UNIVERSIDADE DO ALGARVE

IDENTIFICATION OF NOVEL MOLECULAR DETERMINANTS OF TISSUE MINERALIZATION IN FISH

Joana Alexandra Teixeira Rosa

Tese para a obtenção do grau de Doutor em Ciências Biomédicas

Trabalho efetuado sob a orientação de:

Professora Doutora Leonor Cancela

Doutor Vincent Laizé

2015

UNIVERSIDADE DO ALGARVE

IDENTIFICATION OF NOVEL MOLECULAR DETERMINANTS OF TISSUE MINERALIZATION IN FISH

Joana Alexandra Teixeira Rosa

Tese para a obtenção do grau de Doutor em Ciências Biomédicas

Trabalho efetuado sob a orientação de:

Professora Doutora Leonor Cancela

Doutor Vincent Laizé

2015

Identification of novel molecular determinants of tissue mineralization in fish

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

_____

À Mãe

…What is essential is invisible to the eye.

Antoine de Saint Exupéry

AGRADECIMENTOS

O culminar de vários anos de trabalho figura hoje sob a forma desta dissertação de doutoramento, cuja realização e conclusão não teria sido possível sem a ajuda de várias (muitas) pessoas que, direta ou indiretamente contribuíram para este final feliz. A todas elas, um enorme Obrigada!

Gostaria de começar por agradecer à Professora Leonor Cancela por me ter acolhido, já lá vão alguns anos, no seu laboratório, e por posteriormente me ter incentivado e dado a oportunidade para realização deste doutoramento. Foi um privilégio poder contar com os seus ensinamentos, e estou-lhe grata por todo o empenho, preocupação e partilha do saber.

Ao Doutor Vincent Laizé, meu co-orientador, a quem o espaço limitado desta secção de agradecimentos jamais chegaria para expressar toda a minha gratidão. Obrigada pela infindável paciência, por toda a motivação e incentivos constantes, pelos conhecimentos transmitidos, por toda a ajuda durante o processo de escrita, mas principalmente por nunca ter duvidado que seria possível chegar até aqui. Muito mais do que um "chefe" ou um líder, um amigo. O meu Muitíssimo Obrigada Vincent! (A, eternamente agradecida, Calimero!)

Aos meus colegas de Laboratório, que ao fim de tantos anos se tornaram nos meus amigos. Vocês foram a razão, mesmo em dias menos bons, da minha alegria e sorrisos no trabalho. Obrigada por me terem aturado! Obrigada pelas amizades saudáveis que contruímos!

Sem qualquer ordem em especial gostaria de agradecer à Natércia, ao Gavaia, ao Daniel e ao Ricardo L. as ajudas prestadas, a disponibilidade e as discussões científicas que tanto me enriqueceram. À Anabela, por falar sobre ciência com o coração!

À Iris, à Helena e à Andreia por serem umas miúdas daquelas que se querem sempre por perto, que nos contagiam com boa disposição! Ao Nacho pelas nossas conversas infindáveis sobre tudo e sobre nada e pelo incentivo que me deu na realização deste trabalho. À Gi, de quem tenho tantas saudades! Ao meu Mike, pelos sorrisos, abraços, conforto e cumplicidade! Por ontem, por hoje, e por tudo! À Bri por me mostrar que a bondade é ainda existe!

À Cátia, a minha conterrânea do coração. Foi preciso a ciência para nos juntar, mas fomos nós que contruímos o resto. E o resto é hoje tão grande que até custa a crer que não nos tenhamos "encontrado" antes! À Vânia, que precisou de tempo para entrar na minha vida, mas de onde

não vai sair! Obrigada por tanta partilha, pelos sorrisos e alegrias, pelos copos de vinho e conversas demoradas, pelos abraços e carinho, pelas frases chave e pelos conselhos! À Marta, a minha madrinha no laboratório que se tornou uma amiga para a vida, e que mesmo distante está sempre presente! À Sara, que deixou de ser colega mas que a vida encarregou de voltar a pôr no meu caminho! Obrigada pela tua ajuda, disponibilidade e compreensão, de antes e de agora! O meu obrigada também ao pessoal do CCVL por me terem recebido tão bem e ajudado nesta fase.

Aos meus amigos de sempre e para sempre! Aqueles que me conseguiram fazer desligar do trabalho e proporcionar tantos momentos inesquecíveis! Um obrigada a todos vocês por colorirem o meu caminho! À Isabel, à Melanie, as minhas pipocas mais boas! À Inês e à Rita, as amigas de uma vida! À Verónica! Ao André, ao Rodrigo e ao Nuno! Ao meu Quintas e ao Zé! Ao pessoal dos Artistas, e ao Zéi em particular pela paciência e carinho dos últimos tempos! Aos amigos da vila mais bonita de Portugal! Ao Ricardo, que me aturou tantos anos e a quem tanto devo! Que esteve sempre presente, que suportou comigo as alegrias e as tristezas da vida! Um companheiro que nunca deixará de o ser! Um Viva à nossa amizade! E um Viva à Liberdade que voltou para vos compensar por tantos momentos em que não me foi possível estar presente!

Um obrigada especial à minha família! E o maior de todos ao meu Pai! O melhor Pai do mundo (mesmo)! Aquele quem tem sido Pai, Mãe, amigo, companheiro, confidente, que me tem guiado e dado força, que me ensinou que desistir nunca é opção! A ele, devo tudo! Esta tese é, no entanto, dedicada à minha Mãe, a pessoa que mais orgulho teria por me ver chegar até aqui e que em mim deixou todo o seu amor e toda a sua força!

