted (Hultman, Bahary et al. 2007).

Gene duplication and, ultimately, whole genome duplications are thought to be major evolutionary mechanisms (Ohno 1970) and in fact, there is evidence that genome duplications preceded major species diversification, both in animals and plants (reviewed in (De Bodt, Maere et al. 2005)). It was suggested that after duplication of a gene, one copy keeps the original function relieving the other copy from evolutionary constrictions, usually resulting is loss of function (nonfunctionalization) or, less frequently, giving rise to a new function (neofunctionalization) (Ohno 1970). The redundancy of the duplicated gene protein products relieves selective pressure and enables the repeated accumulation of missense substitutions at high rate shortly after duplication, the frequency of these missense

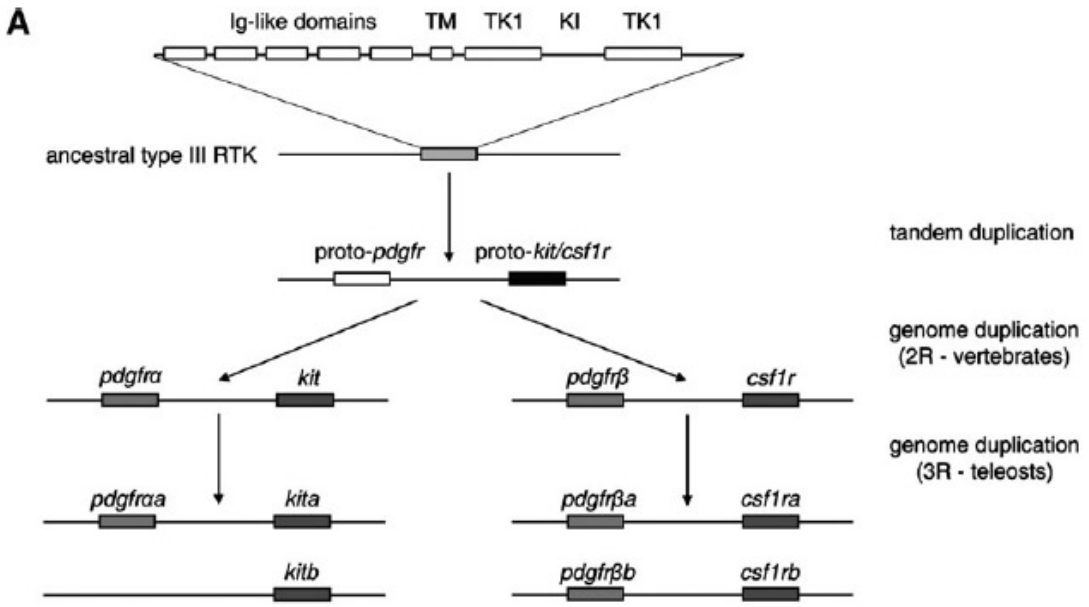


Figure 3. Receptor tyrosine kinases duplication. The third genome duplication, specific of teleosts, gave rise to two CSF1R forms (adapted from (Braasch, Salzburger et al. 2006))

substitutions diminishes as the paralog proteins become increasingly dissimilar; when the paralog proteins reach a divergence degree that eliminates functional redundancy, the pressure to preserve function greatly increases in one paralog, hence tightly restricting the accumulation of missense substitutions (reviewed in (Sidow 1996)). This protein divergence may also be attained by different mechanisms that have more discrete results, such as alternative splicing or transcriptional regulation. Expression patterns of redundant proteins can also shift gradually and result in tissue/structure specification of the proteins.

Neofunctionalization can occur in a discrete and complete form in only one of the paralogs, or it can occur in both paralogs simultaneously resulting in a sub-specialization of each paralog. This model, known as subfunctionalization, is thought to be more frequent than discrete neofunctionalization (Force, Lynch et al. 1999).

When multiple gene copies are beneficial for the organism, like in the case of plant pesticide resistance genes whose increase in protein expression can enhance pesticide resistance, the pressure to keep the multiple gene copies is huge and no differential gene evolution occurs.

CSF-1 genes were identified and characterized for several fish species, namely rainbow trout, zebrafish and goldfish (Hanington, Wang et al. 2007; Wang, Hanington et al. 2008). As expected, and in accordance with what was observed for other related genes, two copies of

the CSF-1 gene were identified for rainbow trout and zebrafish (Wang, Hanington et al. 2008). The protein products of these fish genes have low identity with the known mammalian counterparts but conserve cysteine residues required for dimerization and histidine residues required for interaction with CSF1R (Wang, Hanington et al. 2008); furthermore, functional studies performed with goldfish revealed that this growth factor, as in mammals, mediates macrophage development (Hanington, Hitchen et al. 2009) and the pro-inflammatory response of macrophages (Grayfer, Hanington et al. 2009). Treatment of macrophages with a recombinant form of CSF-1 increased the expression of pro-inflammatory molecules such as TNF-α (isoform 1 and 2), IL-1β (isoform 1 and 2), IL-10, IFNγ and IL12 (p35 and p40) (Hanington, Hitchen et al. 2009). The respiratory burst and the phagocytic activity, antimicrobial key features of macrophages, and the chemotactic response of macrophages, essential for macrophage recruitment and migration, were also enhanced by treatment with CSF-1; furthermore, these effects were abrogated in the presence of a recombinant soluble form of the CSF-1 receptor (Hanington, Hitchen et al. 2009).

The CSF1R genes were identified and characterized for several fish species, namely goldfish (Barreda, Hanington et al. 2005), pufferfish (Williams, Brenner et al. 2002), rainbow trout (Honda, Nishizawa et al. 2005), zebrafish (Parichy and Turner 2003) and gilthead sea bream (Roca, Sepulcre et al. 2006). The two copies of the CSF1R gene were identified for pufferfish (Williams, Brenner et al. 2002). The protein products of the genes share relatively high identity and similarity with the known mammalian counterparts, namely in the predicted CSF-1 binding domains and intracellular kinase domain (Barreda, Hanington et al. 2005; Roca, Sepulcre et al. 2006).

As in mammals, CSF1R is considered to be a specific marker of macrophages in fish and several biological properties of CSF-1 were also shown to be mediated by CSF1R (Herbomel, Thisse et al. 2001; Roca, Sepulcre et al. 2006; Mulero, Pilar Sepulcre et al. 2008; Hanington, Hitchen et al. 2009). Studies performed in zebrafish revealed that CSF1R also plays an important role in the pigment pattern development of fish, being essential for the maintenance of xanthophores and for the organization of melanophores (Parichy and Turner 2003).

It was suggested that, as for many other growth factors, production of a soluble CSF-1 receptor form with ligand binding affinity might be an alternative regulatory mechanism of CSF-1 activity (Barreda, Hanington et al. 2004). In fact, short transcript that encoded a putative soluble form of CSF1R was identified in goldfish self-renewing macrophages. The production of this transcript results from alternative splicing of a full-length mRNA, a

premature stop codon is inserted and the consequent premature termination of the transcript results in a shorter transcript that encode a putative soluble form of the CSF-1 receptor that comprises the two N-terminal Ig-like domains, that have some ligand binding ability. This transcript was found associated with a senescent phase of macrophages, which further indicates that it may have a regulatory role by decreasing the free CSF-1 and hence reducing CSF1R signaling and finally inhibiting cell proliferation (Barreda, Hanington et al. 2005). A recombinant protein corresponding to this predicted protein was shown to inhibit macrophage growth at very low concentrations and in a dose-dependent manner (Barreda, Hanington et al. 2005).

Fish macrophages are more prone to regulate their own growth than mammalian macrophages. Indeed, several macrophage self renewing cultures have been established *in vitro* from head kidney precursors without the aid of exogenous growth factor or other cell types (reviewed in (Barreda, Hanington et al. 2005)). The progenitor cells isolated from the head kidney are able to, by their own, give rise to different cell types that maintain themselves. Fish macrophages and head kidney leukocytes are able to produce growth factors that induce proliferation and differentiation of macrophages from hematopoietic progenitors (Neumann, Barreda et al. 1998).

Main purpose of this work

Macrophages are key players of the fish immune defenses. Fish photobacteriosis, is characterized by the massive destruction of macrophages and neutrophils, a process triggered by *Phdp*'s exotoxin AIP56. Taking into consideration that AIP56's action is highly specific for macrophages and neutrophils, specific markers for these cells would be of valuable assistance in the research of this infectious disease. Molecules such as CD14, CD68, F4/80 or CD11b have been widely used to label mammalian macrophages but these molecules are, at the moment, poorly characterized in fish. CSF1R has already been used to specifically label fish monocytes and macrophages (Roca, Sepulcre et al. 2006; Mulero, Pilar Sepulcre et al. 2008) and has been characterized for several fish species (Williams, Brenner et al. 2002; Parichy and Turner 2003; Barreda, Hanington et al. 2005; Honda, Nishizawa et al. 2005; Roca, Sepulcre et al. 2006). This work aims to characterize the sea bass CSF1R proteins and to develop antibodies that are specific to these molecules.

Materials and methods

Fish

European sea bass (*Dicentrarchus labrax* L.), purchased from a commercial fish farm, were kept in recirculating aerated seawater (20-25 $^{0}/_{00}$) at 20 ºC and fed commercial pelleted food *ad libitum*. Water quality was maintained with mechanical and biological filtration, complemented with ozone disinfection.

Fish were anesthetized with ethylene glycol monophenyl ether (0.03% v/v; Merck) before collecting blood and euthanized thereafter with a lethal dose of glycol monophenyl ether (0.06% v/v) and complete exsanguination by cutting the ventral aorta. To collect organs and peritoneal macrophages, fish were euthanized as described above.

cDNA cloning and sequencing

Multiple alignments of *Sparus aurata*, *Takifugu rubripes*, *Oncorhynchus mykiss*, *Danio rerio* and *Homo sapiens* CSF1R amino acid sequences were performed with ClustalW (Larkin, Blackshields et al. 2007) using standard parameters and degenerate primers were designed based on the conserved regions.

The head kidney of a fish that was stimulated with 60 µg of *E. coli* 026:B6 LPS (intraperitoneal injection) was collected 1 hour after stimulation and used to extract total RNA according to the RNAqueous-4PCR Kit (Ambion) protocol. RNA was reverse transcribed using BioScript RNase H^- reverse transcriptase (Bioline) according to the manufacturer's protocol and using either the APv2 primer (Table 1).

The cDNA was amplified with the primers CSF1RFW3 and CSF1RRV2 (Table 1 and Fig. 4) in a PCR reaction adjusted to a final volume of 50 µL, containing 200 µM of each dNTP, 1.5 mM $MgCl_2$, 5 µL of 10x PCR buffer (Bioline), 0.8 µM of each primer, 2 µL of cDNA and 1.25 U of *Taq* polymerase (Biotaq DNA polymerase; Bioline), with the following conditions: initial denaturing at 94 ºC for 2 min; 30 cycles of amplification (denaturing at 94 ºC for 45 s; annealing at 52 ºC for 1 min; extension at 72 ºC for 1 min 30 s); final extension at 72 ºC for 5 min. PCR products around 750 bp were

purified using the QIAquick gel extraction kit (Qiagen) after gel electrophoresis, cloned into pGEM-T Easy vector (Promega) and sequenced (Eurofins MWG Operon). Two different sequences of 740 and 773 bp were obtained and, analysis of the translated protein

sequences by BLAST (Altschul, Madden et al. 1997), showed that the 773 bp clone was similar to CSF1R1/a and the 740 bp band was similar to CSF1R2/b.

Specific reverse primers were designed and used together with 5'-upstream degenerate primers to obtain new PCR products in the 5'direction of the previous ones. Head kidney Total RNA of the head kidney from a fish stimulated with *Photobacterium damselae* subsp. *piscicida* strain PP3 (do Vale, Marques et al. 2003) was used to synthesize cDNA using the

Table 1. List of primers used in this study with their respective sequences and applications.

Primer	Sequence[a]	Application
APV	5′-GGC CAC GCG TCG ACT AGT ACT TTT TTT TTT TTT TTT TTV-3′	cDNA synthesis and amplification
APV2	5′-GAC TCA GGA CTT CAG GAC TTA GTT TTT TTT TTT TTT TTT V-3′	cDNA synthesis and amplification
AUAP	5′-GGC CAC GCG TCG ACT AGT AC-3′	cDNA amplification
AUAP2	5′-GAC TCA GGA CTT CAG GAC TTA G-3′	cDNA amplification
CSF1RFW1	5′-TTC ACY GGR ACM TAY ARR TG-3′	cDNA amplification
CSF1RFW2	5′-CCC TGC CTG CTG ACC GAC CC-3′	cDNA amplification
CSF1RFW3	5′-TAC AAY GAG AAG TGG GAG TT-3′	cDNA amplification
CSF1RRV2	5′-CTC TCT GGA GCC ATC CAC TT-3′	cDNA amplification
DLCSF1CSF1R1FW4	5′-GGC ATT CTG GCC ATC CTC CT-3′	cDNA amplification
DLCSF1CSF1R1FW6	5′-ATG CAG TCC TAC CTC ACT CT-3′	cDNA amplification
DLCSF1CSF1R1RV1	5′-CCA GGA TCT TCC CTA GCT TC-3′	cDNA amplification
DLCSF1CSF1R1RV2	5′-CAC AGC CAC ACG CAT CGC AT-3′	cDNA amplification
DLCSF1CSF1R1RV3	5′-ACG CTG GTG CCG TTG TCC AT-3′	cDNA synthesis
DLCSF1CSF1R1RV4	5′-GCA GGG CAG CAG GTA GTC TT-3′	cDNA amplification
DLCSF1CSF1R1RV5	5′-GAT GTG CTG CTG GTC CAG AA-3′	cDNA amplification
DLCSF1CSF1R1RV9	5′-ACA AAA CTG GTA GTT GTT GGT-3′	cDNA amplification
DLCSF1CSF1R2FW1	5′-CCA ACA TTG ACC AGA AAG GAG A-3′	cDNA amplification
DLCSF1CSF1R2FW4	5′-GAA GTG ACA GGA GAG GCT TTC-3′	cDNA amplification
DLCSF1CSF1R2FW5	5′-CCG GCC AGC ACG CCA GAC-3′	cDNA amplification
DLCSF1CSF1R2RV2	5′-CTT GAC AGC AAC TCG TGT GA-3′	cDNA amplification
DLCSF1CSF1R2RV3	5′-CGT CCT TCC AGC CTG AAC AA-3′	cDNA amplification
DLCSF1CSF1R2RV4	5′-GCC AGG TCT GGA TCT GAA CT-3′	cDNA synthesis
DLCSF1CSF1R2RV5	5′-GGG GTC AAT GGT CAC AAT CA-3′	cDNA amplification
DLCSF1CSF1R2RV6	5′-TTT TTT TTC CGT TCG TCT GAA T-3′	cDNA amplification
DLCSF1CSF1R1FW1NdeI	5′-GGA ATT CCA TAT GGC GGA ATG GAG GCG TCC GGT AAT CA-3′	recombinant protein
DLCSF1CSF1R1RV1XhoI	5′-CCG CTC GAG CTT GTC AGA AAC CTC CAT GGC-3′	recombinant protein
DLCSF1CSF1R2FW1NdeI	5′-GGA ATT CCA TAT GAA AGA TCC TCC TGG TCC TCC ATC G-3′	recombinant protein
DLCSF1CSF1R2RV1XhoI	5′-CCG CTC GAG CGT CTG AGG TTC TTC TGC AGA ACG-3′	recombinant protein
DLCSF1CSF1R2FW1EcoRI	5′-GGA ATT CAT GAA AGA TCC TCC TGG TCC TCC ATC G-3′	recombinant protein

[a] M = C/A; R = G/A; Y = T/C; V = G/C/A

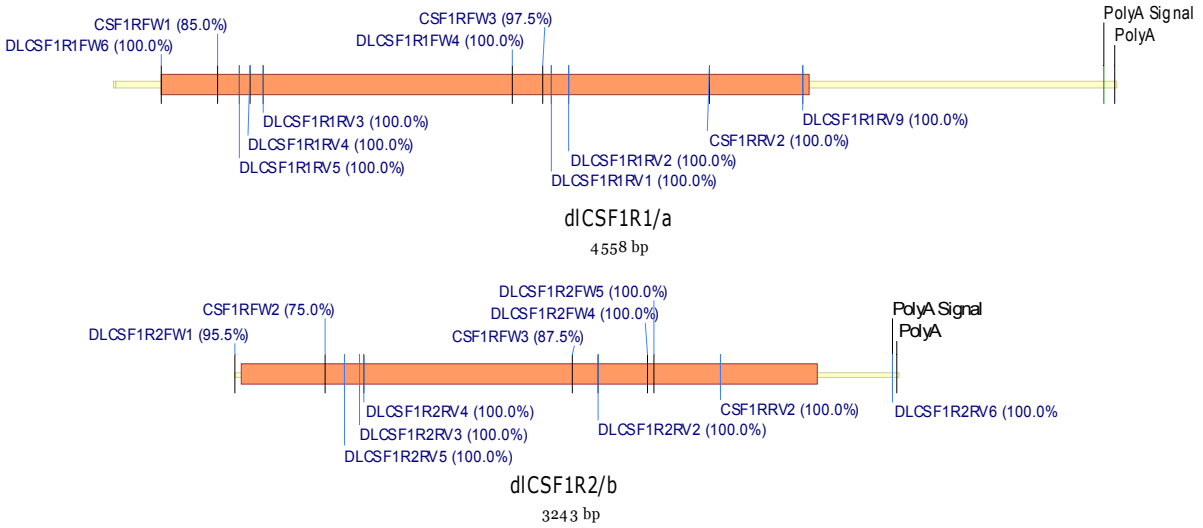


Figure 4. CSF1R1/a and CSF1R2/b sequences. Schematic representation of the sea bass CSF1R1/a and CSF1R2/b consensus cDNA sequences with the relative position of the primers (sequence identity in brackets) used in this study is depicted for each cDNA; the non-specific primers APv, APv2, AUAP and AUAP2 were not included in this scheme (see Table 1). The coding sequences of the cDNAs are shown in orange.

APv2 primer (Table 1). For the CSF1R1/a molecule, 2 µL of a 10-fold dilution of cDNA was used as a template for a first PCR reaction performed using the same conditions as described above, except for annealing temperature (55 ºC) and extension time (1 min), and primers CSF1RFW1 (0.8 µM) and DLCSF1R1RV2 (0.4 µM) primers (Table 1 and Fig. 4). A 50-fold dilution of the previous PCR was used as template for a semi-nested PCR reaction with the CSF1RFW1 (0.8 µM) and DLCSF1R1RV1 (0.4 µM) primers, using the same conditions as in the previous PCR. The obtained band of approximately 1500 bp was purified, cloned and sequenced as described before. For the CSF1R2/b molecule, a first PCR reaction was performed as for CSF1R1/a, but with the primers CSF1RFW2 (0.8 µM) and DLCSF1R2RV2 (0.4 µM) (Table 1 and Fig. 4). A 50-fold dilution of the previous PCR was used as template in a PCR reaction using the same primers and conditions. The obtained PCR product of approximately 1500 bp was purified, cloned and sequenced as described before.

Newly designed reverse primers were used to perform 5' RACE for both molecules.

For CSF1R1/a, cDNA was synthesized as decribed before but from the head kidney of a non-stimulated fish, using the specific reverse primer DLCSF1R1RV3 (Table 1 and Fig. 4); cDNA

was diluted to a final volume of 100 µL and purified using a Sephadex G-50 column. Purified cDNA was dATP tailed in a 25 µL volume reaction
with 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 200 µM dATP and 80 U of Terminal Transferase (Roche), the mixture was incubated at 37 ºC for 20 min and heat inactivated at 70 ºC for 10 min. The dATP-tailed cDNA (5 µL of a 10-fold dilution) was used as template in a PCR reaction using the same conditions as described above, except for the annealing temperature (61 ºC) and extension time (30 s), and using primers DLCSF1R1RV4 and APv primers (0.4 µM each). A nested PCR was performed using 1 µL of the previous PCR product as template under the same conditions, except for the annealing temperature (59 ºC), and primers DLCSF1R1RV5 and AUAP. A band of approximately 700 bp was purified, cloned and sequenced as described before.
For the CSF1R2/b 5' RACE, the same strategy applied for CSF1R1/a 5' RACE was followed, but using the specific reverse primer DLCSF1R2RV4 (Table 1 and Fig. 4) for cDNA synthesis. The dATP-tailed cDNA (2.5 µL) was used as template for a first PCR reaction under the same conditions as for CSF1R1/a 5' RACE but adjusted to a volume of 25 µL, annealing temperature at 59 ºC and 1 min extension time, using the primers DLCSF1R2RV3 and APv2 (0.4 µM each). A nested PCR was performed as the previous one, except for the annealing temperature (57 ºC), using 2.5 µL of a 50-fold dilution of the previous PCR as template and the primers DLCSF1R2RV5 and AUAP2 (0.4 µM each). The obtained band of approximately 600 bp was purified, cloned and sequenced as described before.

Several attempts to obtain full cDNA sequences of each CSF1R molecule were performed using primers in the 5'-UTR and an oligo-dT primer. However, as no PCR product could be obtained, the remaining of the sequence of each CSF1R molecule was obtained by 3' RACE as follows. Head kidney RNA (5 µg) of a non-stimulated fish was used to synthesize cDNA with the APv2 primer. This cDNA (1 µL) was used as template in the CSF1R1/a 3' RACE PCR reaction as first described, except for the annealing temperature (59 ºC) and extension time (3 min), and using primers DLCSF1R1FW4 and AUAP2 (0.4 µM each). A band of approximately 2500 bp was purified, cloned and sequenced as described before.
Spleen RNA (5 µg) of a non-stimulated fish was used to synthesize cDNA with the APv2 primer. This cDNA was used as template (0.25 µL) for the first CSF1R2/b 3' RACE PCR reaction using the same conditions as the previous one, but adjusted to a total reaction volume of 15 µL, annealing temperature at 57 ºC and extension time of 2 min, using primers

DLCSF1R2FW4 and AUAP2 (0.4 µM each). A semi-nested PCR was performed using 1 µL of the previous PCR product as template in a 50 µL reaction using the same conditions as before, except for the annealing temperature (63 ºC) and primers DLCSF1R2FW5 and AUAP2 (Table 1). A band of approximately 1200 bp was purified, cloned and sequenced as described before.