A produção deste trabalho teve o financiamento da Fundação para a Ciência e Tecnologia, à qual expresso os meus agradecimentos. Referência da bolsa SFRH/BD/47433/2008.

O meu Muito Obrigada a todos!

ABSTRACT

The identification of genes involved in signaling and regulatory pathways, and matrix formation is paramount to the better understanding of the complex mechanisms of bone formation and mineralization, and critical to the successful development of therapies for human skeletal disorders. To achieve this objective, *in vitro* cell systems derived from skeletal tissues and able to mineralize their extracellular matrix have been used to identify genes differentially expressed during mineralization and possibly new markers of bone and cartilage homeostasis. Using cell systems of fish origin and techniques such as suppression subtractive hybridization and microarray hybridization, three genes never associated with mechanisms of calcification were identified: the calcium binding protein *S100-like*, the short-chain dehydrogenase/reductase *sdr-like* and the betaine homocysteine S-methyltransferase *bhmt3*. Analysis of the spatial-temporal expression of these 3 genes by qPCR and *in situ* hybridization revealed: (1) the up-regulation of *sdr-like* transcript during *in vitro* mineralization of gilthead seabream cell lines and its specificity for calcified tissues and differentiating osteoblasts; (2) the up-regulation of *S100-like* and the down-regulation of *bhmt3* during *in vitro* mineralization and the central role of both genes in cartilaginous tissues undergoing endo/perichondral mineralization in juvenile fish. While expression of *S100-like* and *bhmt3* was restricted to calcified tissues, *sdr-like* transcript was also detected in soft tissues, in particular in tissues of the gastrointestinal tract. Functional analysis of gene promoters revealed the transcriptional regulation of the 3 genes by known regulators of osteoblast and chondrocyte differentiation/mineralization: RUNX2 and RAR (*sdr-like*), ETS1 (*s100-like; bhmt3*), SP1 and MEF2c (*bhmt3*). The evolutionary relationship of the different orthologs and paralogs identified within the scope of this work was also inferred from taxonomic and phylogenetic analyses and revealed novel protein subfamilies (S100-like and Sdr-like) and the explosive diversity of Bhmt family in particular fish groups (Neoteleostei). Altogether our results contribute with new data on SDR, S100 and BHMT proteins, evidencing for the first time the role for these three proteins in mechanisms of mineralization in fish and emphasized their potential as markers of mineralizing cartilage and bone in developing fish.

Keywords: S100-like, S100 calcium binding-protein; Sdr-like, short-chain dehydrogenase/ reductase; Bhmt3, betaine homocysteine S-methyltransferase; gene expression patterns; transcriptional regulation; molecular phylogeny; bone; cartilage.

RESUMO

Nos vertebrados, o sistema esquelético representa uma das mais importantes inovações ocorridas durante a evolução, estando na base de múltiplos mecanismos de adaptação. É o sistema responsável pelo suporte e proteção dos órgãos vitais, pela locomoção (através da interação com músculos, tendões e ligamentos), permite o armazenamento e balanço homeostático de minerais, a produção de fatores de crescimento e, nos mamíferos, funciona ainda como o principal órgão hematopoiético. O esqueleto dos vertebrados é maioritariamente composto por dois tipos distintos de tecidos conectivos, o osso e a cartilagem, que diferem não só nos tipos celulares que os constituem mas também na composição da sua matriz extracelular, nos tipos de vascularização e nas propriedades químicas e físicas. O esqueleto é assim um sistema bastante complexo, rigorosamente regulado por várias vias de sinalização e inúmeros fatores que ditam diversas decisões celulares para um desenvolvimento coordenado. Tal como noutros sistemas, a desregulação de um ou mais fatores moleculares leva ao aparecimento de patologias, como por exemplo a osteoporose, diferentes tipos de escolioses, osteoartrites e tumores. Apesar do sistema ósseo ser já bastante estudado, existem ainda vários intervenientes pouco caracterizados, e outros ainda desconhecidos, o que resulta numa falta de terapêutica adequada a algumas doenças. Assim, é essencial conhecer melhor os mecanismos fisiológicos (e patológicos) do osso, bem como as suas vias de sinalização e regulação e respetivos genes associados.

Durante muitos anos, o estudo do osso e da cartilagem teve como modelo preferencial os mamíferos, no entanto, mais recentemente, o peixe mostrou ter inúmeras vantagens, sendo hoje reconhecido pela comunidade científica como um modelo biológico para o estudo da esqueletogénese. A sua grande progenia, o desenvolvimento externo e estados larvares translúcidos são apenas algumas das vantagens do peixe como modelo que, apesar das diferenças que possui relativamente aos mamíferos como consequência da evolução, tem ainda uma conservação marcante a nível genético. Os peixes possuem não só genes ortólogos para a grande maioria dos genes de mamíferos, como também os mecanismos moleculares são preservados, tendo-se tornado num modelo emergente quer para o desenvolvimento quer para estudos genéticos e funcionais. Apesar do uso preferencial do peixe enquanto modelo biológico recair sobre o peixe zebra (*Danio rerio*), dadas as diversas vantagens a ele associadas, este não é o único modelo reconhecido. Um exemplo é o caso da dourada (*Sparus aurata*), um teleósteo marinho com uma importância económica muito relevante na indústria piscícola, e que tem sido largamente utilizado na investigação do osso, uma vez que o seu cultivo em larga escala leva