Finally, all obtained sequences were aligned and consensus sequences of CSF1R1/a and CSF1R2/b were assembled including the start and stop codons and the polyA signals. Based on the above consensus sequences, complete CDS sequences of CSF1R1/a and CSF1R2/b were obtained and cloned as follows.

The cDNA that was previously used for the CSF1R1/a 3' RACE was used as template in a first PCR amplification with the primers DLCSF1R1W6 and DLCSF1R1RV9 (0.4 µM each), using the same conditions as described before, except for the annealing temperature (54 ºC), extension time (2 min), total reaction volume (25 µL) and polymerase (an extra 0.08 U of *Pfu* DNA polymerase (Promega) was added). A 1000-fold dilution of the PCR product was used in a second PCR amplification with the same conditions except for the total reaction volume (50 µL). The resulting 2943 bp PCR product was purified, cloned and sequenced.

The CSF1R2/b CDS was amplified a single PCR reaction with the primers DLCSF1R2FW1 and DLCSF1R2RV6 using as template the cDNA that was previously used for the CSF1R1/a 3' RACE. Reaction conditions were the same as for amplification of CSF1R1/a except for the annealing temperature (starting at 60 ºC with 1 ºC decrements in the first 5 cycles and 55 ºC in the remaining 25 cycles), extension time (3 min 30 s), total volume (25 µL). The resulting 3234 bp PCR product was purified, cloned and sequenced.

Sequence analysis

Raw nucleotide sequences were visualized with Chromas 1.45. Contig sequences were assembled using the ContigExpress component of Vector NTI Advance10 (Invitrogen) and curated manually (Lu and Moriyama 2004). Amino acid sequences were deduced using the ExPASy translate tool (www.expasy.org/tools/dna.html;(Gasteiger, Gattiker et al. 2003)). Nucleotide and deduced amino acid sequences were compared to the respective DNA and protein databases using the BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi; (Altschul, Madden et al. 1997)). Protein molecular weight was predicted using the ExPASy compute pI/Mw tool (www.expasy.org/tools/pi_tool.html; (Gasteiger, Gattiker et al. 2003)). Proteins domains were

predicted using the InterProScan tool (www.ebi.ac.uk/Tools/InterProScan/; (Mulder and Apweiler 2007)) and SMART (http://smart.embl-heidelberg.de/; (Schultz, Milpetz et al. 1998)). Putative *N*-glycosylation sites were predicted with the ScanProsite tool (www.expasy.org/tools/scanprosite/; (de Castro, Sigrist et al. 2006)). Transmembrane domains were predicted using the TMpred tool (www.ch.embnet.org/software/TMPRED_form.html) and curated manually based on alignments with characterized proteins. Signal peptide was predicted using SignalP 3.0 (Emanuelsson, Brunak et al. 2007). Multiple alignments were made with Clustal W (Larkin, Blackshields et al. 2007) using standard parameters. The similarity and identity percentages were calculated using MatGAT 2.02 (Campanella, Bitincka et al. 2003) with the BLOSUM64 scoring matrix and the values 16 for first gap and 4 for extended gap penalties. Neighbour-joining phylogenetic tree was constructed using MEGA 4 program (Tamura, Dudley et al. 2007) with p-distance and complete deletion of gaps. Exons were predicted by aligning the obtained cDNA sequences with the stickleback (*Gasterosteus aculeatus*) genomic CSF1R sequences found at the Ensembl Genome Browser (http://www.ensembl.org/), alignments were performed using the Spidey alignment tool (http://www.ncbi.nlm.nih.gov/spidey/; (Wheelan, Church et al. 2001)).

Protein 3D modelling

Modelling was performed using the SWISS-MODEL Workspace (http://swissmodel.expasy.org/workspace/; (Arnold, Bordoli et al. 2006)).

Structural templates were obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB; www.pdb.org; (Berman, Westbrook et al. 2000)).

The structure of the human Kit receptor extracellular domain (PDB ID: 2EW9 chain B (Yuzawa, Opatowsky et al. 2007)) was used as template to model the full-length extracellular domains of sea bass CSF1R1/a (from amino acid E^{19} to M^{517}; 26% sequence identity) and sea bass CSF1R2/b (from amino acid P^{22} to D^{476}; 20% sequence identity). The structure of the three N-terminal Ig-like domains (D1, D2 and D3) of the murine CSF1R protein (PDB ID: 3EJJ chain X (Chen, Liu et al. 2008)) were used as template to model the three N-terminal domains of sea bass CSF1R1/a (from amino acid R^{22} to V^{310}; 27% sequence identity) and CSF1R2/b (from amino acid P^{25} to V^{314}; 27% sequence identity). The structure of the autoinhibited human CSF1R kinase (PDB ID: 2OGV chain A (Walter, Lucet et al. 2007)) was

used as template to model intracellular juxtamembrane (JM), tyrosine kinase (TK) and kinase insert (KI) domains of sea bass CSF1R1/a (from amino acid K^{552} to Q^{933}: 62% sequence identity) and sea bass CSF1R2/b (from amino acid K^{514} to I^{868}; 60% sequence identity).

Given that there are significant structural differences between the D1-D2 segment of the Kit template (2E9W) and the D1-D2 segment of the CSF1R template (3EJJ) (Chen, Liu et al. 2008), the obtained sea bass homology models were used to create chimeric models comprising the D1, D2 and D3 domains obtained with the murine CSF1R template (3EJJ) and the D4 and D5 domains obtained with the human Kit template (2EW9); this was done by superposing the alpha carbon chain of the D3 domain of both homology models and subsequent deletion of the D3 domain corresponding to the homology models obtained with the Kit template (2E9W).

The obtained homology models were compared with the respective templates using the secondary-structure matching (SSM v2.36) service at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/ssm; (Krissinel and Henrick 2004)) and accordingly edited using the PyMOL Molecular Graphics System, Version 1.2r2, Schrödinger, LLC.

Pictures were rendered using the software CCP4mg 2.4.1 (Potterton, McNicholas et al. 2004).

Recombinant proteins

A fragment corresponding to the extracellular domain of CSF1R1/a (A^{18} to L^{523}) was amplified by PCR from a full CSF1R1/a coding clone (1 µL of a 1000-fold dilution of a standard miniprep). Forward (DLCSF1R1FW1NdeI) and reverse (DLCSF1R1RV1XhoI) primers (Table 1) containing the NdeI (includes a start codon) and XhoI restriction sites, respectively, were designed and used at a concentration of 0.4 µM each in a PCR reaction adjusted to a final volume of 50 µL, containing 200 µM of each dNTP, 1.5 mM $MgCl_2$, 10 µL of 5x GoTaq Flexi buffer (Promega), 1.25 U of GoTaq Flexi DNA polymerase (Promega) and 0.15 U of *Pfu* DNA polymerase (Promega). PCR conditions were: 94 ºC for 2 min; 30 cycles of amplification (94 ºC for 45 s; 65 ºC for 1 min; 72 ºC for 1 min 30 s); 72 ºC for 5 min. A PCR product of 1537 bp, that included terminal XhoI and NdeI restriction sites was purified with the QIAquick gel extraction kit and partially digested with NdeI and XhoI restriction enzymes (Fermentas). The 1521 bp semi-digested product was isolated by gel electrophoresis, purified as before and ligated to pET23a(+) expression vector previously digested with NdeI and XhoI. The resulting

plasmid, pET23DLCSF1Re, codes for the extracellular domain of CSF1R1/a with a C-terminal histidine tag.

The extracellular domain of CSF1R2/b (K^{19} to T^{488}) was amplified by PCR from a clone containing the CSF1R2/b cDNA's 5' sequence (1 µL of a 1000-fold dilution of a standard miniprep). Primers DLCSF1R2FW1NdeI and DLCSF1R2RV1XhoI primers were used at a concentration of 0.4 µM each in a PCR reaction as described for CSF1R1/a. A PCR product of 1432 bp, that included a terminal XhoI restriction site was purified and blunted using T4 DNA Polymerase (Promega), the polymerase was heat inactivated and DNA was digested with XhoI (Fermentas). The resulting 1426 bp product was purified and ligated in frame with a C-terminal His-tag into pET23a(+) that was previously digested with NdeI (blunted with T4 DNA Polymerase) and XhoI. The resulting plasmid, pET23DLCSF2Re, codes for the extracellular domain of CSF1R2/b with a C-terminal His-tag.

The same extracellular domain of CSF1R2/b was also cloned in pET28a(+). A forward primer (DLCSF1R2FW1EcoRI; Table 1) containing the EcoRI restriction site was designed and used with DLCSF1R2RV1XhoI (Table 1) in a PCR reaction similar to the one performed for the CSF1R1/a recombinant protein but with an annealing temperature of 63 ºC. A PCR product of 1429 bp, that included terminal EcoRI and XhoI restriction sites was purified with the QIAquick gel extraction kit and digested with EcoRI and XhoI restriction enzymes (Fermentas). The 1419 bp digested DNA was purified using the QIAquick gel extraction kit after gel electrophoresis and ligated in frame with the N- and C-terminal His-tags of pET28a(+) that was previously digested with EcoRI and XhoI. The resulting plasmid, pET28DLCSF2Re, codes for the extracellular domain of CSF1R2/b with N-terminal and C-terminal histidine tags.

After sequencing, clones were transformed into *Escherichia coli* BL21 cells (Star, Rosetta, DE3 and CodonPlus strains).

E. coli carrying the recombinant proteins encoding plasmids were used to inoculate 10 mL of LB media supplemented with the respective antibiotic, and the cultures incubated overnight at 37 ºC with shaking. Fresh media without antibiotic (10 mL) was inoculated with 100 µL of the saturated overnight culture and incubated at 37 ºC with shaking until the OD_{600} reached around 0.6. IPTG was added to a final concentration of 0.4 mM to induce recombinant protein expression and the cultures further incubated until a maximum of 16 hours at 17, 24 or 37 ºC. Samples of the cultures were taken at different time points during induction, pelleted at 1800

g for 6 min at 4 ºC, resuspended in wash buffer (50 mM phosphate buffer pH 8.0 containing 150 mM NaCl), pelleted again and frozen at -20 ºC.

To test the protein solubility, bacterial pellets were resuspended in wash buffer containing 1% (v/v) Triton X-100 and sonicated eight times for 15 s, allowing the samples to cool down on ice after each sonication. Lysates were centrifuged at 20000 *g* for 15 min at 4 ºC and the resulting supernatants (soluble fraction) and pellets (insoluble fraction) frozen at -20 ºC.

Production of CSF1R1/a and CSF1R2/b antibodies

Rabbit monospecific polyclonal antibodies were produced against peptides of the sea bass CSF1R1/a (SLQLKRMNAQEQGEY) and CSF1R2/b (FVLRRVRQEDRGRYSF) molecules, homologous to human counterparts known to contain immunogenic epitopes to which antibodies have been already produced and characterized (Lifespan LS-C5837), and predicted by aligning sea bass and human CSF1R amino acid sequences. Synthetic peptides, pre-immune sera, immune whole sera and affinity-purified IgG fractions were provided by Davids Biotechnologie.

Collection of peritoneal leukocytes

To recruit a greater number of macrophages to the peritoneal cavity, fish were stimulated by injecting intraperitoneally 100 µL of Freund's Incomplete Adjuvant (FIA; Sigma). At day 12-20 post-stimulus, fish were euthanized and, after exsanguination, the abdominal side of the fish was washed with ethanol. L-15 medium (Gibco; supplemented with 10% (v/v) fetal bovine serum (FBS), 1.2 mM NaCl, 1% (v/v) Pen-Strep and 20 $U.mL^{-1}$ heparin) or sea bass PBS (sbPBS; 1.8 mM NaH_2PO_4, 2.2 mM Na_2HPO_4, 184 mM NaCl, pH 7.2) was injected into the peritoneal cavity (5 mL per fish) and, after a brief massage of the abdominal side, the wash medium containing enriched macrophage peritoneal leukocytes was drained from the peritoneal cavity, cells were counted and kept on ice until processed.

Collection of peripheral blood leukocytes

Blood was collected from the caudal vein of anesthetized fish with a heparinised syringe. Blood was diluted in ice-cold sbPBS (1:1) and centrifuged for 15 min at 200 *g* (4 ºC) to pellet erythrocytes. The supernatant was centrifuged at 550 *g* for 10 min (4 ºC), the pellet was resuspended in sbPBS, carefully layered onto 3 mL of Lymphoprep (Axis-shield; 1.077 $^{+}/_{-}$ 0.001 mg mL^{-1} density) and centrifuged at 1100 *g* (4 ºC) for 30 min with minimal deceleration.

Cells at the interface were collected, washed twice with sbPBS by centrifuging for 10 min at 550 g (4 ºC), resuspended in sbPBS, counted and kept on ice until processed.

Collection of leukocytes from organs

Head kidney, spleen or thymus was collected from euthanized and exsanguinated fish.

Cell suspension was obtained by macerating the organ through a 100 µm nylon mesh and rinsing it with sbPBS. Cells were then washed twice with sbPBS by centrifuging for 10 min at 550 g (4 ºC) and finally resuspended in 2 mL of sbPBS. The cell suspension was then carefully layered onto 3 mL of Lymphoprep (Axis-shield, Oslo, Norway; 1.077 $^{+}/_{-}$ 0.001 mg mL^{-1} density) and centrifuged at 1100 *g* (4 ºC) for 30 min with minimal deceleration. Cells at the interface were collected, washed twice as described before, resuspended in sbPBS, counted and kept on ice until processed.

SDS-PAGE

Cells were resuspended or diluted in gel loading buffer (50 mM Tris-HCl, pH 8.8, 2% (w/v) SDS, 10% (v/v) glycerol, 0.017% bromophenol blue, 2 mM EDTA, 100 mM DTT), boiled for 5 min and run in discontinuous SDS-PAGE mini-gels (Mini-PROTEAN 2; Biorad) of variable acrylamide concentrations (8-12%) depending on the target protein size. Electrophoresis was buffered with 25 mM Tris, 250 mM glycine and 0.1% SDS and performed at 150 V. Proteins were either transferred to nitrocellulose membranes or stained directly in the gel with Coomassie Brilliant Blue G-250.

Western blot

Cell lysates were subjected to SDS-PAGE and proteins were transferred to 0.45 µm nitrocellulose membrane (Whatman) using a semi-dry electrophoretic system (Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell; Bio-rad); transfers were buffered with 25 mM Tris, 192 mM glycine and 20% (v/v) methanol in distilled H_2O (dH_2O) and performed for 1 hour with an electric intensity of 2 mA cm^{-2} with a maximum voltage value of 19V. After transferring, the membranes were rinsed with dH2O and stained with Ponceau S to confirm the transfer efficiency. Ponceau S was removed by washing with dH2O and membranes were blocked with 5% (w/v) non-fat powdered milk in TBS-T (20 mM Tris, pH 7.6, 137 mM NaCl, 0.1% (v/v) Tween-20) for 1 hour at room temperature (RT) with shaking or overnight at 4 ºC with shaking. After rinsing with TBS-T, membranes were incubated for 1 hour with the primary

antibody (anti-CSF1R1 or anti-CSF1R2) diluted in TBS-T (1:1000, 1:2000, 1:5000, 1:10000 and 1:20000) and, after washing (3 x 10 min) with TBS-T, were incubated (1 h at RT with shaking) in TBS-T with a 1:10000 dilution of the secondary antibody goat anti-rabbit Ig Horseradish Peroxidase (GAR-HRP; Amersham Pharmacia Biotech UK Limited). Membranes were washed again as before and rinsed with dH2O, the excess water was drawn and membranes were incubated with the detection reagent (SuperSignal West Pico Chemiluminescent Substrate; Thermo Scientific) according to the manufacturer's instructions.

Immunoprecipitation

Peritoneal macrophage lysates were used to perform immunoprecipitation using Roche's Immunoprecipitation Kit (protein G-agarose; Roche) according to the manufacturer's instructions.

Prior to the immunoprecipitation protocol, the proteins' availability in different buffers were assessed by western blot. Cells ($1x10^7$ cells per immunoprecipitation reaction) were resuspended in 1 mL of either the cell lysis buffer of the kit (50mM Tris-HCl pH 7.5; 150mM NaCl; 1% Nonidet P40; 0.5% sodium deoxycholate and protease inhibitors) or the RIPA buffer (lysis buffer of the kit supplemented with 0.1% SDS) and were lysed by incubating in a test tube rotator for 30 min at 4 ºC; denaturing/reducing lysis buffer (1% SDS; 5 mM EDTA and freshly added 1mM dithiothreitol) was also used to lyse cells ($1x10^7$ cells in 100 µL of buffer) by vortexing for 2/3 s, heating at 100 ºC for 5 min, diluting 10-fold in the cell lysis buffer of the kit and passing 10 times through a 26G needle and finally incubating for 5 min on ice. Lysates were then centrifuged at 12000 g for 10 min (4 ºC) and aliquots of the pellets and supernatants screened for the presence of the target proteins by western blot.

Immunoprecipitations were performed using the kit's buffer and the denaturing/reducing buffer lysates according to Roche's Immunoprecipitation Kit instructions. Briefly, 1 mL of cell lysate was pre-cleared by incubating in a test tube rotator for 3 hours at 4 ºC with 50 µL of protein G-agarose suspension after which the beads were pelleted and discarded. Affinity-purified rabbit polyclonal antibody (1-5 µg) was added to the supernatants and the mixtures incubated in a test tube rotator for 1 hour or overnight at 4 ºC. Then, 50 µL of protein G-agarose suspension was added and the mixtures further incubated overnight at 4 ºC with shaking. Beads were pelleted, washed repeatedly with decreasing ionic strength buffers, resuspended in gel loading buffer and boiled for 5 min. The beads were then pelleted and supernatants subjected to SDS-PAGE.

Flow cytometry

Peritoneal macrophages, peripheral blood leukocytes, head kidney, spleen and thymus leukocytes were collected as described before and maintained at 4 ºC during all steps of the procedure. Cells were placed into wells of round bottom 96-well microtitre plates ($5x10^5$ cells per well; Nunclon) and incubated for 45 min at 4 ºC with rabbit polyclonal antibodies anti-CSF1R1 or anti-CSF1R2 diluted in sbPBS (1:10, 1:25, 1:50) supplemented with 0.1% (w/v) sodium azide and 0.5% (w/v) BSA. After washing twice with sbPBS, by centrifuging at 250 *g* for 2 min, cells were incubated for 45 min at 4 ºC in the dark with the secondary antibody Alexa Fluor 488 goat anti-rabbit (Molecular Probes) diluted in sbPBS (1:2000) containing with 0.1% [w/v] sodium azide and 0.5% [w/v] BSA. Cells were washed twice as described above and resuspended in sbPBS. Propidium iodide (1 µg mL^{-1}) was added just before analysis to exclude dead cells. Analysis was performed in a FACSCalibur flow cytometer (BD Biosciences) and the results were edited using FlowJo software.

Immunofluorescence

Head kidney or peritoneal leukocyte suspensions obtained as described above were centrifuged were centrifuged at 1000 rpm for 5 min (Shandon Cytospin 2; Shandon) onto glass slides, fixed for 15 min in cold 4 % paraformaldehyde in sbPBS and stored at -20 ºC in PBS (1.8 mM NaH_2PO_4, 2.2 mM Na_2HPO_4, 150 mM NaCl, pH 7.2) containing 65 % glycerol until immuno-detection. For immuno-detection, slides were washed with PBS (3 x 5 min) and blocked for 30 min with 10% FBS and 0.1% Tween-20 in PBS. Blocking solution was aspirated and slides were incubated with the primary antibody rabbit anti-CSF1R1 or anti-CSF1R2 IgG (1:50 – 1:2500 dilutions in PBS) for 1 hour at RT or overnight at 4 ºC, washed with PBS-T (3 x 5 min; PBS containing 0.1% (v/v) Tween-20) and incubated with Alexa Fluor 488 goat anti-rabbit IgG (1:2000 dilution in PBS) for 1 hour at RT. Slides were then washed with PBS-T (3 x 10 min) and incubated with DAPI (10 µg mL^{-1} in PBS) for 15 min, mounted using Vectashield Fluorescent Mounting Media (Vector Labs) and sealed with nail polish. Preparations were observed under an Axiovert 200M microscope (Carl Zeiss).

Results and discussion

Cloning and characterization of sea bass CSF1R1/a

The 4535 bp CSF1R1/a full-length cDNA (Annex I, Fig. A.1) comprises a 209 bp 5'untranslated region (UTR), a 2946 bp open reading frame (ORF) encoding a predicted protein of 981 amino acids and a 3'UTR of 1380 bp, excluding the polyadenylation tail (poly(A) tail). The ATTAAA polyadenylation signal is located 27 bp upstream of the poly(A) tail.

The Interproscan, SMART and Scanprosite tools were used to characterize the domains and motifs of the predicted amino acid sequence (Fig. 5). A 17-amino acid signal peptide was predicted, meaning that the mature protein is 964 amino acids long and has a theoretical molecular mass of 109.5 kDa. The extracellular domain of 505 amino acids comprises 5 Ig-like domains, two of the constant-2 (C-2) subtype. The alignment with the human CSF1R (NCBI and UniProt accession number: P07333) revealed that the cysteine residues (C) responsible for the disulfide bridges that stabilize the extracellular Ig-like domain 1 (D1; C^{49} and C^{93}), domain 2 (D2; C^{140} and C^{189}), domain 3 (D3; C^{236} and C^{292}) and domain 5 (D5; C^{433} and C^{502}) are conserved; the predicted Ig-like domain 4 (D4), as the human and murine (NCBI and UniProt accession number: P09581) counterparts, lacks cysteine residues and the respective disulfide bridge. The extracellular portion of sea bass CSF1R1/a protein is predicted to be glycosylated at 9 asparagine residues (N), this extensive predicted glycosylation is in agreement to what is also predicted for the human and murine counterparts.

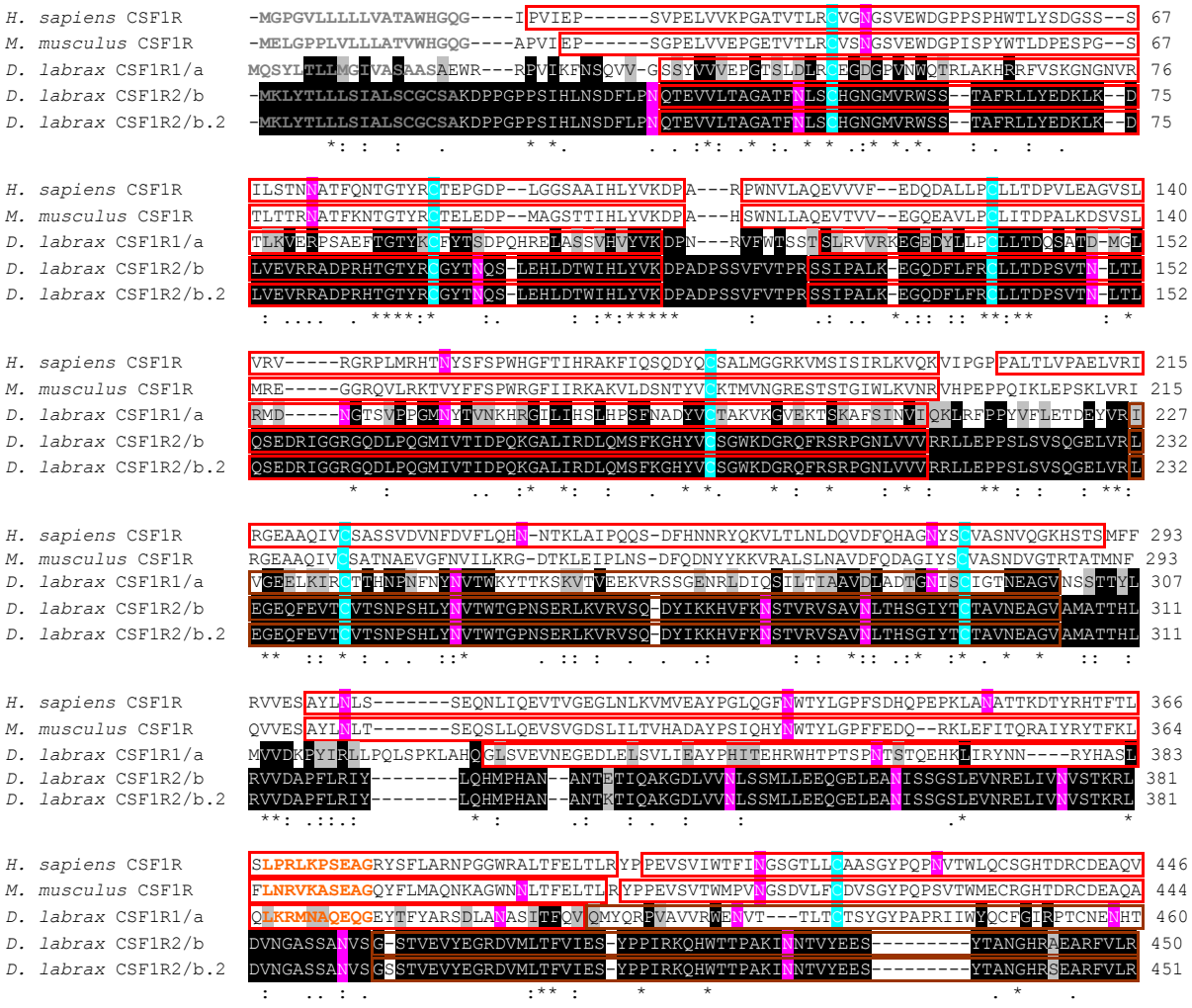
The intracellular portion of the the sea bass CSF1R1/a protein is 434 amino acids long and includes a juxtamembrane (JM) domain, the tyrosine kinase (TK) domain, a kinase insert (KI) domain and a C-terminal tail (Fig. 5).

The majority of the intracellular tyrosine residues that, in the mammalian counterparts, were reported to undergo phosphorylation and be functionally relevant (reviewed in (Pixley and Stanley 2004)) are conserved in the sea bass CSF1R1/a molecule, namely Y^{555}, Y^{570} (major mediator of proliferation (Takeshita, Faccio et al. 2007) and Src family kinases interaction putative site (Alonso, Koegl et al. 1995)), Y^{709} (Grb2, Mona and Socs1 interactions; (van der Geer and Hunter 1993; Bourette, Arnaud et al. 1998; Bourette, De Sepulveda et al. 2001)), Y^{733} (p85 PI3K, PLCγ2 and Socs1 interactions; (Reedijk, Liu et al. 1992; van der Geer and

Hunter 1993; Bourette, Myles et al. 1997; Bourette, De Sepulveda et al. 2001)), Y^{820} (required for mitogenesis; (Roussel, Shurtleff et al. 1990)), Y^{936} (Grb2 interaction; (Mancini, Niedenthal et al. 1997)) and Y^{978} (Cbl interaction and receptor degradation; (Mancini, Koch et al. 2002; Wilhelmsen, Burkhalter et al. 2002)).

Cloning and characterization of sea bass CSF1R2/b

Two CSF1R2/b clones were isolated, a 3229 bp cDNA (named CSF1R2/b; Annex I, Fig. A.2) and a 3097 bp cDNA (named CSF1R2/b.2; Annex I, Fig. A.3). The shorter cDNA seems to be an alternative spliced variant, whose most significant difference is the deletion of part of the sequence in the final part of the ORF, resulting in a smaller protein that has a much shorter C-terminal tail (35 amino acids shorter; Fig. 5), lacking a poly-glutamate stretch whose function is unknown.



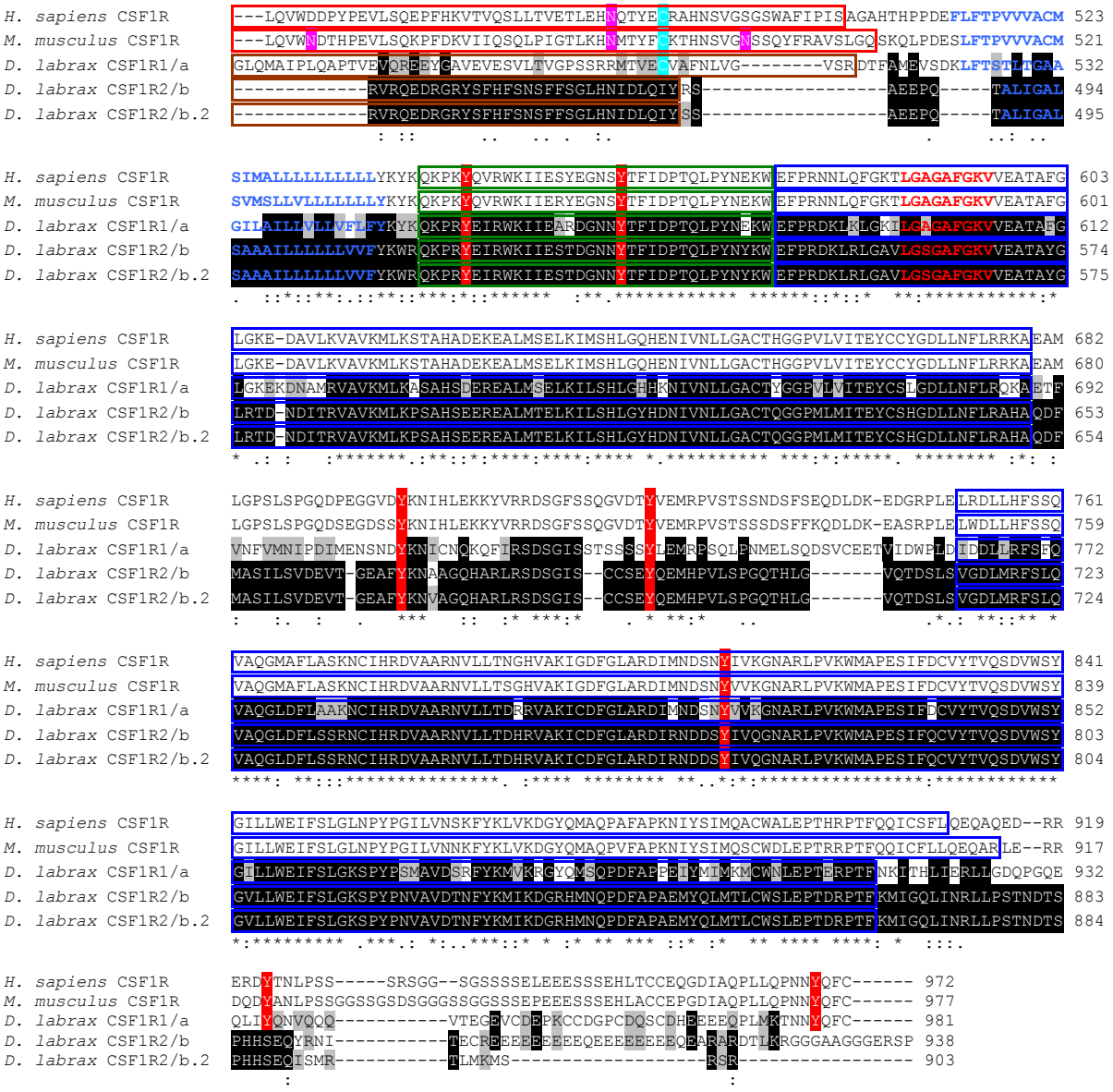


Figure 5. Multiple sequence alignment of the CSF1R amino acid sequences. Sea bass (*D. labrax*), human (*H. sapiens*; accession number P07333) and mouse (*M. musculus*; accession number P09581) multiple sequence alignment. The gray letters in bold correspond to the predicted signal peptide; the predicted glycosylated residues are shaded in pink; the structural and functional relevant cysteine (C) and tyrosine (Y) residues are shaded in turquoise and red respectively; the extracellular predicted Ig-like domains are depicted by red or brown (constant-2 subtype), the juxtamembrane (JM) domain is denoted by a green box, the kinase domain by a blue box (interrupted by the kinase insert (KI) domain); the transmembrane domains are denoted by blue letters in bold; the ATP binding site is denoted by red letters in bold; the dimerization motif is denoted by orange letters in bold. The numbers indicate the amino acid positions and the dashes indicate gaps introduced to optimize similarity between sequences. Asterisks (*) denote identical residues (black shading for the sea bass sequences) and the colon (:) and dot (.) denote chemical similarity between amino acids (gray shading for the sea bass sequences). The human and murine CSF1R proteins are annotated as described in the UniProt database.

The 3229 bp CSF1R2/b full-length cDNA comprises a 28 bp 5'UTR, a 2817 bp ORF encoding a predicted protein of 938 amino acids and a 3'UTR of 384 bp, excluding the poly(A) tail. The polyadenylation signal (AATAAA) is located13 bp upstream of the poly(A) tail.

Analysis of the predicted amino acidic sequence (Fig. 5) revealed a 18-aa signal peptide, resulting in a 920 amino acids long mature protein with a theoretical molecular mass of 103.5 kDa. The Interproscan and SMART analysis of the 470 amino acid extracellular domain predicted the existence of 4 Ig-like domains, one of them of the constant-2 (C-2) subtype. The alignment with the human CSF1R (P07333) revealed that the cysteine residues responsible for the disulfide bridges that stabilize the extracellular Ig-like domain 1 (D1; C^{52} and C^{92}), domain 2 (D2; C^{140} and C^{194}) and domain 3 (D3; C^{241} and C^{296}) are conserved but the predicted fourth Ig-like domain, the membrane proximal domain whose sequence and relative position is similar to the mammalian fifth Ig-like domain, lacks cysteine residues at all.

The extracellular portion of sea bass CSF1R2/b protein is predicted to be glycosylated at 12 asparagine residues (N).

The intracellular portion of the sea bass CSF1R2/b protein is 430 amino acids long (395 amino acids in the CSF1R2/b.2 protein) and comprises the juxtamembrane (JM) domain, the tyrosine kinase (TK) domain, a kinase insert (KI) domain and a C-terminal tail (Fig. 5).

As for CSF1R1/a, the majority of the CSF1R2/b intracellular tyrosine residues that, in the mammalian counterparts, were reported to undergo phosphorylation and be functionally relevant are conserved (reviewed in (Pixley and Stanley 2004)), namely Y^{517}, Y^{532} (major mediator of proliferation (Takeshita, Faccio et al. 2007) and Src family kinases interaction putative site (Alonso, Koegl et al. 1995)), Y^{669} (Grb2, Mona and Socs1 interactions; (van der Geer and Hunter 1993; Bourette, Arnaud et al. 1998; Bourette, De Sepulveda et al. 2001)), Y^{691} (p85 PI3K, PLCγ2 and Socs1 interactions; (Reedijk, Liu et al. 1992; van der Geer and Hunter 1993; Bourette, Myles et al. 1997; Bourette, De Sepulveda et al. 2001)) and Y^{771} (required for mitogenesis; (Roussel, Shurtleff et al. 1990)). Surprisingly, CSF1R2/b lacks an equivalent of the murine C-terminal Y^{974}, known to be essential for the receptor degradation (reviewed in (Pixley and Stanley 2004)).

The sea bass CSF1R2/b protein has a marked decrease in similarity with the CSF1R and CSF1R1/a sequences of other organisms (Annex I, Fig. A.4) from aspartate 315 (D^{315}) to valine 396 (V^{396}). This restricted dissimilarity is verified for all fish CSF1R2 and CSF1Rb

proteins (*Dicentrarchus labrax*, *Gasterosteus aculeatus*, *Takifugu rubripes*, *Tetraodon nigroviridis*, *Haplochromis burtoni* and *Oryzias latipes*).

The differences with the murine and human proteins are greater in CSF1R2/b, mainly in the extracellular domain. The low sequence similarity from D^{315} to V^{396} and the related predicted structural differences in the extracellular membrane proximal region may have implications in the CSF1R2/b dimerization and internalization ability and thus, in its signal transduction ability, as dimerization is mandatory for the activation of the intracellular kinase.

The intracellular kinase domain is highly conserved both in CSF1R1/a and CSF1R2/b (Fig. 5), which might indicate that its intracellular enzymatic activity and related signaling functions are preserved in sea bass.

Phylogenetic analysis of the CSF1R proteins

The sea bass CSF1R1/a exhibits very high similarity and identity to the CSF1R proteins of other vertebrates (Table 2), mainly with other fish CSF1R, CSF1R1 and CSF1Ra proteins. All of the proteins already identified as CSF1R1 or CSF1Ra (that resulted from the fish specific genome duplication), namely those of the bony fishes *Gasterosteus aculeatus*, *Haplochromis burtoni*, *Tetraodon nigroviridis* and *Oryzias latipes*, have similarity values of more than 90.0 % and identity values of more than 83.5 %. Interestingly, the protein that revealed to be the most dissimilar of sea bass CSF1R1/a was the *O. latipes* CSF1R2 (48.8 % similarity and 32.4 % identity), an organism that also has one of the most similar proteins (*O. latipes* CSF1R1; 91.7 % similarity and 83.7 % identity), a fact which should indicate a marked evolutionary divergence within this specie.

The sea bass CSF1R2/b also exhibits high similarity and identity to the CSF1R proteins of other vertebrates (Table 2), though significantly less than CSF1R1/a. It has similarity values of more than 71.0 % and identity values of more than 59.0 % with the CSF1R2 or CSF1Rb proteins of *T. nigroviridis*, *Takifugu rubripes* and *H. burtoni*, having the maximum similarity and identity values with the *G. aculeatus* CSF1R2 protein (73.8 % similarity and 64.2 % identity). The similarity and identity of the sea bass CSF1R2/b with the mammalian counterparts are lower than the obtained for CSF1R1/a; sea bass CSF1R2/a has 51.7 % similarity and 37.4 % identity with the human CSF1R protein and 53.7 % similarity and 36.6 % identity with the murine CSF1R.

Table 2. Percentages of similarity and identity of CSF1R1/a or CSF1R2/b from sea bass and other species. Values were determined with MatGAT 2.02 using the BLOSUM64 scoring matrix and the values 16 for first gap and 4 for extended gap penalties.

	CSF1R1/a				CSF1R2/b		
Species	Similarity (%)	Identity (%)	Accession numbers	Species	Similarity (%)	Identity (%)	Accession numbers
Sparus aurata CSF1R	96,3	93,0	CAJ18352	*Gasterosteus aculeatus* CSF1R2	73,8	64,2	ENSGACP00000027653
Gasterosteus aculeatus CSF1R1	93,0	87,4	ENSGACP00000023794	*Haplochromis burtoni* CSF1Rb	73,4	62,5	ABD48799
Haplochromis burtoni CSF1Ra	91,6	84,7	ABD48801	*Takifugu rubripes* CSF1R2	72,3	61,5	Q8UVR8
Tetraodon nigroviridis CSF1R1	92,0	84,2	ENSTNIP00000017239	*Tetraodon nigroviridis* CSF1R2	71,2	59,1	ENSTNIP00000010425
Oryzias latipes CSF1R1	91,7	83,7	ENSORLP00000006110	*Oryzias latipes* CSF1R2	62,6	46,3	ENSORLP00000016040
Takifugu rubripes CSF1R1	91,0	83,7	P79750	*Salmo salar* CSF1R	62,1	46,0	ACU57063
Oncorhynchus mykiss CSF1R	85,7	75,6	Q8AXU0	*Oncorhynchus mykiss* CSF1R	61,1	45,2	Q76KE8
Carassius auratus CSF1R	78,3	65,2	AAT08795	*Tetraodon nigroviridis* CSF1R1	57,7	41,9	ENSTNIP00000017239
Danio rerio CSF1R	78,2	64,6	Q9I8N6	*Oryzias latipes* CSF1R1	57,6	41,9	ENSORLP00000006110
Oncorhynchus mykiss CSF1R	70,7	56,6	Q76KE8	*Takifugu rubripes* CSF1R1	59,0	41,9	P79750
Salmo salar CSF1R	70,8	56,5	ACU57063	*Oncorhynchus mykiss* CSF1R	56,9	41,7	Q8AXU0
Danio rerio CSF1R	63,3	47,5	XP_001919824	*Sparus aurata* CSF1R	58,0	41,6	CAJ18352
Gallus gallus CSF1R	62,8	45,5	XP_414597	*Danio rerio* CSF1R	58,2	41,6	Q9I8N6
Haplochromis burtoni CSF1Rb	60,8	45,3	ABD48799	*Carassius auratus* CSF1R	57,4	41,3	AAT08795
Homo sapiens CSF1R	60,1	44,6	P07333	*Dicentrarchus labrax* CSF1Ra	57,5	41,2	
Takifugu rubripes CSF1R2	58,3	42,9	Q8UVR8	*Haplochromis burtoni* CSF1Ra	57,6	41,1	ABD48801
Xenopus tropicalis CSF1R	60,0	42,8	NP_001008181	*Gasterosteus aculeatus* CSF1R1	56,4	40,2	ENSGACP00000023794
Mus musculus CSF1R	59,4	42,7	P09581	*Danio rerio* CSF1R	56,2	39,0	XP_001919824
Tetraodon nigroviridis CSF1R2	58,9	41,5	ENSTNIP00000010425	*Homo sapiens* CSF1R	51,7	37,4	P07333
Gasterosteus aculeatus CSF1R2	56,8	41,3	ENSGACP00000027653	*Gallus gallus* CSF1R	53,4	37,3	XP_414597
Dicentrarchus labrax CSF1Rb	57,5	41,2		*Mus musculus* CSF1R	53,7	36,6	P09581
Danio rerio Kit-b	55,7	39,4	CAX14292	*Danio rerio* Kit-b	54,4	35,6	CAX14292
Homo sapiens Kit	55,5	37,6	AAC50968	*Xenopus tropicalis* CSF1R	55,2	35,2	NP_001008181
Danio rerio Kit-a	52,2	36,5	AAI62502	*Danio rerio* Kit-a	50,8	34,7	AAI62502
Oryzias latipes CSF1R2	48,8	32,4	ENSORLP00000016040	*Homo sapiens* Kit	52,3	34,4	AAC50968

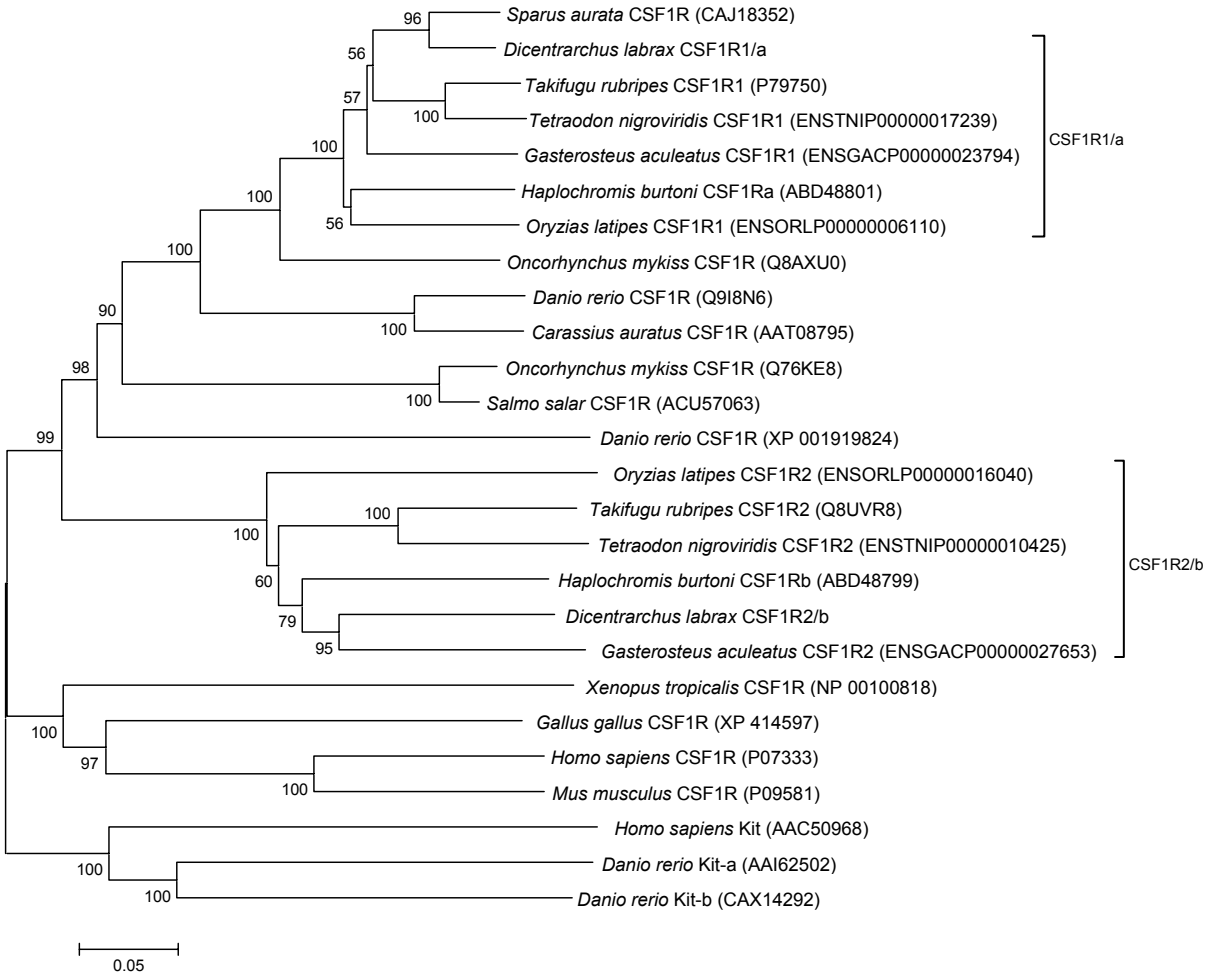


Figure 6. Neighbour-joining tree of CSF1R proteins. Sequences were aligned using ClustalW with default parameters; the tree was construted using the MEGA4 software with p-distance and complete deletion of gaps; the numbers in the branches nodes denote the bootstrap percentages of 1000 replicates. Sequences identified as CSF1R1/a or CSF1R2/b are grouped by brackets together with the sea bass sequences. The outgroup (Kit) is denoted by a doted branch. Accession numbers are indicated for each respective sequence; accession numbers starting with the letters ENS refer to Ensembl genomes and the remaining sequences are from the NCBI databases.