ao desenvolvimento de deformações esqueléticas. Na última década, têm sido desenvolvidas várias ferramentas laboratoriais para melhor estudar este modelo, quer a nível genético (construção de bibliotecas de DNA, microarrays, caracterização funcional de genes) quer a nível de culturas *in vitro* (linhas celulares com capacidade de mineralização; derivadas do osso, VSa16, e da cartilagem, VSa13). Em particular, estas linhas celulares (VSa13 e VSa16) permitiram também a utilização de técnicas, como a hibridização subtrativa supressiva (HSS) e microarrays, ferramentas moleculares que permitem evidenciam genes diferencialmente expressos em diferentes condições. Com base nestas técnicas, tem sido possível a identificação de novos genes marcadores do desenvolvimento dos componentes do esqueleto. Estudos anteriores permitiram a identificação de genes ortólogos de várias proteínas em dourada (*p.ex.*: *mgp*, *oc*, *fhl2*, *bmp2*), sendo possível estudar a sua expressão durante o processo de diferenciação celular e mineralização *in vitro*, não só confirmando a conservação das funções já descritas para mamíferos no peixe, mas também validando a dourada como modelo biológico.

Este trabalho pretende assim, caracterizar genes ainda não descritos para a dourada e com uma potencial função no processo de mineralização. A descoberta de novos genes potencialmente envolvidos neste processo partiu de um grupo de genes diferencialmente expressos durante o processo de mineralização extracelular das células VSa16 e Vsa13 identificados pelos métodos acima descritos (HSS e microarray). De entre esses genes, três suscitaram particular interesse pelas consideráveis diferenças de expressão dos seus transcritos no processo de mineralização: a proteína de ligação de cálcio S100 – *S100-like*, a desidrogenase/reductase de cadeia curta - *sdr-like* – e a betaína-homocisteína metil-transferase – posteriormente designada de *bhmt3*. Uma vez que a função destes 3 genes durante a mineralização extracelular permanece por caracterizar, e dois deles (*sdr-like* e *bhmt3*) nunca haviam sido antes associados a mecanismos celulares do osso e da cartilagem, este trabalho tem como principal objetivo o estudo e a caracterização das suas funções em processos de mineralização/esqueletogénese. A caracterização molecular destes genes envolveu: 1) o estudo da expressão de cada gene em tecidos e durante os diferentes estádios de desenvolvimento da dourada, através de técnicas de PCR em tempo real (ou northern blot para *S100-like*) e hibridação *in situ*; 2) o estudo da sua regulação transcricional, focado em fatores de transcrição associados aos processos de diferenciação do osso e da cartilagem, com recurso à transfecção de células, e 3) o conhecimento da sua evolução filogenética e taxonómica, através de técnicas bioinformáticas. A expressão genética espacial e temporal revelou que enquanto os genes *s100-like* e *bhmt3*, são altamente específicos de tecidos cartilagíneos em processo de mineralização, o gene de *sdr-like*, é encontrado em tecidos pré-calcificados com associação à diferenciação de

osteoblastos. Estes resultados sugerem assim uma função no esqueleto para os 3 genes. No entanto, se os dois primeiros são genes altamente específicos de tecidos em processo de mineralização, a expressão do gene *sdr-like* em tecidos moles, em particular nos tecidos associados ao trato gastrointestinal, pressupõe uma função não restrita aos processos de desenvolvimento do esqueleto. Com base na sua expressão, propomos que os genes *s100-like* e *bhmt3* são marcadores da ossificação endocondral em dourada, podendo ser utilizados em estudos posteriores de caracterização dos processos de diferenciação da cartilagem. Relativamente ao estudo da regulação transcricional, a nossa análise revelou que os 3 genes parecem ser regulados por fatores de transcrição com um papel importante na diferenciação de células do osso (osteoblastos) e/ou da cartilagem (condrócitos): RAR e Runx2 (*sdr-like*); ETS1 (*S100-like* e *bhmt3*); SP1 e MEF2c (*bhmt3*), confirmando mais uma vez a sua possível função na esqueletogénese. Por fim, a relação evolucionária dos diferentes ortólogos e parálogos aqui identificados, revelou novas subfamílias proteicas para S100-like e Sdr-like, as quais parecem não ter membros em mamíferos, e uma explosão de diversidade de genes Bhmt para o grupo específico de peixes Neoteleósteos, com até pelo menos seis isoformas identificadas para esta proteína.

Ao longo deste trabalho foram recolhidos novos dados que contribuíram para uma melhor caracterização das proteínas S100, SDR e BHMT, tendo sido evidenciado pela primeira vez novas funções para as duas últimas no processo de mineralização. Foi ainda demonstrado que a existência de mais do que uma isoforma para as proteínas BHMT não é exclusiva de mamíferos como até aqui se pressuponha. Finalmente, no seu conjunto os resultados obtidos neste trabalho contribuíram para a validação dos peixes como um modelo alternativo, em particular da dourada e suas ferramentas, na investigação de mecanismos moleculares envolvidos em processos de mineralização dos tecidos.

Palavras-chave: proteína de ligação de cálcio S100 – *S100-like*; desidrogenase/reductase de cadeia curta, *sdr-like*; betaína-homocisteína metil-transferase – *bhmt3*; padrões de expressão genética; regulação transcricional; filogenia molecular; osso; cartilagem.