The neighbor-joining tree (Fig. 6) clearly recapitulates what is thought to be the evolution of the proteins and confirms the nature of the identified sea bass sequences, CSF1R1/a and CSF1R2/b. An outgroup branch comprises the Kit proteins that are though to have derived from the same ancestor of the CSF1R proteins. The CSF1R proteins of non-fish vertebrates cluster in a branch that is parallel to the fish CSF1R protein branches; in accordance with the FSGD event that occurred after the separation of bony fish and tetrapods, the fish CSF1R

line is split in two branches, one that comprises the CSF1R2/b proteins and another that clusters the CSF1R1/a proteins with similar CSF1R proteins that presumably are CSF1R1/a.

A comparative analysis of the similarity and identity values of both molecules, CSF1R1/a and CSF1R2/b (for the organisms were both proteins were identified), reveals that the CSF1R2/b proteins have systematically less similarity and identity with the mammalian (human and mouse) CSF1R proteins; this fact is in concordance with the analysis of the amino acid sequences of the sea bass CSF1R1/a and CSF1R2/b, where the first protein revealed more similarities in structural and functional key features of the receptor with the mammalian counterparts.

The differences observed between the two receptors suggest a functional differentiation of CSF1R1/a and CSF1R2/b proteins, not only in sea bass but also in the other bony fish considered in this study. CSF1R2/b function is, presently, not clear, but the similarity and identity values between fish CSF1R2/b sequences suggest that this gene is under some kind of evolutionary pressure whose nature is still unknown but which should be different, at least its intensity, to the evolutionary pressure exerted over the CSF1R1/a gene whose protein is more strictly conserved. Sequencing and analysis of the sea bass CSF1R1/a and CSF1R2/b genomic sequences may give some more insights into the mechanism by which the proteins are diverging.

Structure models

The obtained sea bass CSF1R1/a D1-D5 homology model has a final total energy of -13856.626 KJ/mol and comparative analysis with the template (Kit D1-D5) revealed 28% structural identity over 461 C^{α} atoms and a RMSD value of 0.682 Å; the CSF1R2/b D1-D5 homology model has a final total energy of -499.248 KJ/mol and comparative analysis with the template (Kit D1-D5) revealed 24% structural identity over 406 C^{α} atoms and a RMSD value of 0.758 Å. The sea bass CSF1R1/a D1-D3 homology model has a final total energy of -8379.371 KJ/mol and comparative analysis with the template (3EJJ) revealed 30% structural identity over 269 C^{α} atoms and a RMSD value of 0.508 Å; the CSF1R2/b D1-D3 homology model has a final total energy of -8766.259 KJ/mol and comparative analysis with the template (3EJJ) revealed 30% structural identity over 267 C^{α} atoms and a RMSD value of 0.540 Å. The sea bass CSF1R1/a intracellular homology model has a final total energy of -15847.888 KJ/mol and comparative analysis with the template (2OGV) revealed 75%

structural identity over 317 C^{α} atoms and a RMSD value of 0.292 Å; the CSF1R2/b intracellular homology model has a final total energy of -13085.956 KJ/mol and comparative analysis with the template (2OGV) revealed 70% structural identity over 299 C^{α} atoms and a RMSD value of 0.203 Å.

The CSF1R1/a and CSF1R2/b ectodomain structures (Fig. 7.A) have a serpentine outline that have an abrupt curve in their N-terminal region, resulting from the sharp angle formed by the N-terminal domains D1 and D2. All five domains (D1, D2, D3, D4 and D5) of the CSF1R1/a model display an immunoglobulin fold, with 7 to 9 *β*-strands forming two antiparallel *β*-sheets (Fig. 7.B and 7.E). In the case of the domains D1, D2, D3 and D4, the *β*-sandwiches formed by the *β*-sheets are further stabilized by disulfide bonds (Fig. 7.B and 7.E). The CSF1R2/b model lacks the immunoglobulin fold in its D4 domain and the disulfide bridge in its D5 domains (Fig. 7.F).

Even though the overall conformation of the 3 N-terminal domains is conserved both in CSF1R1/a and CSF1R2/b, the majority of the residues that, in the mammalian counterparts, were identified as important for ligand binding are not conserved (Chen, Liu et al. 2008). This fact may suggest that the receptors affinity has changed. To give some more insight into this question, further modeling should be done including the ligand. The structure of the D4 domain and relative position of the dimerization motif is conserved in CSF1R1/a (Fig. 7.H) but not in CSF1R2/b, which suggests that CSF1R2/b may have lost its dimerization ability.

The amino acid sequences of the intracellular kinase domains of both receptors is highly conserved and, accordingly, the obtained homology models are highly similar to the mammalian counterparts. Both display the typical bi-lobed form (Fig. 8.B), composed by the N-terminal lobe that comprises the nucleotide binding loop (Fig. 8.A) and by the C-terminal lobe that comprises the activation and catalytic loop (Fig. 8.A). The kinase insert (KI) domain, which is mainly disordered, is located outside of the catalytic bi-lobed structure; this structure is not necessary for kinase activity but it is thought to be important for the recognition of some proteins (Taylor, Reedijk et al. 1989).

Sequence and structural conservation suggests that the intracellular kinase activity of both receptors is conserved.

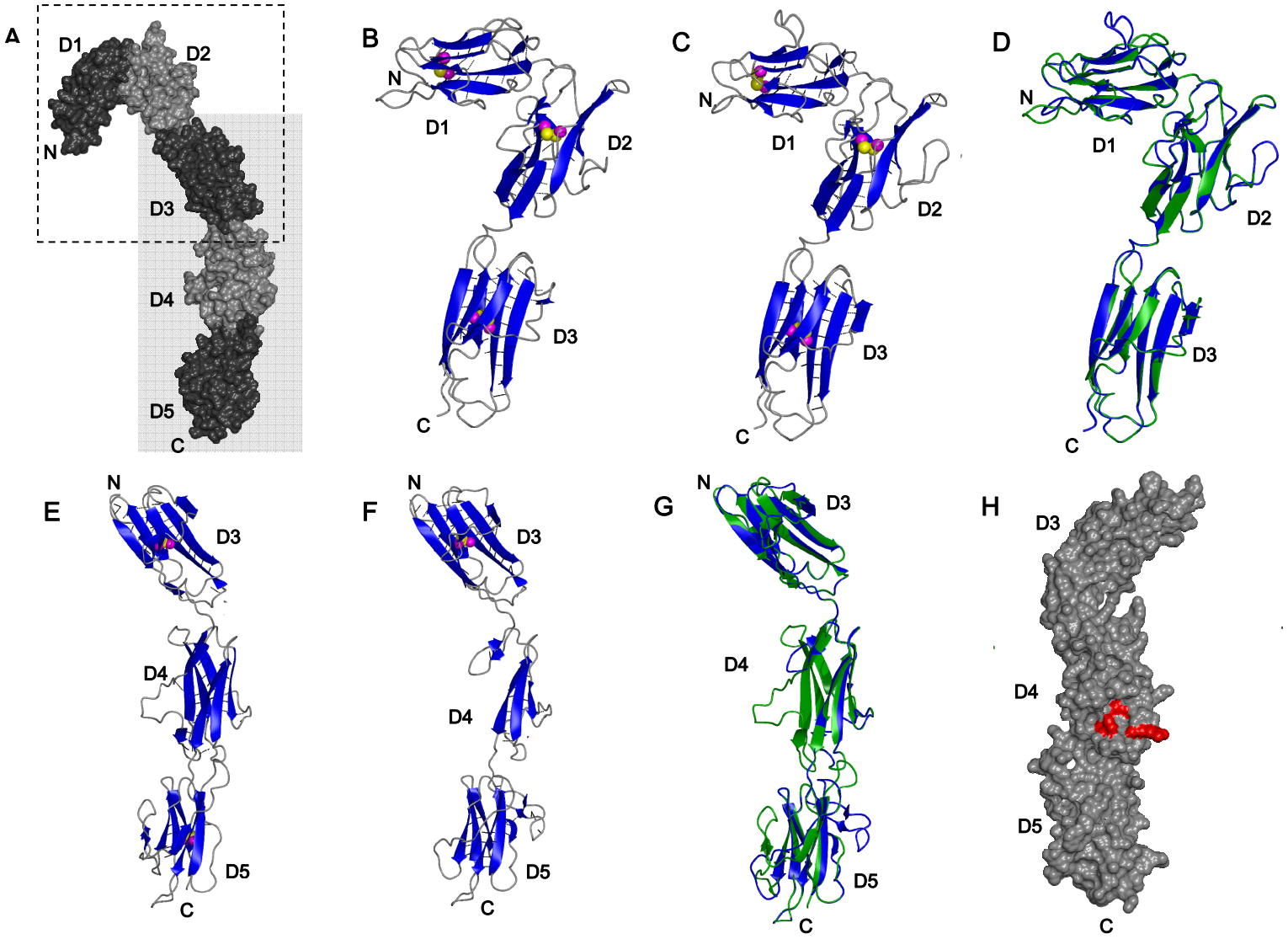

A
D1
D2
N
D3
D4
D5
C
B
N
D1
D2
D3
C
C
N
D1
D2
D3
C
D
N
D1
D2
D3
C
E
N
D3
D4
D5
C
F
N
D3
D4
D5
C
G
N
D3
D4
D5
C
H
D3
D4
D5
C

Figure 7. Extracellular domain of sea bass CSF1R proteins. (A) Model of the full-lenght sea bass CSF1R1/a extracellular domain. Ribbon diagrams of the D1-D3 domains of sea bass (B) CSF1R1/a and (C) CSF1R2/b and (D) the respective superposition (CSF1R1/a in green and CSF1R2/b in blue). Ribbon diagrams of the D3-D5 domains of sea bass (E) CSF1R1/a and (F) CSF1R2/b and (G) the respective superposition (CSF1R1/a in green and CSF1R2/b in blue). (H) The sea bass CSF1R1/a conserved residues that were identified as being important for class III RTKs' dimerization are shown in red. Structure elements are shown in red (*α*-helices) and blue (*β*-strands), disulfide bonds are denoted by pink and yellow balls and the hydrogen bonds between *β*-strands are depicted by dashed black lines. The domains delimited in A by the dashed rectangle were modelled according to the murine CSF1R structure (B, C and D) and the domains shaded in gray were modelled according to the human Kit protein (E, F, G and H) as decribed in the Materials and methods section.

CSF1R antibodies

It was verified, in cell cultures of *C. auratus*, that CSF1R is expressed throughout macrophage development since the early progenitors, a characteristic that makes this receptor a putative specific marker of the fish monocytic-macrophagic lineage as it is for other vertebrates (Barreda, Hanington et al. 2005).

A specific marker of sea bass macrophages (or macrophage progenitors) would be a valuable tool to study sea bass immunology and infection. In that sense we produced rabbit monospecific polyclonal antibodies against peptides of sea bass CSF1R1/a and CSF1R2/b. The affinity-purified antibodies were tested by western blot (Fig. 9) against extracts of bacteria expressing the recombinant proteins and extracts of sea bass head kidney leukocytes, spleen leukocytes, peripheral blood leukocytes, erythrocytes and peritoneal cavity macrophages.

The antibody directed against the sea bass CSF1R1/a detected the recombinant form of the protein but did not detect a protein with a suitable MW in all sea bass cells tested. The antibody directed against the sea bass CSF1R2/b did not mark the recombinant form of the protein but detected a single protein of a suitable MW in all tested sea bass cells, mainly in head kidney and spleen leukocytes. The single band that was detected with the CSF1R2/b antibody has an estimated MW in the range of 130 - 150 kDa, 25 – 45 kDa in excess with what is predicted for the amino acid sequence but in agreement with what is observed in mammals, as the CSF1R proteins are extensively glycosylated. The CSF1R2/b antibody dilution was optimized by performing western blots with peritoneal cavity macrophages and the protein was detected with dilutions as far as 1:10000.

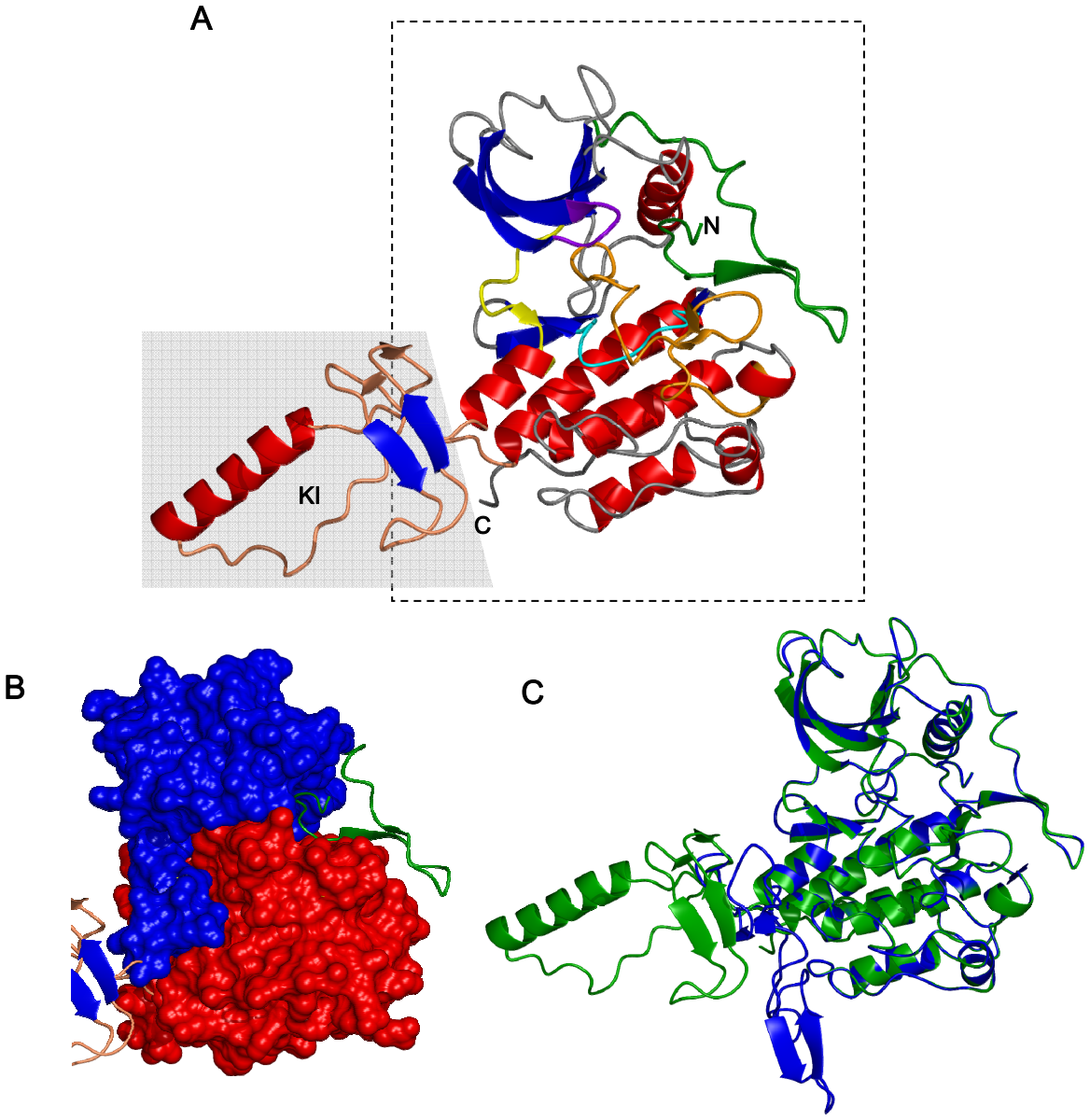


Figure 8. Intracellular kinase domain of sea bass CSF1R proteins. (A) The sea bass CSF1R1/a intracellular kinase domain depicted in a ribbon diagram; the core kinase is delimited by the dashed square and the kinase insert (KI) is shaded in gray; secondary structure elements are shown in red (*α*-helices) and blue (*β*-strands); the juxtamembrane domain is shown in green; the activation loop is shown in orange; the glycine rich nucleotide binding loop is shown in purple; the hinge region that coordinates the two lobes is shown in yellow; the catalytic loop is shown in cyan; the kinase insert (KI) domain is denoted in light brown (gray shaded area). (B) The surface of the two lobes that comprise the kinase itself. (C) Superposition of the sea bass CSF1R1/a (green) and CSF1R2/b (blue) protein models.

To verify the authenticity of the marked protein, immunoprecipitation experiments were performed. Even though they were repeated and optimized several times, it was at all impossible to immunoprecipitate a valid target and thus, impossible to confirm the protein identity. Nevertheless, it may be useful to perform more studies, such as assess the protein expression under several stimuli known to affect CSF1R protein expression, which can be done *ex vivo* with cells from the peritoneal cavity (highly enriched in macrophages) or with leukocytes from other organs that are immunological relevants such as the head kidney and spleen.

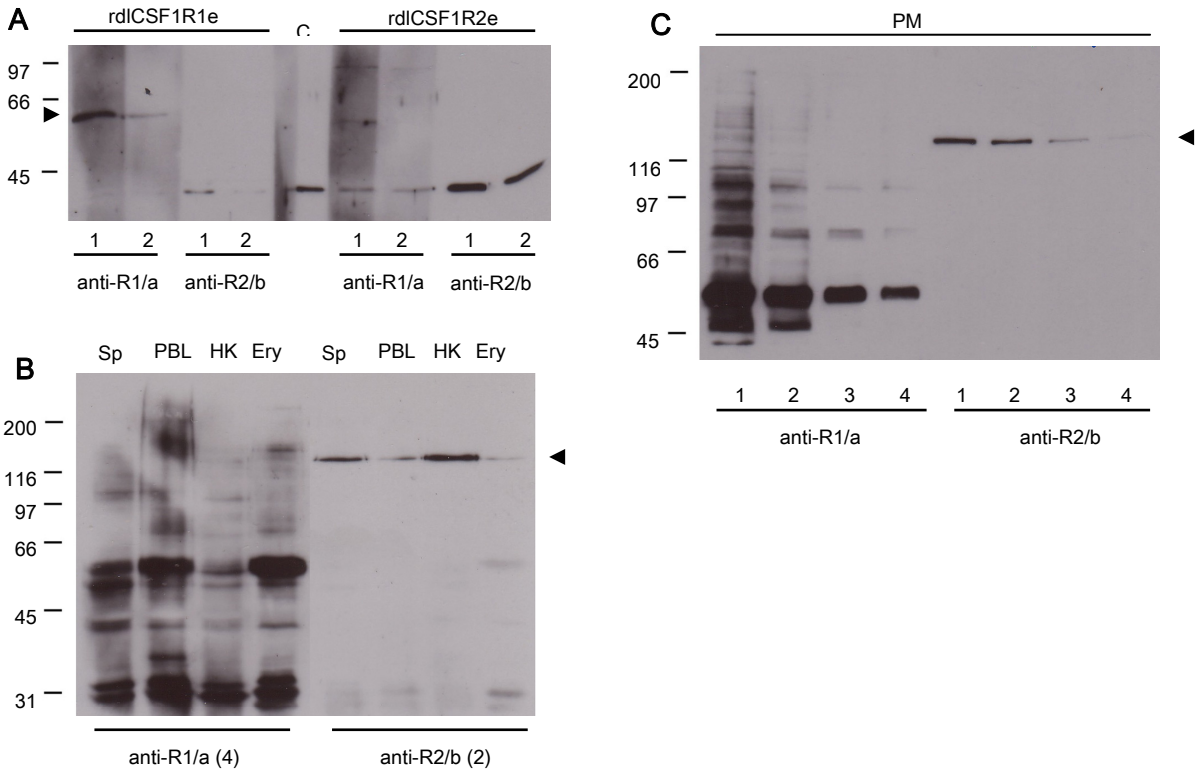


Figure 9. Western blot of CSF1R proteins. The affinity-purified antibodies produced against sea bass CSF1R1/a (anti-R1/a) and CSF1R2/b (anti-R2/b) peptide sequences were tested against [A] extracts of *E. coli* expressing sea bass CSF1R1/a recombinant protein (rdlCSF1R1e) and sea bass CSF1R2/b recombinant protein (rdlCSF1R2e), [B] spleen (Sp) leukocytes, head kidney (HK) leukocytes, peripheral blood leukocytes (PBL), erythrocytes (Ery) (1 x 10^6 cells per lane) and [C] peritoneal cavity macrophages (PM). The expected molecular weights are 60.1 kDa for rdlCSF1R1e, 55.8 kDa for rdlCSF1R2e, 109.5 kDa for the sea bass CSF1R1/a protein and 103.5 kDa for the sea bass CSF1R2/b protein. The numbers indicate the antibody dilution (1) 1:1000, (2) 1:2000, (3) 1:5000 and (4) 1:10000; arrowheads indicate results of interest; molecular weights (MW) are indicated on the left in kDa.