TABLE OF CONTENTS

ABBREVIATION LIST

BHMT	Betaine homocysteine S-methyltransferase
cDNA	Coding DNA
DPF	Days post-fertilization
DPH	Days post-hatching
ECM	Extracellular matrix
ETS1	V-ets avian erythroblastosis virus E26 oncogene homolog 1
FLuc	Firefly luciferase
MEF2	Myocite enhancer Factor 2
PBS	Phosphate-buffered saline solution
PFA	Paraformaldehyde
qPCR	Quantitative real-time PCR
RA	Retinoic acid
RACE	Rapid amplification of cDNA ends
RLuc	Renilla luciferase
RPL27a	Ribosomal protein L27a
RUNX	Runt-related transcription factor
S100	S100 calcium-binding protein
SDR	Short-chain dehydrogenase/reductase;
SP1	Specific Protein 1 transcription factor
SOX	Sex determining region Y box
SSH	Suppression subtractive hybridization
TF	Transcription factor
UTR	Untranslated region

PREAMBLE

This thesis is divided into five chapters and includes a list of references common to all chapters. The first chapter presents information useful to the understanding of the data collected within the scope of this work, as well as a short description of our objectives. The second chapter is divided into two parts addressing the characterization of the gilthead seabream S100-like calcium-binding protein and presenting data on gene expression and transcriptional regulation. Chapter 2 is based on two manuscripts published in Gene Expression Patterns and in the Journal of Applied Ichthyology. The third chapter aimed at the characterization of the gilthead seabream short chain dehydrogenase/reductase protein Sdr-like regarding gene expression and molecular and is based on a manuscript submitted to Cellular and Molecular Life Sciences. The fourth chapter present data on the gilthead seabream betaine homocysteine S-methyltransferase Bhmt3 from a molecular and evolutionary perspective and the data collected was submitted to Molecular Biology and Evolution. Finally, chapter five gathers the main conclusions drawn from the data presented in this thesis and presents perspectives for future works.

CHAPTER 1

GENERAL INTRODUCTION

CHAPTER 1. GENERAL INTRODUCTION

1.1 Fish as a suitable model to study mechanisms of skeleton formation

For many years, vertebrate development has been investigated using mainly mammalian model animals - e.g. human, mouse and rat - and although mammalian genetics have largely expanded our understanding on vertebrate development during the last decades, current knowledge on mechanisms underlying physiological and pathological processes remains often insufficient to develop successful therapies targeting human diseases. The interest of the scientific community in using alternative model organisms – e.g. chicken, *Xenopus* and zebrafish – has recently emerged. Reasons are multiple but mostly related to technical or ethical advantages but also associated to the need of understanding the complexity of vertebrate diversity and evolution. Because it shares with mammals numerous structural, physiological and molecular features and because it offers many technical and financial advantages over mammalian models, fish has become a very popular and promising alternative to substitute or complement the traditional model organisms. The possibility of easily producing mutant lines (mainly from the zebrafish) mimicking human disorders, unveiling, for example, molecular mechanisms behind the development of cancer, cardiovascular or degenerative disorders (Chico et al., 2008; Feitsma and Cuppen, 2008; Ingham, 2009) or transgenic lines to study gene function or screen for molecules with therapeutic potential (Tamplin et al., 2012; North and Zon, 2003; Quach et al., 2015; Wehner et al., 2015; Zheng et al., 2014) have also stimulated the interest of pharmaceutical companies in using fish systems. It is safe to say today that our knowledge on the molecular basis of human pathologies but also on the mechanisms underlying vertebrate development has greatly improved since the implementation of fish as a lab model.

Despite small differences that may be the consequence of evolutionary distance (last common ancestor existed approximately 420 million years ago), cellular components, molecular pathways and mechanisms involved in the onset of patterning and development of skeletal structures in fish are very similar to those observed in mammals, thus skeletal development and bone/cartilage formation have been largely studied in fish. As a consequence,

in vitro (e.g. mineralogenic cell lines), *ex vivo* (e.g. scale culture) and *in vivo* (e.g. mutant lines with skeletal phenotype and transgenic lines for skeletal genes) have been developed (reviewed in Laizé et al. (2015) Apschner et al. (2011) and McGonnell and Fowkes (2006)). Although the Japanese medaka and the guppy have been used in some studies, the zebrafish (*Danio rerio*; a teleost and a tropical freshwater fish) is currently the most used fish model for scientific research related to skeletogenesis and bone/cartilage formation (Haffter et al., 1996; Lieschke and Currie, 2007) because of many technical advantages: the possibility to monitor each step of the skeleton formation due to transparent embryos and external development, a rapid growth with almost all the body structures visible at 48 hours post-fertilization, an easy maintenance due to its small size (up to 4-cm long) and remarkable robustness (it can adapt to a wide range of environments), a large progeny (hundreds of eggs per spawning), a short generation time (adulthood attained in 3–4 months) and the availability of various genomic tools (zebrafish genome is almost completely sequenced and annotated, with most human genes having orthologs in zebrafish) have reinforced the attractiveness of the zebrafish over other fish as a laboratory model. In addition, approximately 70% of human disease genes appears to have functional homologs in zebrafish (Langheinrich, 2003) and key regulators orthologs of bone formation shares significant sequence similarities and overlapping of expression patterns (Spoorendonk et al., 2010). Marine teleost fish have also been used to get insights into the mechanisms of skeleton/osteogenesis and tissue mineralization. Because they are important species for aquaculture and because they suffer a high rate of skeletal abnormalities when farmed under intensive conditions, the Atlantic salmon (*Salmo salar*), the Senegalese sole (*Solea senegalensis*), the European seabass (*Dicentrarchus labrax*) and the gilthead seabream (*Sparus aurata*) have been extensively used in bone research (Gil Martens et al., 2010; Cardeira et al., 2015; Boglino et al., 2013; Benitez-Santana et al., 2013; Wang et al., 2013).