The antibodies were also tested for immunofluorescence microscopy with fixed cells, an application for which the antibodies revealed to be unsuitable.

In vitro culture of macrophages or macrophage-like cells, has been an invaluable tool for infection and immunology studies. Salt water fish macrophages, sea bass' included, have been showing to be extremely hard to culture *in vitro*. Taking that into account, the antibodies were also tested for flow citometry, aiming to have an antibody that is suitable to perform cell sorting and thus, that can be used to label and enrich macrophage cultures. Unfortunately, both antibodies showed none or very little ability to stain and isolate a specific cell population from mixed leukocyte populations from head kidney, spleen, thymus, blood and peritoneal cavity.

The CSF1R2/b antibody should be further evaluated to confirm its specificity and another approach can be used to design new antibodies to these proteins that can be used in important applications such immunofluorescence, flow citometry and fluorescence activated cell sorting (FACS).

The sea bass CSF1R proteins

Altogether, the observed differences between the two receptors, in agreement with what was observed for other bony fish species, suggest that the two receptors might have different functions in ligand affinity and even in intracellular signaling. In that sense, further studies, such as the analysis of the expression of both receptors, and possibly their identified ligands CSF-1 and IL-34, is of extreme importance. Gene expression should be accessed for tissues and organs of resting fish and fish submitted to stimuli that are known to affect CSF1R expression. This expression data might give some more insights into the receptors' function in fish.

The CSF1R gene expression was previously assessed for *Oncorhynchus mykiss* (Honda, Nishizawa et al. 2005), *T. rubripes* (Williams, Brenner et al. 2002), *Sparus aurata* (Roca, Sepulcre et al. 2006; Chaves-Pozo, Liarte et al. 2008) and *Carassius auratus* (Barreda, Hanington et al. 2005). CSF1R was observed to be expressed mainly in the head kidney, spleen, gills, thymus, and liver and at less extent in the blood and peritoneal exsudate of *S. aurata* (Roca, Sepulcre et al. 2006). Besides confirming the expression of the CSF1R genes in immunological relevant organs such as the head kidney, spleen, gills and thymus, the previous study revealed that both CSF1R1/a and CSF1R2/b expression were altered upon

infection, which might indicate that these genes and proteins have in fish, as in other vertebrates, immunological relevant functions. Indeed, studies performed in some fish species, equivalent to those performed in mammals, showed that several biological properties of the CSF-1 growth factor were mediated by CSF1R (Herbomel, Thisse et al. 2001; Roca, Sepulcre et al. 2006; Mulero, Pilar Sepulcre et al. 2008; Hanington, Hitchen et al. 2009).

Altogether, this data suggest that CSF1R (both CSF1R1/a and CSF1R2/b) perform functions similar to those of the mammalian counterparts. The extent of this functional similarity and the bony fish, particularly the sea bass, specificities should be addressed in more extensive studies.

Conclusions

Two cDNAs that code for sea bass proteins homologous to known vertebrate CSF1R proteins were identified and characterized. In agreement to what is described for other fish species, two CSF1R cDNAs were isolated that code for different proteins, the sea bass CSF1R1/a protein that has high similarity and identity with other fish CSF1R1 and CSF1Ra proteins and the sea bass CSF1R2/b protein that has high similarity and identity with other fish CSF1R2 and CSF1Rb proteins.

Of these two proteins, CSF1R1/a is the most similar to the mammalian counterpart (CSF1R), both in its extracellular domain and intracellular domain. Analysis of the protein alignments and of the modeled structures show that key elements of the protein are conserved such as the dimerization motif, the regulatory juxtamembrane domain and the kinase domain, suggesting that this protein should perform, in sea bass, the functions of the mammalian CSF1R. CSF1R2/b most significant difference is the absence of the dimerization motif; even though this molecule may still be able to bind the ligand and have some regulatory function.

The attempts to confirm the identity of the protein products of both cDNAs were unfruitful to date, which hindered further studies on the proteins functions. Nevertheless, valuable insight to this question can be given by other means such as gene expression analysis, which has already been done for some fish species and gave some clues about the genes' function.

Even though the efforts to develop antibodies to these proteins have been unproductive, they should be continued since the antibodies would be of several important uses. Knowledge about the biology of CSF1R itself and their ligands is important to understand the development and action of macrophages and thus, the immunological response to invading pathogens, which is of extra value for the study of the sea bass infection by *Phdp*. From a research technical perspective, they would allow for specific macrophage labeling, thus enabling a more controlled establishment and maintenance of fish cell cultures, namely mononuclear phagocytic cell cultures that are indispensable research tools.

Bibliography

Abdel-Aziz, E.-S. H., S. B. S. Abdu, et al. (2009). "Haemopoiesis in the head kidney of tilapia, Oreochromis niloticus (Teleostei: Cichlidae): a morphological (optical and ultrastructural) study." Fish Physiol Biochem.

Acosta, F., J. Vivas, et al. (2009). "Invasion and survival of Photobacterium damselae subsp. piscicida in non-phagocytic cells of gilthead sea bream, Sparus aurata L." J Fish Dis **32**(6): 535-41.

Alonso, G., M. Koegl, et al. (1995). "Sequence requirements for binding of Src family tyrosine kinases to activated growth factor receptors." J Biol Chem **270**(17): 9840-8.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.

Alvarez, F., E. Flano, et al. (1994). "Seasonal intrathymic erythropoietic activity in trout." Dev Comp Immunol **18**(5): 409-20.

Arnold, K., L. Bordoli, et al. (2006). "The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling." Bioinformatics **22**(2): 195-201.

Auffray, C., M. H. Sieweke, et al. (2009). "Blood monocytes: development, heterogeneity, and relationship with dendritic cells." Annu Rev Immunol **27**: 669-92.

Aviles-Trigueros, M. and J. A. Quesada (1995). "Myelopoiesis in the thymus of the sea bass, Dicentrarchus labrax L. (teleost)." Anat Rec **242**(1): 83-90.

Bakopoulos, V., A. Hanif, et al. (2004). "The effect of in vivo growth on the cellular and extracellular components of the marine bacterial pathogen Photobacterium damsela subsp. piscicida." J Fish Dis **27**(1): 1-13.

Barreda, D. R., P. C. Hanington, et al. (2004). "Regulation of myeloid development and function by colony stimulating factors." Dev Comp Immunol **28**(5): 509-54.

Barreda, D. R., P. C. Hanington, et al. (2005). "A novel soluble form of the CSF-1 receptor inhibits proliferation of self-renewing macrophages of goldfish (Carassius auratus L.)." Dev Comp Immunol **29**(10): 879-94.

Berman, H. M., J. Westbrook, et al. (2000). "The Protein Data Bank." Nucleic Acids Res **28**(1): 235-42.

Bjerkas, E., O. Evensen, et al. (2006). Systemic Pathology of Fish: A Text and Atlas of Normal Tissues in Teleost and their Responses in Disease, Scotian Press.

Bourette, R. P., S. Arnaud, et al. (1998). "Mona, a novel hematopoietic-specific adaptor interacting with the macrophage colony-stimulating factor receptor, is implicated in monocyte/macrophage development." EMBO J **17**(24): 7273-81.

Bourette, R. P., P. De Sepulveda, et al. (2001). "Suppressor of cytokine signaling 1 interacts with the macrophage colony-stimulating factor receptor and negatively regulates its proliferation signal." J Biol Chem **276**(25): 22133-9.

Bourette, R. P., G. M. Myles, et al. (1997). "Sequential activation of phoshatidylinositol 3-kinase and phospholipase C-gamma2 by the M-CSF receptor is necessary for differentiation signaling." EMBO J **16**(19): 5880-93.

Braasch, I., W. Salzburger, et al. (2006). "Asymmetric evolution in two fish-specifically duplicated receptor tyrosine kinase paralogons involved in teleost coloration." Mol Biol Evol **23**(6): 1192-202.

Branch, D. R., A. R. Turner, et al. (1989). "Synergistic stimulation of macrophage proliferation by the monokines tumor necrosis factor-alpha and colony-stimulating factor 1." Blood **73**(1): 307-11.

Campanella, J. J., L. Bitincka, et al. (2003). "MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences." BMC Bioinformatics **4**: 29.

Cecchini, M. G., M. G. Dominguez, et al. (1994). "Role of colony stimulating factor-1 in the establishment and regulation of tissue macrophages during postnatal development of the mouse." Development **120**(6): 1357-72.

Chaves-Pozo, E., S. Liarte, et al. (2008). "Pattern of expression of immune-relevant genes in the gonad of a teleost, the gilthead seabream (Sparus aurata L.)." Mol Immunol **45**(10): 2998-3011.

Chen, X., H. Liu, et al. (2008). "Structure of macrophage colony stimulating factor bound to FMS: diverse signaling assemblies of class III receptor tyrosine kinases." Proc Natl Acad Sci U S A **105**(47): 18267-72.

Chihara, T., S. Suzu, et al. (2010). "IL-34 and M-CSF share the receptor Fms but are not identical in biological activity and signal activation." Cell Death Differ.

Chitu, V. and E. R. Stanley (2006). "Colony-stimulating factor-1 in immunity and inflammation." Curr Opin Immunol **18**(1): 39-48.

Dai, X. M., G. R. Ryan, et al. (2002). "Targeted disruption of the mouse colony-stimulating factor 1 receptor gene results in osteopetrosis, mononuclear phagocyte deficiency, increased primitive progenitor cell frequencies, and reproductive defects." Blood **99**(1): 111-20.

De Bodt, S., S. Maere, et al. (2005). "Genome duplication and the origin of angiosperms." Trends Ecol Evol **20**(11): 591-7.

de Castro, E., C. J. Sigrist, et al. (2006). "ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins." Nucleic Acids Res **34**(Web Server issue): W362-5.

do Vale, A., C. Costa-Ramos, et al. (2007). "Systemic macrophage and neutrophil destruction by secondary necrosis induced by a bacterial exotoxin in a Gram-negative septicaemia." Cell Microbiol **9**(4): 988-1003.

do Vale, A., F. Marques, et al. (2003). "Apoptosis of sea bass (Dicentrarchus labrax L.) neutrophils and macrophages induced by experimental infection with Photobacterium damselae subsp. piscicida." Fish Shellfish Immunol **15**(2): 129-44.

do Vale, A., M. T. Silva, et al. (2005). "AIP56, a novel plasmid-encoded virulence factor of Photobacterium damselae subsp. piscicida with apoptogenic activity against sea bass macrophages and neutrophils." Mol Microbiol **58**(4): 1025-38.

Douglass, T. G., L. Driggers, et al. (2008). "Macrophage colony stimulating factor: not just for macrophages anymore! A gateway into complex biologies." Int Immunopharmacol **8**(10): 1354-76.

Ellis, A. E. (1999). "Immunity to bacteria in fish." Fish Shellfish Immunol **9**: 291-308.

Emanuelsson, O., S. Brunak, et al. (2007). "Locating proteins in the cell using TargetP, SignalP and related tools." Nat Protoc **2**(4): 953-71.

Esteban, M. A., J. Meseguer, et al. (1989). "Erythropoiesis and thrombopoiesis in the head-kidney of the sea bass (Dicentrarchus labrax L.): an ultrastructural study." Arch Histol Cytol **52**(4): 407-19.

FAO. (2004-2010, 6 June 2003). "SIDP Home Page. FAO Species Identification and Data Programme (SIDP). FI Programme Websites." Retrieved 30 March 2010, 2010, from http://www.fao.org/fishery/culturedspecies/Dicentrarchus_labrax/en.

Force, A., M. Lynch, et al. (1999). "Preservation of duplicate genes by complementary, degenerative mutations." Genetics **151**(4): 1531-45.
Froese, R. and D. Pauly. (2010). "FishBase." Retrieved 30 March 2010, 2010, from http://www.fishbase.org/Summary/speciesSummary.php?ID=63&genusname=Dicentrarchus&speciesname=labrax.
Gasteiger, E., A. Gattiker, et al. (2003). "ExPASy: The proteomics server for in-depth protein knowledge and analysis." Nucleic Acids Res **31**(13): 3784-8.
Gauthier, G., B. Lafay, et al. (1995). "Small-subunit rRNA sequences and whole DNA relatedness concur for the reassignment of Pasteurella piscicida (Snieszko et al.) Janssen and Surgalla to the genus Photobacterium as Photobacterium damsela subsp. piscicida comb. nov." Int J Syst Bacteriol **45**(1): 139-44.
Gilbert, H. S., V. Praloran, et al. (1989). "Increased circulating CSF-1 (M-CSF) in myeloproliferative disease: association with myeloid metaplasia and peripheral bone marrow extension." Blood **74**(4): 1231-4.
Grayfer, L., P. C. Hanington, et al. (2009). "Macrophage colony-stimulating factor (CSF-1) induces pro-inflammatory gene expression and enhances antimicrobial responses of goldfish (Carassius auratus L.) macrophages." Fish Shellfish Immunol **26**(3): 406-13.
Guilbert, L. J. and E. R. Stanley (1986). "The interaction of 125I-colony-stimulating factor-1 with bone marrow-derived macrophages." J Biol Chem **261**(9): 4024-32.
Hanamura, T., K. Motoyoshi, et al. (1988). "Quantitation and identification of human monocytic colony-stimulating factor in human serum by enzyme-linked immunosorbent assay." Blood **72**(3): 886-92.
Hanington, P. C., S. J. Hitchen, et al. (2009). "Macrophage colony stimulating factor (CSF-1) is a central growth factor of goldfish macrophages." Fish Shellfish Immunol **26**(1): 1-9.
Hanington, P. C., T. Wang, et al. (2007). "Growth factors of lower vertebrates: characterization of goldfish (Carassius auratus L.) macrophage colony-stimulating factor-1." J Biol Chem **282**(44): 31865-72.
Herbomel, P., B. Thisse, et al. (2001). "Zebrafish early macrophages colonize cephalic mesenchyme and developing brain, retina, and epidermis through a M-CSF receptor-dependent invasive process." Dev Biol **238**(2): 274-88.
Honda, T., T. Nishizawa, et al. (2005). "Molecular cloning and expression analysis of a macrophage-colony stimulating factor receptor-like gene from rainbow trout, Oncorhynchus mykiss." Mol Immunol **42**(1): 1-8.
Horiuchi, K., T. Miyamoto, et al. (2007). "Cell surface colony-stimulating factor 1 can be cleaved by TNF-alpha converting enzyme or endocytosed in a clathrin-dependent manner." J Immunol **179**(10): 6715-24.
How, G. F., B. Venkatesh, et al. (1996). "Conserved linkage between the puffer fish (Fugu rubripes) and human genes for platelet-derived growth factor receptor and macrophage colony-stimulating factor receptor." Genome Res **6**(12): 1185-91.
Hubbard, S. R. and J. H. Till (2000). "Protein tyrosine kinase structure and function." Annu Rev Biochem **69**: 373-98.
Hultman, K. A., N. Bahary, et al. (2007). "Gene Duplication of the zebrafish kit ligand and partitioning of melanocyte development functions to kit ligand a." PLoS Genet **3**(1): e17.
Krissinel, E. and K. Henrick (2004). "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2256-68.
Larkin, M. A., G. Blackshields, et al. (2007). "Clustal W and Clustal X version 2.0." Bioinformatics **23**(21): 2947-8.

Lin, H., E. Lee, et al. (2008). "Discovery of a cytokine and its receptor by functional screening of the extracellular proteome." Science **320**(5877): 807-11.
Lopez-Doriga, M. V., A. C. Barnes, et al. (2000). "Invasion of fish epithelial cells by Photobacterium damselae subsp. piscicida: evidence for receptor specificity, and effect of capsule and serum." Microbiology **146 (Pt 1)**: 21-30.
Lu, G. and E. N. Moriyama (2004). "Vector NTI, a balanced all-in-one sequence analysis suite." Brief Bioinform **5**(4): 378-88.
Magariños, B., J. L. Romalde, et al. (1996). "Adherence and invasive capacities of the fish pathogen Pasteurella piscicida." FEMS Microbiol Lett **138**(1): 29-34.
Magariños, B., Y. Santos, et al. (1992). "Pathogenic activities of live cells and extracellular products of the fish pathogen Pasteurella piscicida." J Gen Microbiol **138**(12): 2491-8.
Magariños, B., A. E. Toranzo, et al. (1996). "Phenotypic and pathobiological characteristics of *Pasteurella piscicida*." Annual Review of Fish Diseases **6**: 41-64.
Mancini, A., A. Koch, et al. (2002). "c-Cbl associates directly with the C-terminal tail of the receptor for the macrophage colony-stimulating factor, c-Fms, and down-modulates this receptor but not the viral oncogene v-Fms." J Biol Chem **277**(17): 14635-40.
Mancini, A., R. Niedenthal, et al. (1997). "Identification of a second Grb2 binding site in the v-Fms tyrosine kinase." Oncogene **15**(13): 1565-72.
Meseguer, J., M. A. Esteban, et al. (1990). "Granulopoiesis in the head-kidney of the sea bass (Dicentrarchus labrax L.): an ultrastructural study." Arch Histol Cytol **53**(3): 287-96.
Meseguer, J., A. Lopez-Ruiz, et al. (1995). "Reticulo-endothelial stroma of the head-kidney from the seawater teleost gilthead seabream (Sparus aurata L.): an ultrastructural and cytochemical study." Anat Rec **241**(3): 303-9.
Mulder, N. and R. Apweiler (2007). "InterPro and InterProScan: tools for protein sequence classification and comparison." Methods Mol Biol **396**: 59-70.
Mulero, I., M. Pilar Sepulcre, et al. (2008). "Characterization of macrophages from the bony fish gilthead seabream using an antibody against the macrophage colony-stimulating factor receptor." Dev Comp Immunol **32**(10): 1151-9.
Neumann, N. F., D. Barreda, et al. (1998). "Production of a macrophage growth factor(s) by a goldfish macrophage cell line and macrophages derived from goldfish kidney leukocytes." Dev Comp Immunol **22**(4): 417-32.
Noya, M., B. Magariños, et al. (1995). "Sequential pathology of experimental pasteurellosis in gilthead seabream Sparus aurata. A light- and electron-microscopic study." Diseases of Aquatic Organisms **21**(3): 177-186.
Ohno, S. (1970). Evolution by gene duplication. New York, Springer-Verlag.
Oster, W., A. Lindemann, et al. (1987). "Tumor necrosis factor (TNF)-alpha but not TNF-beta induces secretion of colony stimulating factor for macrophages (CSF-1) by human monocytes." Blood **70**(5): 1700-3.
Parichy, D. M. and J. M. Turner (2003). "Temporal and cellular requirements for Fms signaling during zebrafish adult pigment pattern development." Development **130**(5): 817-33.
Pixley, F. J. and E. R. Stanley (2004). "CSF-1 regulation of the wandering macrophage: complexity in action." Trends Cell Biol **14**(11): 628-38.
Potterton, L., S. McNicholas, et al. (2004). "Developments in the CCP4 molecular-graphics project." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2288-94.
Press, C. M. and O. Evensen (1999). "The morphology of the immune system in teleost fishes." Fish Shellfish Immunol **9**: 309-318.

Reedijk, M., X. Liu, et al. (1992). "Tyr721 regulates specific binding of the CSF-1 receptor kinase insert to PI 3'-kinase SH2 domains: a model for SH2-mediated receptor-target interactions." EMBO J **11**(4): 1365-72.

Roberts, W. M., A. T. Look, et al. (1988). "Tandem linkage of human CSF-1 receptor (c-fms) and PDGF receptor genes." Cell **55**(4): 655-61.

Roca, F. J., M. A. Sepulcre, et al. (2006). "The colony-stimulating factor-1 receptor is a specific marker of macrophages from the bony fish gilthead seabream." Mol Immunol **43**(9): 1418-23.

Roussel, M. F., S. A. Shurtleff, et al. (1990). "A point mutation at tyrosine-809 in the human colony-stimulating factor 1 receptor impairs mitogenesis without abrogating tyrosine kinase activity, association with phosphatidylinositol 3-kinase, or induction of c-fos and junB genes." Proc Natl Acad Sci U S A **87**(17): 6738-42.

Scapigliati, G., N. Romano, et al. (2002). "The immune system of sea bass, Dicentrarchus labrax, reared in aquaculture." Dev Comp Immunol **26**(2): 151-60.

Schultz, J., F. Milpetz, et al. (1998). "SMART, a simple modular architecture research tool: identification of signaling domains." Proc Natl Acad Sci U S A **95**(11): 5857-64.

Sherr, C. J. (1990). "Colony-stimulating factor-1 receptor." Blood **75**(1): 1-12.

Sherr, C. J., M. F. Roussel, et al. (1988). "Colony-stimulating factor-1 receptor (c-fms)." J Cell Biochem **38**(3): 179-87.

Sidow, A. (1996). "Gen(om)e duplications in the evolution of early vertebrates." Curr Opin Genet Dev **6**(6): 715-22.

Silva, M. T., N. M. dos Santos, et al. (2010). "AIP56: A novel bacterial apoptogenic toxin." Toxins **2**(4): 905-918.

Suzuki, M., I. Sekiguchi, et al. (1996). "Expression of c-fms proto-oncogene product by ovarian cancer cell lines with effects of macrophage colony-stimulating factor on proliferation." Oncology **53**(2): 99-103.

Takeshita, S., R. Faccio, et al. (2007). "c-Fms tyrosine 559 is a major mediator of M-CSF-induced proliferation of primary macrophages." J Biol Chem **282**(26): 18980-90.

Tamura, K., J. Dudley, et al. (2007). "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0." Mol Biol Evol **24**(8): 1596-9.

Taylor, G. R., M. Reedijk, et al. (1989). "The unique insert of cellular and viral fms protein tyrosine kinase domains is dispensable for enzymatic and transforming activities." EMBO J **8**(7): 2029-37.

Travers, P., M. Walport, et al. (2008). Janeway's Immunobiology, Garland Science.