1.2 Vertebrate skeleton

A mineralized skeleton represents one of the most important innovations that occurred throughout vertebrate evolution, offering the basis for many adaptive mechanisms. In vertebrates, the internal skeleton performs several important functions: support and protection for vital organs, body movements through interaction with muscles, tendons and ligaments, storage and balance for calcium and phosphate, and production of growth factors (Kronenberg, 2003; Pirraco et al., 2010; Rameshwar and Stegemann, 2013). Beside these well-known

functions, a role of the skeleton in the endocrine regulation of energy metabolism has been proposed recently (Karsenty and Oury, 2014). In mammals, skeleton, through the bone marrow, is also the principal hematopoietic organ and the site of blood cells production (Taichman, 2005; Kronenberg, 2003; Pirraco et al., 2010). Vertebrate skeleton is mainly composed of two distinct supporting connective tissues, cartilage and bone, which differ in a number of important characteristic such as cell types and the composition of their extracellular matrix, but also in the vascularization and the mechanical, chemical and physical proprieties (Kardong, 1998; Marks and Odgren, 2002).

1.2.1 Cartilage

Cartilage is an avascular connective tissue with a low metabolic rate and is characterized by the presence of chondrocytes embedded in a rigid matrix rich in collagen (or fibrous proteins in lamprey and hagfish) and acidic polysaccharides (Person and Mathews, 1967; Cole and Hall, 2004a; Cole and Hall, 2004b). Chondrocytes differentiate from condensed mesenchymal stem cells and undergo a number of maturational stages characterized by the induction of phenotypic marker genes (Stein et al., 1996; Otto et al., 1997; Olsen et al., 2002; Beck et al., 2001; Karsenty, 2001; Provot and Schipani, 2005; Wuelling and Vortkamp, 2010). Cartilage anlagens may remain as cartilage throughout their existence or be replaced by bone, through a process known as endochondral ossification. During this process, chondrocytes in the center of the cartilage anlagen undergo a program of proliferation, differentiation, hypertrophy and cell death. At each step of the differentiation process, chondrocytes are characterized by unique morphologies, gene expression profiles and metabolic activities (Karsenty and Wagner, 2002) that appear sequentially following specific steps (**Figure 1.1**):

1) Differentiated chondrocytes start to produce a matrix rich in type II collagen followed by matrix enlargement;
2) Chondrocytes become hypertrophic and synthetize a matrix rich in type X collagen and blood vessels;
3) Chondroclast (cells responsible for collagen matrix degradation) are recruited and the flattened fibroblastic cells present in the perichondrium surrounding the chondrocytes are directed to differentiate into osteoblast (which are involved in bone collar formation (Olsen et al., 2002; Provot and Schipani, 2005);

4) Hypertrophic chondrocytes undergo apoptosis. The cartilaginous matrix left behind provides a scaffold where minerals will be deposited by bone cells (reviewed in Kronenberg, 2003).

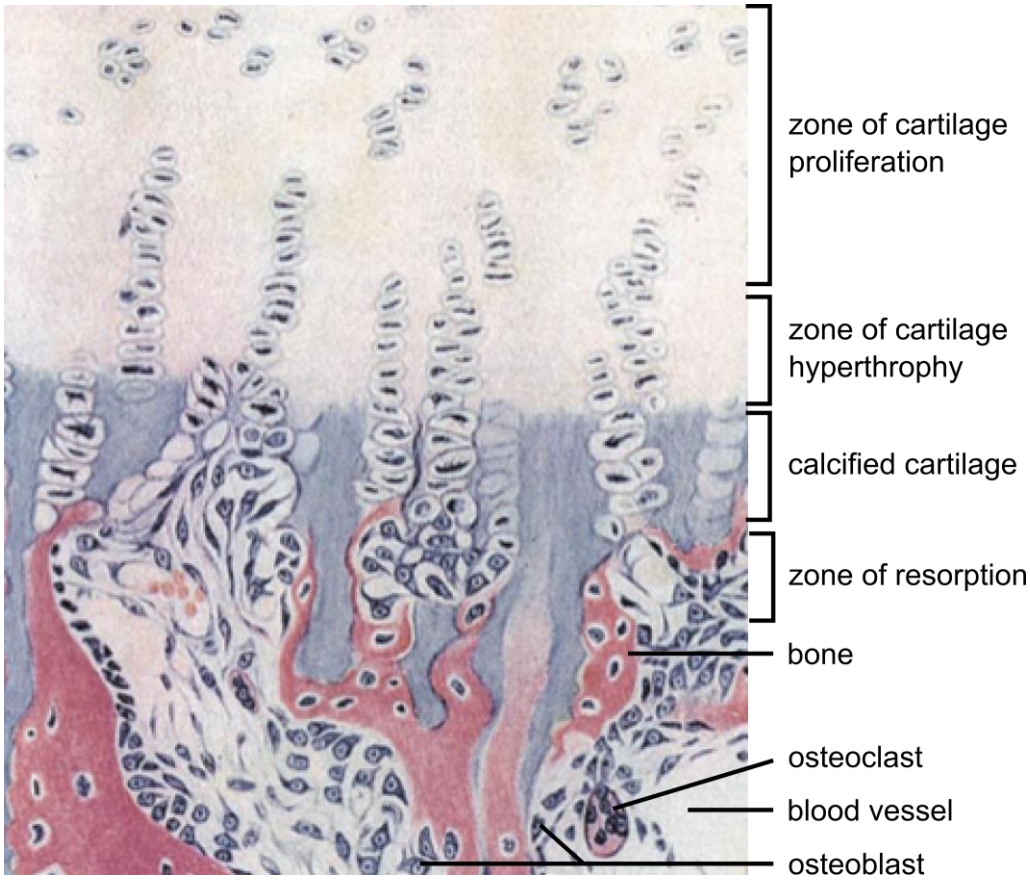


Figure 1.1. Endochondral bone formation and bone remodeling in vertebrate systems. Proliferative and hyperthrophic chondrocytes are observed in the zone of cartilage proliferation and hyperthrophy, respectively. Calcified cartilage is eventually replaced by bone, upon blood vessel invasion and osteoblast recruitment. Bone resorption performed by a multinucleated osteoclast is indicated in the bottom of the image. Adapted from Ross and Pawlina (2011).