Trüper, H. G. and L. De' Clari (1997). "Taxonomic note: Necessary correction of specific epithets formed as substantives (nouns) "in Apposition"." J. Syst. Bacteriol.(47): 908-909.

van der Geer, P. and T. Hunter (1993). "Mutation of Tyr697, a GRB2-binding site, and Tyr721, a PI 3-kinase binding site, abrogates signal transduction by the murine CSF-1 receptor expressed in Rat-2 fibroblasts." EMBO J **12**(13): 5161-72.

Visvader, J. and I. M. Verma (1989). "Differential transcription of exon 1 of the human c-fms gene in placental trophoblasts and monocytes." Mol Cell Biol **9**(3): 1336-41.

Walter, M., I. S. Lucet, et al. (2007). "The 2.7 A crystal structure of the autoinhibited human c-Fms kinase domain." J Mol Biol **367**(3): 839-47.

Wang, T., P. C. Hanington, et al. (2008). "Two macrophage colony-stimulating factor genes exist in fish that differ in gene organization and are differentially expressed." J Immunol **181**(5): 3310-22.

Wang, Z. E., G. M. Myles, et al. (1993). "Identification of the ligand-binding regions in the macrophage colony-stimulating factor receptor extracellular domain." Mol Cell Biol **13**(9): 5348-59.
Webb, S. E., J. W. Pollard, et al. (1996). "Direct observation and quantification of macrophage chemoattraction to the growth factor CSF-1." J Cell Sci **109 (Pt 4)**: 793-803.
Wei, S., S. Nandi, et al. (2010). "Functional overlap but differential expression of CSF-1 and IL-34 in their CSF-1 receptor-mediated regulation of myeloid cells." J Leukoc Biol.
Wheelan, S. J., D. M. Church, et al. (2001). "Spidey: a tool for mRNA-to-genomic alignments." Genome Res **11**(11): 1952-7.
Whyte, S. K. (2007). "The innate immune response of finfish--a review of current knowledge." Fish Shellfish Immunol **23**(6): 1127-51.
Wiktor-Jedrzejczak, W. and S. Gordon (1996). "Cytokine regulation of the macrophage (M phi) system studied using the colony stimulating factor-1-deficient op/op mouse." Physiol Rev **76**(4): 927-47.
Wilhelmsen, K., S. Burkhalter, et al. (2002). "C-Cbl binds the CSF-1 receptor at tyrosine 973, a novel phosphorylation site in the receptor's carboxy-terminus." Oncogene **21**(7): 1079-89.
Williams, H., S. Brenner, et al. (2002). "Identification and analysis of additional copies of the platelet-derived growth factor receptor and colony stimulating factor 1 receptor genes in fugu." Gene **295**(2): 255-64.
Yuzawa, S., Y. Opatowsky, et al. (2007). "Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor." Cell **130**(2): 323-34.
Zapata, A., B. Diez, et al. (2006). "Ontogeny of the immune system of fish." Fish Shellfish Immunol **20**(2): 126-36.
Zuasti, A. and C. Ferrer (1988). "Granulopoiesis in the head kidney of Sparus auratus." Arch Histol Cytol **51**(5): 425-31.

Annex I

```
- 209  AGACTGTAAACAAGCAGAAGGGAGAAGAAGGGCAGAGGCAAACAGAAAGGAAGAAGGGACTCTTGAGTCAGCTCTGGATGC

- 128  TGGAAACATCAAACCCAGGGAAGTTTTCGTACCCTTCTGTTAGGCTGAAGACATCTACAGATCAAAAGAAGAGAGTTTTCG

 - 47  TAGCACACACATCAACATTCAACCTGCAGGTTTGAAGTTGATCCAGG

        M  Q  S  Y  L  T  L  L  M  G  I  V  A  S  A  A  S  A  E  W  R  R  P  V  I  K  F
    1   ATGCAGTCCTACCTCACTCTGCTGATGGGGATTGTGGCCTCTGCGGCTTCAGCGGAATGGAGGCGTCCGGTAATCAAGTTC

        N  S  Q  V  V  G  S  S  Y  V  V  V  E  P  G  T  S  L  D  L  R  C  E  G  D  G  P
   82   AACTCTCAGGTGGTAGGGAGTTCTTACGTGGTGGTCGAACCTGGGACCTCTCTTGATCTGAGGTGTGAGGGGGACGGGCCT

        V  N  W  Q  T  R  L  A  K  H  R  R  F  V  S  K  G  N  G  N  V  R  T  L  K  V  E
  163   GTAAACTGGCAAACGAGGCTAGCCAAACACAGGCGCTTTGTGTCCAAGGGCAATGGGAACGTCCGAACCTTAAAGGTGGAA

        R  P  S  A  E  F  T  G  T  Y  K  C  F  Y  T  S  D  P  Q  H  R  E  L  A  S  S  V
  244   CGTCCCTCTGCAGAATTCACTGGAACATACAAGTGTTTTTACACCTCCGATCCGCAGCACCGTGAACTGGCCTCCTCTGTG

        H  V  Y  V  K  D  P  N  R  V  F  W  T  S  S  T  S  L  R  V  V  R  K  E  G  E  D
  325   CATGTGTATGTAAAAGATCCAAACCGCGTTTTCTGGACCAGCAGCACATCCTTGCGGGTTGTGAGGAAGGAAGGTGAAGAC

        Y  L  L  P  C  L  L  T  D  Q  S  A  T  D  M  G  L  R  M  D  N  G  T  S  V  P  P
  406   TACCTGCTGCCCTGCCTGCTGACCGACCAATCAGCCACAGACATGGGCCTCCGCATGGACAACGGCACCAGCGTGCCGCCA

        G  M  N  Y  T  V  N  K  H  R  G  I  L  I  H  S  L  H  P  S  F  N  A  D  Y  V  C
  487   GGGATGAACTACACGGTTAACAAGCACCGTGGTATTCTCATCCACAGCCTCCATCCCAGCTTCAATGCTGACTATGTCTGC

        T  A  K  V  K  G  V  E  K  T  S  K  A  F  S  I  N  V  I  Q  K  L  R  F  P  P  Y
  568   ACAGCCAAGGTCAAGGGAGTGGAGAAAACCTCCAAGGCCTTTTCCATCAACGTCATTCAGAAGCTTCGTTTCCCTCCGTAT

        V  F  L  E  T  D  E  Y  V  R  I  V  G  E  E  L  K  I  R  C  T  T  H  N  P  N  F
  649   GTCTTTTTGGAGACGGATGAGTATGTGCGCATTGTTGGGGAGGAACTCAAGATTCGTTGCACAACACACAACCCCAACTTC

        N  Y  N  V  T  W  K  Y  T  T  K  S  K  V  T  V  E  E  K  V  R  S  S  G  E  N  R
  730   AACTACAATGTCACCTGGAAATACACCACCAAGTCGAAAGTAACGGTGGAGGAGAAGGTTCGCTCCAGTGGAGAAAATCGC

        L  D  I  Q  S  I  L  T  I  A  A  V  D  L  A  D  T  G  N  I  S  C  I  G  T  N  E
  811   CTGGATATACAGAGCATTCTGACTATCGCTGCTGTGGACCTCGCAGACACAGGAAACATTTCCTGCATAGGCACTAATGAA

        A  G  V  N  S  S  T  T  Y  L  M  V  V  D  K  P  Y  I  R  L  L  P  Q  L  S  P  K
  892   GCAGGGGTGAACAGTTCAACAACATACCTGATGGTTGTAGACAAGCCCTATATCAGACTGTTGCCCCAGCTGTCCCCAAAA

        L  A  H  Q  G  L  S  V  E  V  N  E  G  E  D  L  E  L  S  V  L  I  E  A  Y  P  H
  973   CTGGCCCATCAGGGTCTTTCGGTCGAAGTGAACGAGGGAGAAGATCTGGAGCTCAGCGTGCTCATCGAAGCGTACCCCCAC

        I  T  E  H  R  W  H  T  P  T  S  P  N  T  S  T  Q  E  H  K  L  I  R  Y  N  N  R
 1054   ATCACAGAGCACAGATGGCACACCCCAACATCTCCCAACACATCCACACAGGAGCACAAACTCATCAGATACAACAACAGA

        Y  H  A  S  L  Q  L  K  R  M  N  A  Q  E  Q  G  E  Y  T  F  Y  A  R  S  D  L  A
 1135   TACCATGCTAGTCTACAGCTGAAGAGAATGAATGCACAGGAGCAGGGGGAATACACCTTCTATGCCAGGAGTGACTTGGCC

        N  A  S  I  T  F  Q  V  Q  M  Y  Q  R  P  V  A  V  V  R  W  E  N  V  T  T  L  T
 1216   AATGCATCCATCACATTTCAAGTCCAAATGTATCAGAGACCTGTTGCTGTGGTGAGATGGGAAAACGTAACCACACTCACT

        C  T  S  Y  G  Y  P  A  P  R  I  I  W  Y  Q  C  F  G  I  R  P  T  C  N  E  N  H
 1297   TGCACCTCATATGGCTATCCTGCTCCCAGAATCATCTGGTATCAGTGTTTTGGGATACGACCTACGTGCAATGAAAACCAC

        T  G  L  Q  M  A  I  P  L  Q  A  P  T  V  E  V  Q  R  E  E  Y  G  A  V  E  V  E
 1378   ACGGGGCTGCAGATGGCGATCCCTCTCCAGGCTCCTACAGTGGAGGTCCAGAGGGAGGAGTACGGGGCTGTGGAGGTGGAG

        S  V  L  T  V  G  P  S  S  R  R  M  T  V  E  C  V  A  F  N  L  V  G  V  S  R  D
 1459   AGCGTTCTCACTGTGGGGCCGTCCAGCAGGAGGATGACGGTGGAGTGTGTGGCCTTCAACCTCGTCGGTGTCAGCAGAGAC
```

```
      T  F  A  M  E  V  S  D  K  L  F  T  S  T  L  T  G  A  A  G  I  L  A  I  L  L  V
1540  ACTTTTGCCATGGAGGTTTCTGACAAGCTCTTCACTTCCACTTTGACTGGAGCAGCAGGCATTCTGGCCATCCTCCTGGTG

      L  L  V  F  L  F  Y  K  Y  K  Q  K  P  R  Y  E  I  R  W  K  I  I  E  A  R  D  G
1621  CTGCTGGTTTTTCTGTTTTACAAATATAAGCAGAAACCCAGGTATGAGATCCGTTGGAAGATCATTGAAGCAAGAGATGGA

      N  N  Y  T  F  I  D  P  T  Q  L  P  Y  N  E  K  W  E  F  P  R  D  K  L  K  L  G
1702  AACAACTACACCTTCATTGACCCCACTCAGCTGCCTTACAATGAGAAGTGGGAGTTCCCAAGAGACAAGCTGAAGCTAGGG

      K  I  L  G  A  G  A  F  G  K  V  V  E  A  T  A  F  G  L  G  K  E  K  D  N  A  M
1783  AAGATCCTGGGTGCAGGAGCTTTCGGAAAGGTTGTTGAGGCTACAGCCTTTGGTCTGGGAAAGGAGAAGGATAATGCGATG

      R  V  A  V  K  M  L  K  A  S  A  H  S  D  E  R  E  A  L  M  S  E  L  K  I  L  S
1864  CGTGTGGCTGTGAAAATGTTGAAAGCCAGTGCCCATTCAGATGAGAGGGAAGCTCTGATGTCTGAACTGAAGATCCTGAGC

      H  L  G  H  H  K  N  I  V  N  L  L  G  A  C  T  Y  G  G  P  V  L  V  I  T  E  Y
1945  CACCTGGGACACCACAAGAACATTGTCAATCTACTGGGAGCCTGCACCTATGGAGGACCAGTGCTTGTGATCACAGAGTAT

      C  S  L  G  D  L  L  N  F  L  R  Q  K  A  E  T  F  V  N  F  V  M  N  I  P  D  I
2026  TGCAGCCTCGGCGACCTCCTGAACTTCCTTCGCCAGAAGGCAGAGACGTTTGTGAACTTTGTTATGAACATTCCTGATATT

      M  E  N  S  N  D  Y  K  N  I  C  N  Q  K  Q  F  I  R  S  D  S  G  I  S  S  T  S
2107  ATGGAGAACTCGAATGATTACAAGAATATCTGCAATCAGAAACAGTTCATTAGAAGTGACAGTGGGATCTCCAGTACATCC

      S  S  S  Y  L  E  M  R  P  S  Q  L  P  N  M  E  L  S  Q  D  S  V  C  E  E  T  V
2188  TCAAGCAGTTACTTGGAGATGAGACCCAGCCAGCTGCCAAATATGGAATTATCTCAAGACTCTGTGTGTGAGGAGACTGTT

      I  D  W  P  L  D  I  D  D  L  L  R  F  S  F  Q  V  A  Q  G  L  D  F  L  A  A  K
2269  ATTGACTGGCCGTTGGACATTGATGACTTGCTGAGGTTTTCATTTCAAGTGGCTCAGGGCCTTGACTTTCTGGCTGCCAAA

      N  C  I  H  R  D  V  A  A  R  N  V  L  L  T  D  R  R  V  A  K  I  C  D  F  G  L
2350  AATTGTATTCACAGAGACGTTGCTGCTAGGAATGTTCTATTGACTGACCGCAGAGTGGCCAAGATTTGTGACTTTGGTCTG

      A  R  D  I  M  N  D  S  N  Y  V  V  K  G  N  A  R  L  P  V  K  W  M  A  P  E  S
2431  GCACGTGACATCATGAATGACTCCAACTACGTGGTGAAGGGCAATGCACGTCTGCCAGTGAAGTGGATGGCTCCAGAGAGC

      I  F  D  C  V  Y  T  V  Q  S  D  V  W  S  Y  G  I  L  L  W  E  I  F  S  L  G  K
2512  ATCTTCGACTGTGTCTACACCGTCCAGAGTGACGTCTGGTCCTACGGCATCCTCCTGTGGGAGATCTTCTCTTTAGGCAAG

      S  P  Y  P  S  M  A  V  D  S  R  F  Y  K  M  V  K  R  G  Y  Q  M  S  Q  P  D  F
2593  AGCCCCTACCCCAGCATGGCTGTGGACTCCAGGTTCTACAAGATGGTGAAGCGTGGCTACCAGATGTCCCAACCAGACTTT

      A  P  P  E  I  Y  M  I  M  K  M  C  W  N  L  E  P  T  E  R  P  T  F  N  K  I  T
2674  GCCCCACCTGAGATCTACATGATCATGAAGATGTGCTGGAATCTGGAGCCTACAGAGCGTCCAACATTTAACAAGATTACA

      H  L  I  E  R  L  L  G  D  Q  P  G  Q  E  Q  L  I  Y  Q  N  V  Q  Q  Q  V  T  E
2755  CATTTGATAGAAAGACTACTTGGGGACCAACCCGGGCAGGAACAGCTAATCTACCAGAATGTGCAGCAGCAGGTCACAGAG

      G  E  V  C  D  E  P  K  C  C  D  G  P  C  D  Q  S  C  D  H  E  E  E  E  Q  P  L
2836  GGTGAAGTGTGTGATGAGCCCAAGTGCTGCGACGGCCCCTGTGACCAGTCTTGTGACCACGAGGAAGAGGAGCAGCCTCTG

      M  K  T  N  N  Y  Q  F  C
2917  ATGAAGACCAACAACTACCAGTTTTGTTGA

2947  AAGTCCTAAAAGCCAATCAGTCAGCAAACCAATCACTCATCAGCCAGCATCAGGAAAGTGGGATTAGCTGTAGTCCGTGGC

3028  CAGAGTACCACAGCTTGTTCTAACCGCCGGGACAGTTGCAGAGAGTCGCCTTCACAGCTATTTTTTGCCTGTAGTGTCTGC

3109  TCGCCAGACAGACGAGCAGCTGCAGCAAGCATGTGACTTGCTTCTTTGCAGACCATGACTAGAAATTGCCACGTCTAATAC

3190  CAACTGCAGACATGAAAGGAGGAATCATGTCAACAAACCACACAAGCATCCAAAGCTGTACCTTATCCTAAATGTTACTCC

3271  CTCACGGACTGTTTATAACTCCTTCTATTATGTATTTGAGGTATAATCTATTACAGTCTATCCAATGGAGAATCTGTCTGT

3352  CCTCTCCCCCGAGAGTGTCCTTAATTATATGTTGTGCACAATGTTTGTTACTATTTTGTATCACTGCCCCAGGAGAGTGGA

3433  TGCCTTGTTTTTATGGCGACTTCAGCCACTTTTGTCAGCAGGTTCACTTCCTTATCATCATTTTTATTGCATGTAAAGTGA
```

```
3514  AGCACAAACAATGAATCCAAAACTATGCAAGCTCCACATCTTTATGTGCGGACATGACAAATTAGGTTTCATTCTCTGATC
3595  ACCTTAATTGAACACCTAGACAGTGCAAGTGTTTTTACCCTGGTTTAAACATGCTCAATTGACCGTGGAGAAGGCAAAGAA
3676  CGCCCACTTTGTTTTCCTTCTCTATGCTTTGGTCATAGCTGTATTTCTTGTGACACACTTACAGTAGCTGTCAGTTGATGG
3757  CAACATGTGATTGTTCCACTTCAATCTAATTACTTAACCAGTGTTCAGTATGATATAGGAGCAGGGGAATACAACTGAACT
3838  GAGCAGGGAAATCTGTGGATCCGTTATGGATCAGATGAACAATGGTTCATGTTTGACAATGTGAAAGTGTTTTTCTTTTTT
3919  TGGCACAACTTACTCTTCTCATTTCCTGTAGAATTAGACAATGTCTATATGTAGCAGTCAGCCTACACTGCCAACTGTATG
4000  AAGATTTTAGAATAACATATAAGCTATAGTGAGATTTACAGACTGCAAATTTGTTTGATTCAGGCATTTCAAATGTGCTTG
4081  CATCTGAATGAAACACGAGTGAGATGTCTAATAATGTATATATCTCATGATGTAAATCAAGAATTATCTTTATTTTCGCTG
4162  AAATGATGTCGAATAGTAGAACATCAGCATTGCATATATGTGTATGGTTTTTTTGTGCAAATGAGACAATCAATGCAAAAG
4243  CTGCCAATTAAATCACTGAAAAACTTCAGAGGCGTTGAATGAAGCTTTACCATTAAAAACTAAAAAAGGTTTATATGTAAT
4324  ACTAAAAAAAAAAAAAAAAAAAAAA
```

Figure A.1. Nucleotide and deduced amino acid sequences of sea bass CSF1R1/a. The polyadenylation signal is denoted by gray shading.

```
-28  ACTTCAACATTGACCAGAAAGGAGAGAG

      M  K  L  Y  T  L  L  L  S  I  A  L  S  C  G  C  S  A  K  D  P  P  G  P  P  S  I
  1  ATGAAGCTTTATACTCTACTGCTGTCGATCGCCCTCAGCTGCGGCTGCTCTGCCAAAGATCCTCCTGGTCCTCCATCGATC

      H  L  N  S  D  F  L  P  N  Q  T  E  V  V  L  T  A  G  A  T  F  N  L  S  C  H  G
 82  CATCTGAACTCTGACTTCCTGCCGAACCAAACCGAGGTGGTCCTGACCGCTGGCGCCACCTTCAACCTCAGTTGTCATGGT

      N  G  M  V  R  W  S  S  T  A  F  R  L  L  Y  E  D  K  L  K  D  L  V  E  V  R  R
163  AACGGTATGGTCCGTTGGTCCAGTACTGCGTTCCGTCTGCTGTACGAGGACAAACTGAAGGATCTGGTGGAGGTACGGAGG

      A  D  P  R  H  T  G  T  Y  R  C  G  Y  T  N  Q  S  L  E  H  L  D  T  W  I  H  L
244  GCGGACCCCAGACACACCGGAACGTACCGCTGTGGGTACACCAACCAGAGCCTGGAACACCTGGACACCTGGATCCACCTG

      Y  V  K  D  P  A  D  P  S  S  V  F  V  T  P  R  S  S  I  P  A  L  K  E  G  Q  D
325  TATGTCAAAGACCCGGCTGACCCCTCCAGTGTGTTCGTCACTCCTCGCAGCAGCATCCCTGCCCTCAAGGAAGGCCAGGAC

      F  L  F  R  C  L  L  T  D  P  S  V  T  N  L  T  L  Q  S  E  D  R  I  G  G  R  G
406  TTCTTGTTCAGGTGTCTACTGACCGACCCATCAGTCACAAACCTCACCCTCCAATCAGAGGACAGAATTGGGGGGAGGGGG

      Q  D  L  P  Q  G  M  I  V  T  I  D  P  Q  K  G  A  L  I  R  D  L  Q  M  S  F  K
487  CAGGACCTGCCCCAGGGCATGATTGTGACCATTGACCCCCAGAAAGGAGCCCTGATCAGAGACTTGCAGATGTCATTTAAA

      G  H  Y  V  C  S  G  W  K  D  G  R  Q  F  R  S  R  P  G  N  L  V  V  V  R  R  L
568  GGACACTACGTTTGTTCAGGCTGGAAGGACGGACGGCAGTTCAGATCCAGACCTGGCAACCTGGTGGTGGTCCGCAGGCTG

      L  E  P  P  S  L  S  V  S  Q  G  E  L  V  R  L  E  G  E  Q  F  E  V  T  C  V  T
649  CTTGAGCCTCCCTCCCTGTCTGTCAGTCAGGGTGAACTCGTTCGTCTGGAAGGAGAGCAGTTTGAAGTCACCTGTGTGACC

      S  N  P  S  H  L  Y  N  V  T  W  T  G  P  N  S  E  R  L  K  V  R  V  S  Q  D  Y
730  AGTAACCCTTCCCATTTGTACAACGTCACCTGGACGGGCCCAAACTCAGAGAGACTGAAGGTCCGTGTCAGCCAAGACTAT

      I  K  K  H  V  F  K  N  S  T  V  R  V  S  A  V  N  L  T  H  S  G  I  Y  T  C  T
811  ATCAAAAAACACGTGTTCAAAAACAGCACTGTGAGAGTGTCTGCTGTCAACCTGACTCACAGTGGAATCTACACCTGTACT

      A  V  N  E  A  G  V  A  M  A  T  T  H  L  R  V  V  D  A  P  F  L  R  I  Y  L  Q
892  GCTGTCAATGAGGCTGGGGTCGCCATGGCAACCACACATCTCAGAGTTGTGGACGCTCCCTTCCTGAGGATTTACCTGCAG

      H  M  P  H  A  N  A  N  T  E  T  I  Q  A  K  G  D  L  V  V  N  L  S  S  M  L  L
973  CATATGCCGCATGCTAACGCTAACACCGAGACCATACAGGCCAAGGGTGACCTGGTTGTTAACCTTAGTAGCATGCTATTG

      E  E  Q  G  E  L  E  A  N  I  S  S  G  S  L  E  V  N  R  E  L  I  V  N  V  S  T
1054  GAGGAGCAAGGAGAGCTAGAAGCTAACATTAGCAGCGGTTCACTGGAAGTGAACAGGGAATTAATTGTTAACGTCAGCACC

      K  R  L  D  V  N  G  A  S  S  A  N  V  S  G  S  T  V  E  V  Y  E  G  R  D  V  M
1135  AAGAGACTGGATGTGAATGGAGCTAGCAGTGCTAACGTTAGTGGCAGCACAGTGGAGGTGTATGAAGGCCGAGATGTGATG

      L  T  F  V  I  E  S  Y  P  P  I  R  K  Q  H  W  T  T  P  A  K  I  N  N  T  V  Y
1216  CTGACCTTTGTGATAGAGTCATACCCTCCAATCAGGAAGCAGCACTGGACAACACCAGCAAAGATCAACAACACGGTGTAT

      E  E  S  Y  T  A  N  G  H  R  A  E  A  R  F  V  L  R  R  V  R  Q  E  D  R  G  R
1297  GAGGAGAGCTACACTGCTAATGGCCACAGGGCAGAGGCGCGCTTCGTGCTGCGGCGGGTTCGCCAGGAGGATCGCGGTCGT

      Y  S  F  H  F  S  N  S  F  F  S  G  L  H  N  I  D  L  Q  I  Y  R  S  A  E  E  P
1378  TACTCGTTCCACTTCTCTAACTCGTTCTTCAGCGGTTTGCATAACATCGACCTCCAAATTTACCGTTCTGCAGAAGAACCT

      Q  T  A  L  I  G  A  L  S  A  A  A  I  L  L  L  L  L  L  V  V  F  Y  K  W  R  Q
1459  CAGACGGCTCTGATTGGAGCTCTGAGTGCCGCTGCCATCCTCCTTCTGCTCTTATTGGTCGTCTTTTATAAGTGGAGACAG

      K  P  R  Y  E  I  R  W  K  I  I  E  S  T  D  G  N  N  Y  T  F  I  D  P  T  Q  L
1540  AAACCCAGATATGAGATTCGCTGGAAGATCATCGAGAGCACCGACGGAAACAACTACACCTTCATTGACCCCACCCAGCTG

      P  Y  N  Y  K  W  E  F  P  R  D  K  L  R  L  G  A  V  L  G  S  G  A  F  G  K  V
1621  CCGTACAACTACAAGTGGGAGTTTCCTCGAGACAAACTCCGCCTCGGTGCTGTTTTGGGTTCGGGGGCGTTCGGGAAGGTT

      V  E  A  T  A  Y  G  L  R  T  D  N  D  I  T  R  V  A  V  K  M  L  K  P  S  A  H
1702  GTCGAGGCGACGGCGTATGGTCTGAGAACCGACAATGACATCACACGAGTTGCTGTCAAGATGCTCAAACCGAGCGCTCAC
```