It is important to mention at this step that ossification can occurred independently of the pre-existence of a cartilage matrix in a process known as intramembranous ossification also involving the condensation of mesenchymal stem cells (see details in section 2.2.). The cell fate decisions made by the aggregation of mesenchymal cells that ultimately results in cartilage or bone formation are regulated by a complex and elaborately skeletogenic gene network which includes numerous transcription factors, growth factors, signaling pathways, post-transcriptional regulators and epigenetic factors (Gaur et al., 2010; Kobayashi et al., 2008; Oberlender and Tuan, 2008; Akiyama, 2008; Chun et al., 2008; Tuli et al., 2003; Michigami,

2013; Michigami, 2014; Goldring et al., 2006; Zeng et al., 2005; Hill et al., 2005; Yoshida and Komori, 2005), that despite important are not relevant to the topic of this dissertation (**Figure 1.2**).

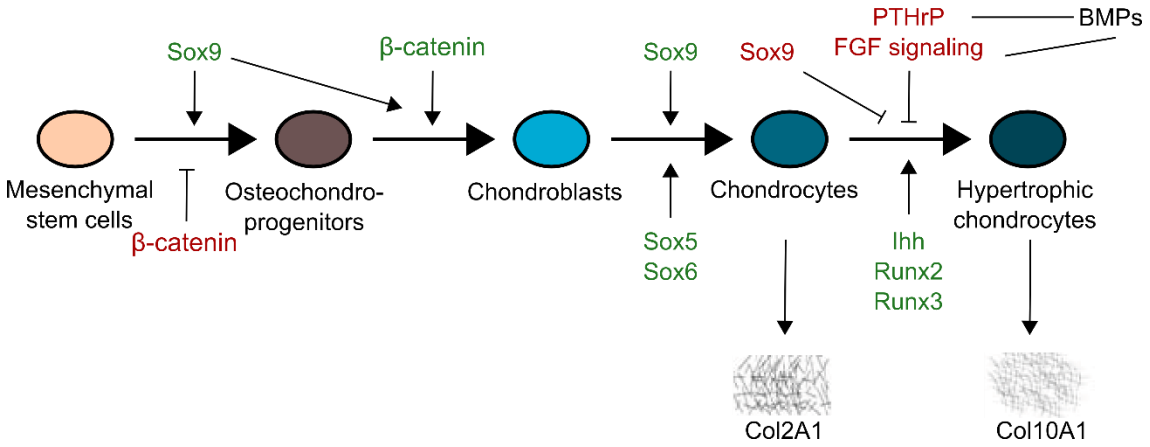


Figure 1.2. Schematic representation of multi-step events that directs mesenchymal cells along chondrogenic differentiation pathways. The different stages of chondrocyte differentiation are represented schematically and main growth and differentiation factors, the transcription factors and at each stage are indicated. Arrows indicate positive regulation, lines indicate interaction, and intersected lines indicate negative regulation (adapted from Zhang et al., 2009).

Mammalian *versus* teleost cartilage

Understanding the relationships among the different types of cartilage found among vertebrates, as well as their biochemical and molecular characterization, has been a challenge. There are generally four kinds of cartilage in vertebrates and invertebrates (matrix-rich cartilage, cell-rich cartilage, vesicular cartilage, and acellular cartilage) however the evolution of those types of cartilage remains unknown (Cole and Hall, 2004a); they could have evolved independently or diversified from a single type of ancestral connective tissue (Stemple, 2004; Zhang and Cohn, 2006). Despite the lack of information regarding this issue, the different types of cartilages found in mammals and teleosts are quite well characterized.

Three major types of cartilage – hyaline, elastic and fibrocartilage – are present in mammals and they can be distinguished following physical characters and matrix components. The hyaline cartilage, named after its semi-transparent and bluish-white color, is the most common and its matrix is rich in glycosaminoglycans and type II collage; it is found in the embryonic models of endochondral bones and in portions of the laryngeal cartilage (Hall, 2005). Elastic cartilage is also rich in glycosaminoglycans and type II collagen fibril, but

additionally contains thick bundles of elastic fibrils and elastin-rich extracellular matrix; it is found mainly in the pinna, larynx, epiglottis and intervertebral discs (Naumann et al., 2002). If the extracellular matrix is rich in type I collagen fibers, which makes it both tensile and tough, cartilage is called fibrocartilage; it is found where ligaments and tendons attach to bone but also in intra-articular discs of joints and as articular cartilages at joint surfaces (Benjamin and Evans, 1990; Benjamin and Ralphs, 2004; Eyre and Wu, 1983). Despite these tree main types of cartilage, some cartilages can demonstrate intermediate tissue properties, not comprised by this tidy classification. As an example, the secondary cartilage, present at stressed joint regions, is formed from osteoblast precursors and besides being similar to hyaline cartilage expresses high amounts of type I collagen (Fukada et al., 1999; Fukuoka et al., 2007).

There are more types of cartilage in teleost fish. Benjamin and co-workers (1990) divided teleost cartilages into at least eight main types (however 16 types can be identified in some teleost), most of them with no counterparts in mammals (**Figure 1.3**). In the lips, rostral folds and other cranial cartilages hyaline-cell cartilage (HCC) is widespread. It is composed of compact chromophobic chondrocytes and hyaline cytoplasm with little content in matrix (Benjamin, 1990). HCC can be further divided into three sub-types depending on matrix content and cell composition: fibrohyaline-cell cartilage with a matrix rich in collagen; elastic hyaline-cell cartilage with an elastin matrix; and lipohyaline-cell cartilage which contains also adipose cells (Benjamin, 1990). Zellknorpel cartilage (ZC) is even more chromophilic than HCC; it is contracted within the large lacunae and usually found in gill filaments, basal plate and others (Benjamin, 1990). If cartilage possess highly cellular elastic fibers and non-hyaline cells it is in turn denominated as elastic/cell-rich cartilage (ECRC), and can be distinguish from HCC and ZC by elastic staining. It is surrounded by a thick fibrous perichondrium and is usually found at the barbels and maxillary oral valves (Benjamin, 1990). Cell-rich hyaline cartilage (CRHC), on the other hand, is also a hyaline-like cartilage but with more cells and lacunae that occupy more than half of the total volume. Parts of neurocranium and Meckel's cartilage (MC) belong to this category (Benjamin, 1990). Teleost have a matrix-rich hyaline cartilage (MRCH) similar to the mammalian hyaline cartilage; it is common in gill arches and part of the neurocranium (Benjamin, 1990). At last, teleosts have a unique type of cartilage located only in the scleral lens named as scleral cartilage (Sc).

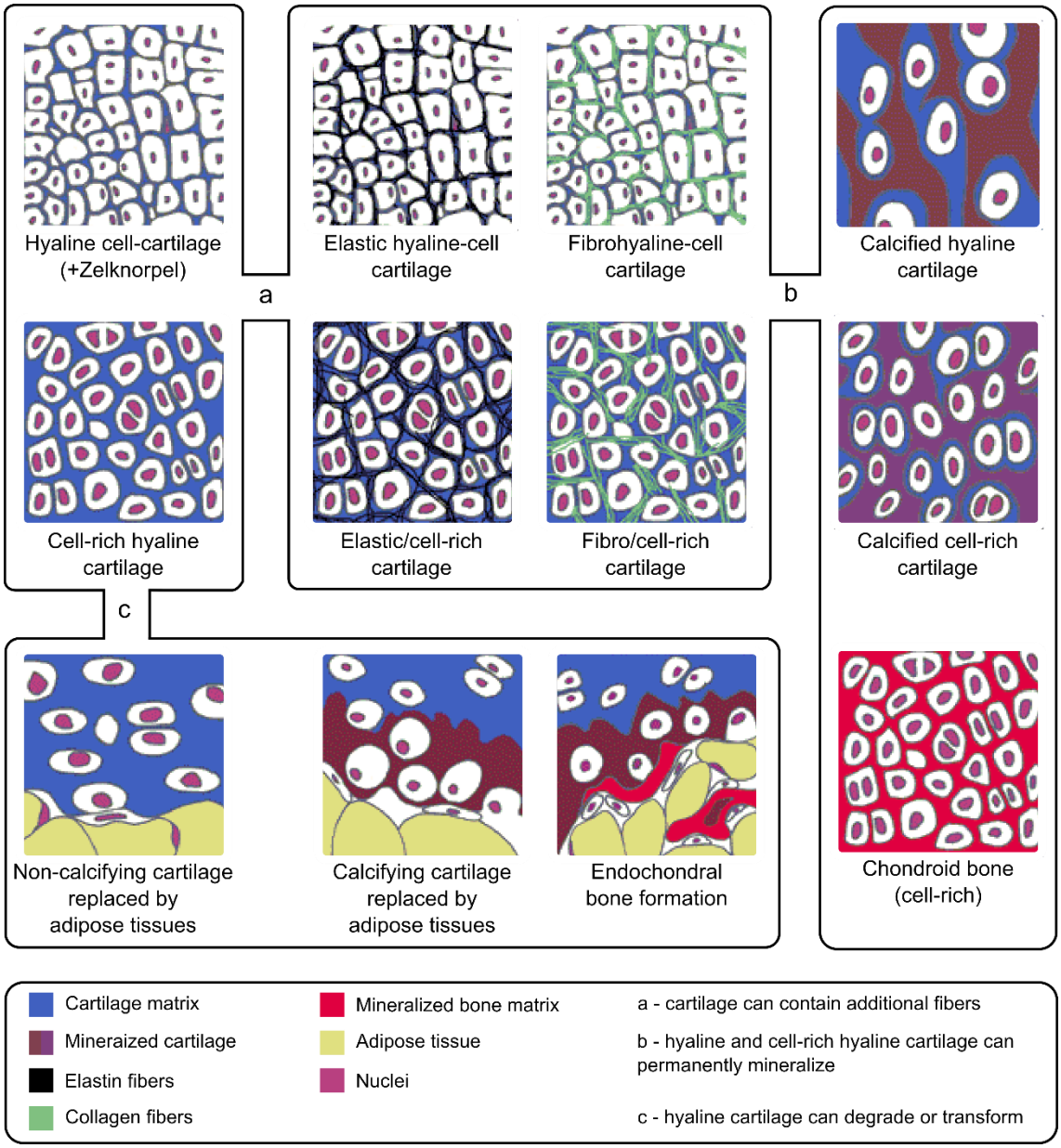


Figure 1.3. Types of differentiated cartilage and cartilaginous tissues in teleost fish. Letters a-c indicate typological and/or developmental relationships that connect the main cartilage categories (adapted from Witten et al., 2010).