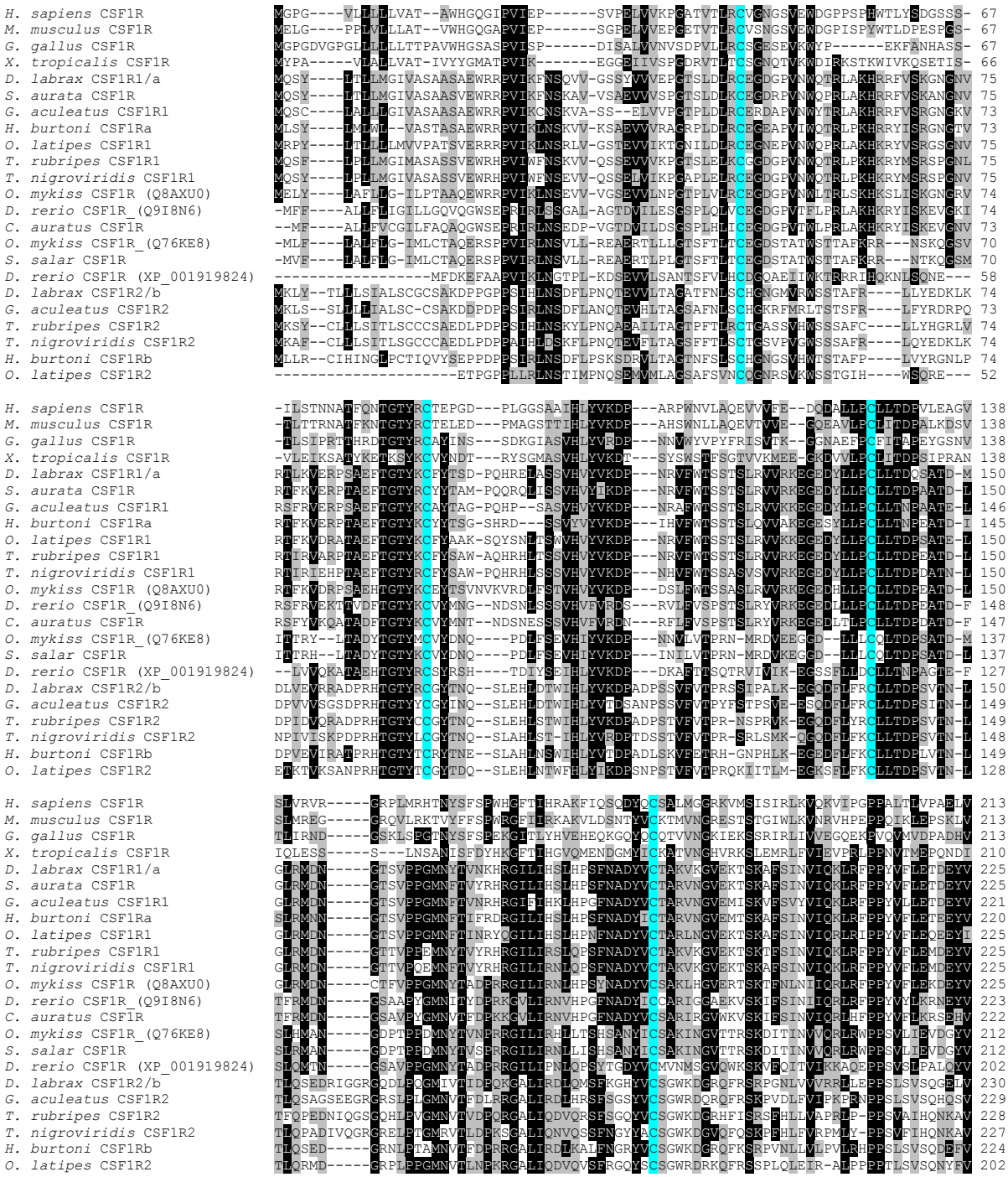
```
       S  E  E  R  E  A  L  M  T  E  L  K  I  L  S  H  L  G  Y  H  D  N  I  V  N  L  L
1783   TCTGAGGAACGGGAAGCTCTGATGACAGAGCTGAAGATCCTCAGCCATCTTGGGTACCATGACAACATCGTCAACCTGCTG

       G  A  C  T  Q  G  G  P  M  L  M  I  T  E  Y  C  S  H  G  D  L  L  N  F  L  R  A
1864   GGCGCATGCACTCAGGGAGGTCCTATGTTGATGATCACAGAGTACTGTAGCCACGGCGACCTGCTCAACTTCCTGCGGGCT

       H  A  Q  D  F  M  A  S  I  L  S  V  D  E  V  T  G  E  A  F  Y  K  N  A  A  G  Q
1945   CACGCTCAGGACTTCATGGCATCCATTTTGAGCGTGGATGAAGTGACAGGAGAGGCTTTCTATAAGAACGCGGCCGGCCAG

       H  A  R  L  R  S  D  S  G  I  S  C  C  S  E  Y  Q  E  M  H  P  V  L  S  P  G  Q
2026   CACGCCAGACTCAGGAGTGACAGTGGGATCTCCTGCTGCTCAGAGTATCAGGAGATGCATCCAGTTCTGAGTCCAGGACAA

       T  H  L  G  V  Q  T  D  S  L  S  V  G  D  L  M  R  F  S  L  Q  V  A  Q  G  L  D
2107   ACACACCTGGGTGTGCAGACAGACAGTCTGTCTGTTGGTGACCTCATGAGGTTTTCCCTCCAGGTGGCTCAGGGTCTGGAC

       F  L  S  S  R  N  C  I  H  R  D  V  A  A  R  N  V  L  L  T  D  H  R  V  A  K  I
2188   TTCCTGTCCTCCAGGAATTGTATCCACAGAGACGTAGCGGCGAGGAACGTCCTGCTGACTGATCATCGTGTTGCAAAGATC

       C  D  F  G  L  A  R  D  I  R  N  D  D  S  Y  I  V  Q  G  N  A  R  L  P  V  K  W
2269   TGTGACTTCGGTTTGGCCCGAGACATCCGTAACGACGACAGCTACATCGTTCAGGGAAATGCTCGTCTTCCTGTGAAGTGG

       M  A  P  E  S  I  F  Q  C  V  Y  T  V  Q  S  D  V  W  S  Y  G  V  L  L  W  E  I
2350   ATGGCTCCAGAGAGTATCTTCCAGTGCGTGTACACCGTCCAGAGCGACGTCTGGTCCTATGGAGTCTTACTGTGGGAGATC

       F  S  L  G  K  S  P  Y  P  N  V  A  V  D  T  N  F  Y  K  M  I  K  D  G  R  H  M
2431   TTCTCTCTGGGTAAGAGTCCCTATCCAAATGTTGCCGTGGATACCAACTTCTACAAGATGATCAAAGATGGCCGCCACATG

       N  Q  P  D  F  A  P  A  E  M  Y  Q  L  M  T  L  C  W  S  L  E  P  T  D  R  P  T
2512   AATCAGCCAGACTTTGCTCCGGCAGAGATGTATCAGCTGATGACGCTCTGCTGGAGTTTGGAGCCCACAGACAGACCCACC

       F  K  M  I  G  Q  L  I  N  R  L  L  P  S  T  N  D  T  S  P  H  H  S  E  Q  Y  R
2593   TTTAAAATGATTGGTCAGCTCATTAACAGGCTCCTCCCCTCCACCAATGACACATCACCACATCACAGTGAGCAGTACAGG

       N  I  T  E  C  R  E  E  E  E  E  E  E  E  E  Q  E  E  E  E  E  E  E  E  Q  E  A
2674   AACATCACAGAGTGCAGAGAGGAAGAGGAGGAGGAAGAGGAGGAGCAAGAAGAGGAAGAGGAGGAAGAGGAGCAGGAGGCC

       R  A  R  D  T  L  K  R  G  G  G  A  A  G  G  G  E  R  S  P
2755   AGAGCAAGAGACACACTGAAGAGAGGAGGAGGAGCAGCAGGAGGAGGAGAGAGATCTCCATGA

2818   GGACATTGATGAAGATGAGCAGGAGCCGATGATGAAGAACATCTACCAGCTGTCCTGATCATCCGTGTTAATGAACAGATT

2899   CAGAGTCTGCTCTCTGTCATGTGATTAGTTTACGATACGTGAAAGATCGCAACGAACCACAACGTACGTTACGATTATATA

2980   ACTCGCTGATTTCCCCTAACTGACTTTTCAGATTAAAAGACCTCAGAGAGTTTAAAGTGCATCATCTTTATAACAGACTCT

3061   GGATCAGTTTCATTTTAACGAACAACACGTGATTTTAATCTTCTTAATCAATTAATCTGATTGATTAGATGTGTGATTTTT

3142   CATTGTTTGTGTCTTTAACAGTTTTATTATCAGAATGTAAAAATAAATTCAGACGAACGGAAAAAAAAAAAAAA
```

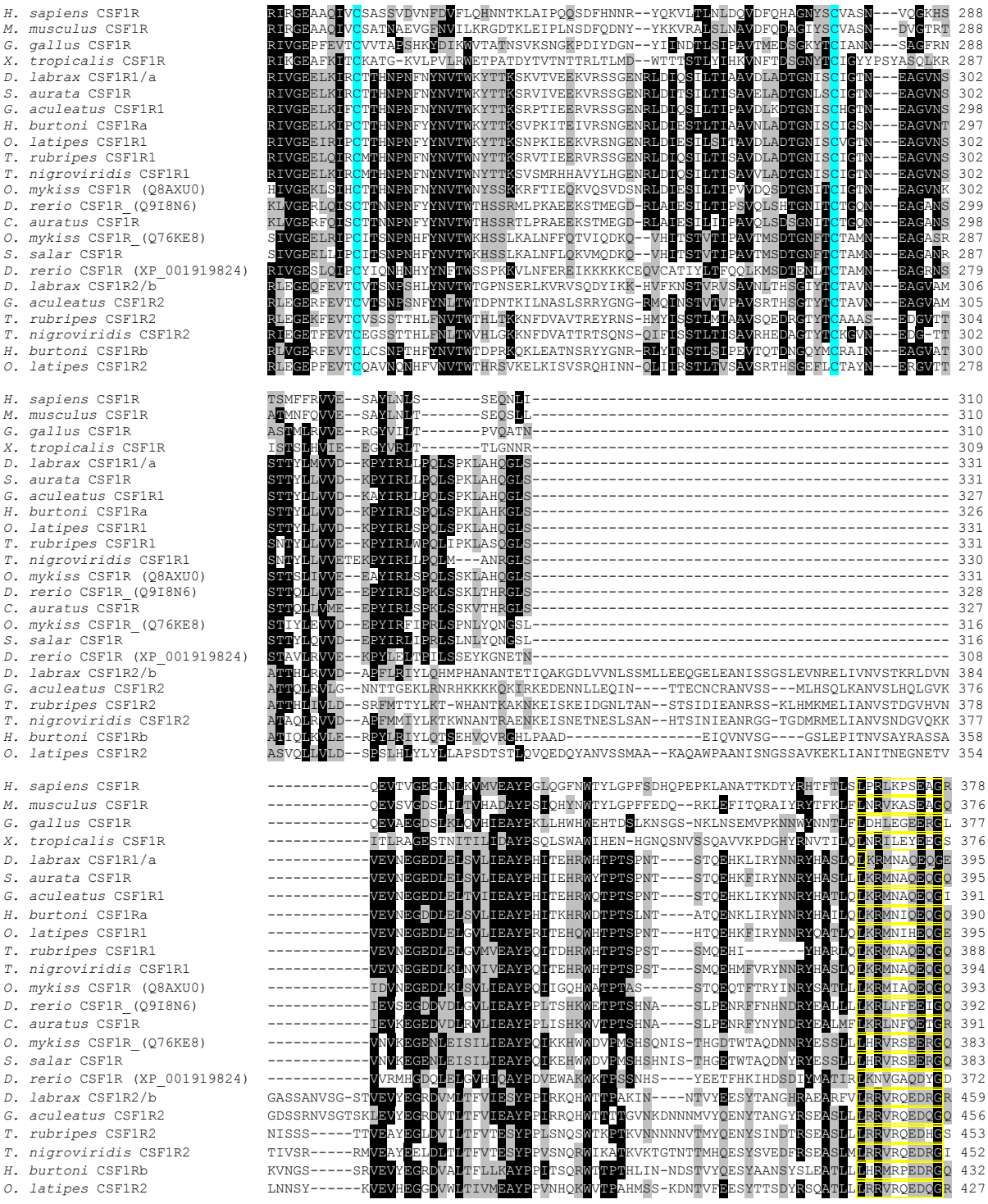
Figure A.2. Nucleotide and deduced amino acid sequences of sea bass CSF1R2/b. The polyadenylation signal is denoted by gray shading.

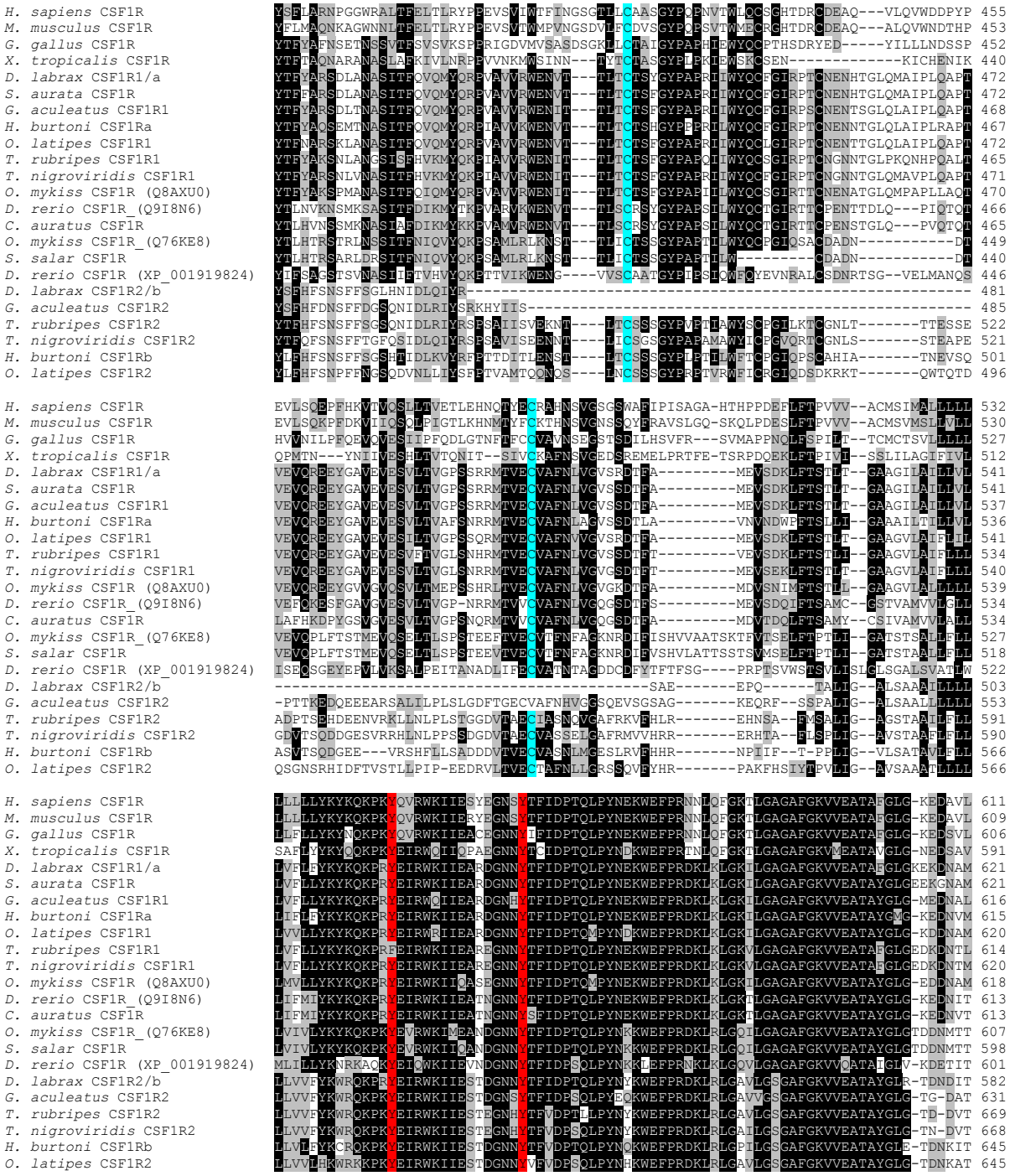
```
-25  CCAACATTGACCAGAAAGGAGAGAG

       M  K  L  Y  T  L  L  L  S  I  A  L  S  C  G  C  S  A  K  D  P  P  G  P  P  S  I
  1  ATGAAGCTTTATACTCTACTGCTGTCGATCGCCCTCAGCTGCGGCTGCTCCGCCAAAGATCCTCCTGGTCCTCCATCGATC

       H  L  N  S  D  F  L  P  N  Q  T  E  V  V  L  T  A  G  A  T  F  N  L  S  C  H  G
 82  CATCTGAACTCTGACTTCCTGCCGAACCAAACCGAGGTGGTCCTGACCGCTGGCGCCACCTTCAACCTCAGTTGTCATGGT

       N  G  M  V  R  W  S  S  T  A  F  R  L  L  Y  E  D  K  L  K  D  L  V  E  V  R  R
163  AACGGTATGGTCCGTTGGTCCAGTACTGCGTTCCGTCTGCTGTACGAGGACAAACTGAAGGATCTGGTGGAGGTACGGAGG

       A  D  P  R  H  T  G  T  Y  R  C  G  Y  T  N  Q  S  L  E  H  L  D  T  W  I  H  L
244  GCGGACCCCAGACACACCGGAACGTACCGCTGTGGGTACACCAACCAGAGCCTGGAACACCTGGACACCTGGATCCACCTG

       Y  V  K  D  P  A  D  P  S  S  V  F  V  T  P  R  S  S  I  P  A  L  K  E  G  Q  D
325  TATGTCAAAGACCCGGCTGACCCCTCCAGTGTGTTCGTCACTCCTCGCAGCAGCATCCCTGCCCTCAAGGAAGGCCAGGAC

       F  L  F  R  C  L  L  T  D  P  S  V  T  N  L  T  L  Q  S  E  D  R  I  G  G  R  G
406  TTCTTGTTCAGGTGTCTACTGACCGACCCATCAGTCACAAACCTCACCCTCCAATCAGAGGACAGAATTGGGGGGAGGGGG

       Q  D  L  P  Q  G  M  I  V  T  I  D  P  Q  K  G  A  L  I  R  D  L  Q  M  S  F  K
487  CAGGACCTGCCCCAGGGCATGATTGTGACCATTGACCCCCAGAAAGGAGCCCTGATCAGAGACTTGCAGATGTCATTTAAA

       G  H  Y  V  C  S  G  W  K  D  G  R  Q  F  R  S  R  P  G  N  L  V  V  V  R  R  L
568  GGACACTACGTTTGTTCAGGCTGGAAGGACGGACGGCAGTTCAGATCCAGACCTGGCAACCTGGTGGTGGTCCGCAGGCTG

       L  E  P  P  S  L  S  V  S  Q  G  E  L  V  R  L  E  G  E  Q  F  E  V  T  C  V  T
649  CTTGAGCCTCCCTCCCTGTCTGTCAGTCAGGGTGAACTCGTTCGTCTGGAAGGAGAGCAGTTTGAAGTCACCTGTGTGACC

       S  N  P  S  H  L  Y  N  V  T  W  T  G  P  N  S  E  R  L  K  V  R  V  S  Q  D  Y
730  AGTAACCCTTCCCATTTGTACAACGTCACCTGGACGGGCCCAAACTCAGAGAGACTGAAGGTCCGTGTCAGCCAAGACTAT

       I  K  K  H  V  F  K  N  S  T  V  R  V  S  A  V  N  L  T  H  S  G  I  Y  T  C  T
811  ATCAAAAAACACGTGTTCAAAAACAGCACTGTGAGAGTGTCTGCTGTCAACCTGACTCACAGTGGAATCTACACCTGTACT

       A  V  N  E  A  G  V  A  M  A  T  T  H  L  R  V  V  D  A  P  F  L  R  I  Y  L  Q
892  GCTGTCAATGAGGCTGGGGTCGCCATGGCAACCACACATCTCAGAGTTGTGGACGCTCCCTTCCTGAGGATTTACCTGCAG

       H  M  P  H  A  N  A  N  T  K  T  I  Q  A  K  G  D  L  V  V  N  L  S  S  M  L  L
973  CATATGCCGCATGCTAACGCTAACACCAAGACCATACAGGCCAAGGGTGACCTGGTTGTTAACCTTAGTAGCATGCTATTG

       E  E  Q  G  E  L  E  A  N  I  S  S  G  S  L  E  V  N  R  E  L  I  V  N  V  S  T
1054  GAGGAGCAAGGAGAGCTAGAAGCTAACATTAGCAGCGGTTCACTGGAAGTGAACAGGGAATTAATTGTTAACGTCAGCACC

       K  R  L  D  V  N  G  A  S  S  A  N  V  S  G  S  S  T  V  E  V  Y  E  G  R  D  V
1135  AAGAGACTGGATGTGAATGGAGCTAGCAGTGCTAACGTTAGTGGCAGCAGCACAGTGGAGGTGTATGAAGGCCGAGATGTG

       M  L  T  F  V  I  E  S  Y  P  P  I  R  K  Q  H  W  T  T  P  A  K  I  N  N  T  V
1216  ATGCTGACCTTTGTGATAGAGTCATACCCTCCAATCAGGAAGCAGCACTGGACAACACCAGCAAAGATCAACAACACGGTG

       Y  E  E  S  Y  T  A  N  G  H  R  S  E  A  R  F  V  L  R  R  V  R  Q  E  D  R  G
1297  TATGAGGAGAGCTACACTGCTAATGGCCACAGGTCAGAGGCGCGCTTCGTGCTGCGGCGGGTTCGCCAGGAGGATCGCGGT

       R  Y  S  F  H  F  S  N  S  F  F  S  G  L  H  N  I  D  L  Q  I  Y  S  S  A  E  E
1378  CGTTACTCGTTCCACTTCTCTAACTCGTTCTTCAGCGGTTTGCATAACATCGACCTCCAAATTTACAGTTCTGCAGAAGAA

       P  Q  T  A  L  I  G  A  L  S  A  A  A  I  L  L  L  L  L  L  V  V  F  Y  K  W  R
1459  CCTCAGACGGCTCTGATTGGAGCTCTGAGTGCCGCTGCCATCCTCCTTCTGCTCTTATTGGTCGTCTTTTATAAGTGGAGA

       Q  K  P  R  Y  E  I  R  W  K  I  I  E  S  T  D  G  N  N  Y  T  F  I  D  P  T  Q
1540  CAGAAACCCAGATATGAGATTCGCTGGAAGATCATCGAGAGCACCGACGGAAACAACTACACCTTCATTGACCCCACCCAG

       L  P  Y  N  Y  K  W  E  F  P  R  D  K  L  R  L  G  A  V  L  G  S  G  A  F  G  K
1621  CTGCCGTACAACTACAAGTGGGAGTTTCCTCGAGACAAACTCCGCCTCGGTGCTGTTTTGGGTTCGGGGGCGTTCGGGAAG

       V  V  E  A  T  A  Y  G  L  R  T  D  N  D  I  T  R  V  A  V  K  M  L  K  P  S  A
1702  GTTGTCGAGGCGACGGCGTATGGTCTGAGAACCGACAATGACATCACACGAGTTGCTGTCAAGATGCTCAAACCGAGCGCT
```

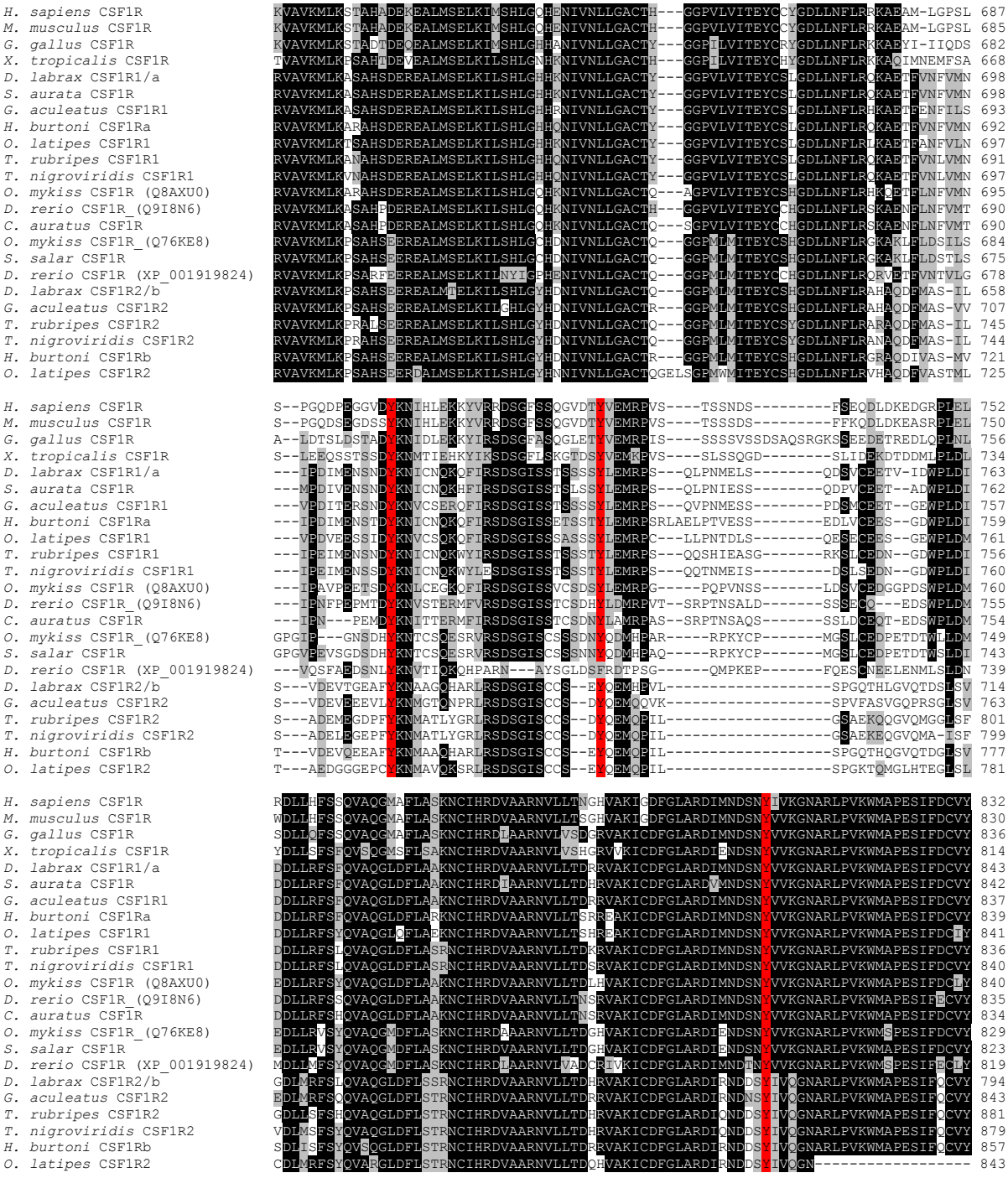
```
       H  S  E  E  R  E  A  L  M  T  E  L  K  I  L  S  H  L  G  Y  H  D  N  I  V  N  L
1783   CACTCTGAGGAACGGGAAGCTCTGATGACAGAGCTGAAGATCCTCAGCCATCTTGGGTACCATGACAACATCGTCAACCTG

       L  G  A  C  T  Q  G  G  P  M  L  M  I  T  E  Y  C  S  H  G  D  L  L  N  F  L  R
1864   CTGGGCGCATGCACTCAGGGAGGTCCTATGTTGATGATCACAGAGTACTGTAGCCACGGCGACCTGCTCAACTTCCTGCGG

       A  H  A  Q  D  F  M  A  S  I  L  S  V  D  E  V  T  G  E  A  F  Y  K  N  V  A  G
1945   GCTCACGCTCAGGACTTCATGGCATCCATTTTGAGCGTGGATGAAGTGACAGGAGAGGCTTTCTATAAGAACGTGGCCGGC

       Q  H  A  R  L  R  S  D  S  G  I  S  C  C  S  E  Y  Q  E  M  H  P  V  L  S  P  G
2026   CAGCACGCCAGACTCAGGAGTGACAGTGGGATCTCCTGCTGCTCAGAGTATCAGGAGATGCATCCAGTTCTGAGTCCAGGA

       Q  T  H  L  G  V  Q  T  D  S  L  S  V  G  D  L  M  R  F  S  L  Q  V  A  Q  G  L
2107   CAAACACACCTGGGTGTGCAGACAGACAGTCTGTCTGTTGGTGACCTCATGAGGTTTTCCCTCCAGGTGGCTCAGGGTCTG

       D  F  L  S  S  R  N  C  I  H  R  D  V  A  A  R  N  V  L  L  T  D  H  R  V  A  K
2188   GACTTCCTGTCCTCCAGGAATTGTATCCACAGAGACGTAGCGGCGAGGAACGTCCTGCTGACTGATCATCGTGTTGCAAAG

       I  C  D  F  G  L  A  R  D  I  R  N  D  D  S  Y  I  V  Q  G  N  A  R  L  P  V  K
2269   ATCTGTGACTTCGGTTTGGCCCGAGACATCCGTAACGACGACAGCTACATCGTTCAGGGAAATGCTCGTCTTCCTGTGAAG

       W  M  A  P  E  S  I  F  Q  C  V  Y  T  V  Q  S  D  V  W  S  Y  G  V  L  L  W  E
2350   TGGATGGCTCCAGAGAGTATCTTCCAGTGCGTGTACACCGTCCAGAGCGACGTCTGGTCCTATGGAGTCTTACTGTGGGAG

       I  F  S  L  G  K  S  P  Y  P  N  V  A  V  D  T  N  F  Y  K  M  I  K  D  G  R  H
2431   ATCTTCTCTCTGGGTAAGAGTCCCTATCCAAATGTTGCCGTGGATACCAACTTCTACAAGATGATCAAAGATGGCCGCCAC

       M  N  Q  P  D  F  A  P  A  E  M  Y  Q  L  M  T  L  C  W  S  L  E  P  T  D  R  P
2512   ATGAATCAGCCAGACTTTGCTCCGGCAGAGATGTATCAGCTGATGACGCTCTGCTGGAGTTTGGAGCCCACAGACAGACCC

       T  F  K  M  I  G  Q  L  I  N  R  L  L  P  S  T  N  D  T  S  P  H  H  S  E  Q  I
2593   ACCTTTAAAATGATTGGTCAGCTCATTAACAGGCTCCTCCCCTCCACCAATGACACATCACCACATCACAGTGAGCAGATC

       S  M  R  T  L  M  K  M  S  R  S  R
2674   TCCATGAGGACATTGATGAAGATGAGCAGGAGCCGATGA

2713   TGAAGAACATCTACCAGCTGTCCTGATCATCCGTGTTAATGAACAGATTCAGAGTCTGCTCTCTGTCATGTGATTAGTTTA

2794   CGATACGTGAAAGATCGCAACGAACCACAACGTACGTTACGATTATATAACTCGCTGATTTCCCCTAACTGACTTTTCAGA

2875   TTAAAAGACCTCAGAGAGTTTAAAGTGCATCATCTTTATAACAGACTCTGGATCAGTTTCATTTTAACGAACAACACGTGA

2956   TTTTAATCTTCTTAATCAATTAATCTGATTGATTAGATGTGTGATTTTTCATTGTTTGTGTTTTTAACAGTTTTATTATCA

3037   GAATGTAAAAATAAATTCAGACGAACGGAAAAAAAA
```

Figure A.3. Nucleotide and deduced amino acid sequences of sea bass CSF1R2/b.2. The polyadenylation signal is denoted by gray shading.


H. sapiens CSF1R
M. musculus CSF1R
G. gallus CSF1R
X. tropicalis CSF1R
D. labrax CSF1R1/a
S. aurata CSF1R
G. aculeatus CSF1R1
H. burtoni CSF1Ra
O. latipes CSF1R1
T. rubripes CSF1R1
T. nigroviridis CSF1R1
O. mykiss CSF1R (Q8AXU0)
D. rerio CSF1R_(Q9I8N6)
C. auratus CSF1R
O. mykiss CSF1R_(Q76KE8)
S. salar CSF1R
D. rerio CSF1R (XP_001919824)
D. labrax CSF1R2/b
G. aculeatus CSF1R2
T. rubripes CSF1R2
T. nigroviridis CSF1R2
H. burtoni CSF1Rb
O. latipes CSF1R2


H. sapiens CSF1R
M. musculus CSF1R
G. gallus CSF1R
X. tropicalis CSF1R
D. labrax CSF1R1/a
S. aurata CSF1R
G. aculeatus CSF1R1
H. burtoni CSF1Ra
O. latipes CSF1R1
T. rubripes CSF1R1
T. nigroviridis CSF1R1
O. mykiss CSF1R (Q8AXU0)
D. rerio CSF1R_(Q9I8N6)
C. auratus CSF1R
O. mykiss CSF1R_(Q76KE8)
S. salar CSF1R
D. rerio CSF1R (XP_001919824)
D. labrax CSF1R2/b
G. aculeatus CSF1R2
T. rubripes CSF1R2
T. nigroviridis CSF1R2
H. burtoni CSF1Rb
O. latipes CSF1R2


H. sapiens CSF1R
M. musculus CSF1R
G. gallus CSF1R
X. tropicalis CSF1R
D. labrax CSF1R1/a
S. aurata CSF1R
G. aculeatus CSF1R1
H. burtoni CSF1Ra
O. latipes CSF1R1
T. rubripes CSF1R1
T. nigroviridis CSF1R1
O. mykiss CSF1R (Q8AXU0)
D. rerio CSF1R_(Q9I8N6)
C. auratus CSF1R
O. mykiss CSF1R_(Q76KE8)
S. salar CSF1R
D. rerio CSF1R (XP_001919824)
D. labrax CSF1R2/b
G. aculeatus CSF1R2
T. rubripes CSF1R2
T. nigroviridis CSF1R2
H. burtoni CSF1Rb
O. latipes CSF1R2


H. sapiens CSF1R
M. musculus CSF1R
G. gallus CSF1R
X. tropicalis CSF1R
D. labrax CSF1R1/a
S. aurata CSF1R
G. aculeatus CSF1R1
H. burtoni CSF1Ra
O. latipes CSF1R1
T. rubripes CSF1R1
T. nigroviridis CSF1R1
O. mykiss CSF1R (Q8AXU0)
D. rerio CSF1R_(Q9I8N6)
C. auratus CSF1R
O. mykiss CSF1R_(Q76KE8)
S. salar CSF1R
D. rerio CSF1R (XP_001919824)
D. labrax CSF1R2/b
G. aculeatus CSF1R2
T. rubripes CSF1R2
T. nigroviridis CSF1R2
H. burtoni CSF1Rb
O. latipes CSF1R2

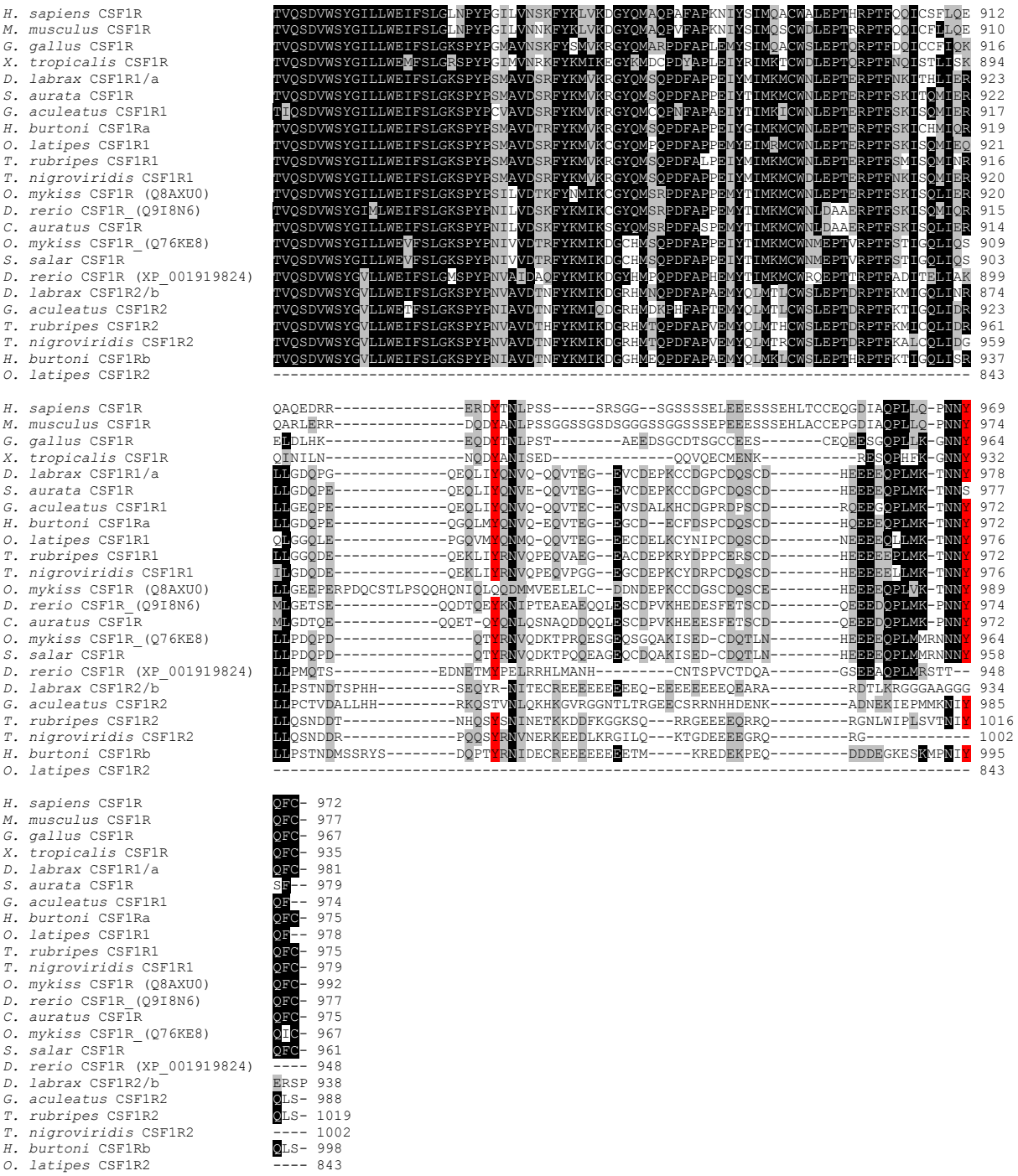

H. sapiens CSF1R
M. musculus CSF1R
G. gallus CSF1R
X. tropicalis CSF1R
D. labrax CSF1R1/a
S. aurata CSF1R
G. aculeatus CSF1R1
H. burtoni CSF1Ra
O. latipes CSF1R1
T. rubripes CSF1R1
T. nigroviridis CSF1R1
O. mykiss CSF1R (Q8AXU0)
D. rerio CSF1R_(Q9I8N6)
C. auratus CSF1R
O. mykiss CSF1R_(Q76KE8)
S. salar CSF1R
D. rerio CSF1R (XP_001919824)
D. labrax CSF1R2/b
G. aculeatus CSF1R2
T. rubripes CSF1R2
T. nigroviridis CSF1R2
H. burtoni CSF1Rb
O. latipes CSF1R2

Figure A.4. Multiple alignment of CSF1R proteins. Amino acid sequences of sea bass, human (*H. sapiens*; P07333), mouse (*M. musculus*; P09581), chicken (*G. gallus*; XP_414597), western clawed frog (*X. tropicalis*; NP_00100818), gilthead seabream (*S.aurata*; CAJ18352), three-spined stickleback (*G. aculeatus*; CSF1R1 ENSGACP00000023794 and CSF1R2 ENSGACP00000027653), medaka (*O. latipes*; CSF1R1 ENSORLP00000006110 and CSF1R2 ENSORLP00000016040), fugu (*T. rubripes*; CSF1R1 P79750 and CSF1R2 Q8UVR8), spotted green pufferfish (*T. nigroviridis*; CSF1R1 ENSTNIP00000017239 and CSF1R2 ENSTNIP00000010425), *Haplochromis burtoni* (CSF1R1 ABD48801 and CSF1R2 ABD48799), rainbow trout (*O. mykiss*, Q8AXU0 and Q76KE8), zebrafish (*D. rerio*; Q9I8N6 and XP_001919824) and atlantic salmon (*Salmo salar*; ACU57063) were used for this alignment. The structural and functional relevant cysteine (C) and tyrosine (Y) residues are shaded in turquoise and red respectively. The yellow box denotes the conserved dimerization motif. The numbers indicate the amino acid positions and the dashes indicate gaps introduced to optimize similarity between sequences. Asterisks (*) denote identical residues and the colon (:) and dot (.) denote chemical similarity between amino acids.