1.2.2 Bone

Bone, in contrast to cartilage, is vascularized and suffers constant and dynamic remodeling. It is composed of a mineralized extracellular matrix and three main types of cells: osteoblasts (bone forming cells), osteocytes (bone sensing cells; they are osteoblasts that cease division and are trapped into the bone matrix) and osteoclasts (bone resorbing cells). Osteoblasts derive from pluripotent mesenchymal stem cells of the bone marrow that undergo

a maturation process where transcription factors such as Runt related factor 2 (RUNX2) and Oxterix (OSX or SP7) play determinant roles (reviewed in Karsenty and Wagner, 2002). They are observed on bone surfaces and are responsible for the deposition of the osteoid, an un-mineralized bone matrix that will progressively mineralize to form bone, trough hydroxyapatite deposition. Once entrapped into the mineralized matrix (composed mainly of type I collagen, the major extracellular matrix component but also of non-collagenous proteins such as osteocalcin, osteopontin and osteonectin (Clarke, 2008), osteoblasts stop dividing; they will mature/differentiate into osteocytes and acquire a star-shaped morphology with extensions that will join and allow the interconnection of neighboring osteocytes (Dallas et al., 2013). They act as mechano-sensors and master regulators of bone remodeling by secreting factors that regulate the activity of both osteoblasts and osteoclasts (Dallas et al., 2013) (**Figure 1.4**).

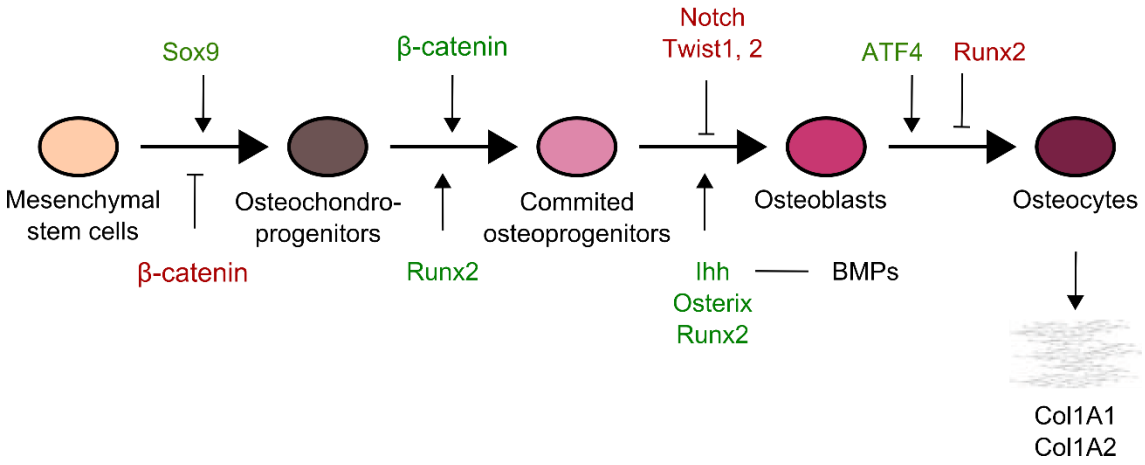


Figure 1.4. Schematic representation of multi-step events that directs mesenchymal cells along osteoblastic differentiation pathways. The different stages of osteoblastogenesis are represented schematically and main growth and differentiation factors, and transcription factors at each stage are indicated. Arrows indicate positive regulation, lines indicate interaction, and intersected lines indicate negative regulation (adapted from Zhang et al., 2009).

Unlike osteoblast, osteoclasts belong to the monocyte-macrophage cell lineage (reviewed in Karsenty and Wagner, 2002); they are large multinucleated cells resulting from the fusion of mononuclear osteoclasts and are responsible for bone resorption through the secretion of hydrolytic enzymes (cathepsin K and matrix metalloproteinases) and the acidification of the resorption compartment, responsible for the dissolution of the organic matrix and consequent release of bone minerals (calcium and phosphorous) (Nakamura et al., 2012). After completing their function osteoclast eventually undergo apoptosis in order to avoid excessive bone

resorption. This bone resorption process is tightly controlled by osteocytes but also by osteoblasts, through the secretion of specific factors such as receptor activator of NF-κB ligand (RANKL) and osteoprotegerin (OPG), recognized by osteoclasts (Dallas et al., 2013; Pirraco et al., 2010; Nakamura et al., 2012; Caetano-Lopes et al., 2007). As mentioned before, bone can form through two distinct processes: 1) endochondral bone ossification (bone is formed from a cartilage template - see details in section 2.1.) and 2) intramembranous ossification that involves the condensation of mesenchymal precursors but occurs directly from their differentiation into osteoblast, independently of the pre-existence of a cartilage matrix. This differentiation process is controlled by transcriptional regulators of osteoblast differentiation (Komori et al., 1997; Otto et al., 1997; Sato et al., 1998; Choi et al., 2001; Hess et al., 2001; Nakashima et al., 2002) and involves the production of a matrix rich in type I collagen (osteoid) that will later mineralize to form for example head bones (e.g. skull flat bones and jaw; Crombrugghe et al., 2001; Karsenty, 2003; Karsenty and Wagner, 2002).

Bone as an active and dynamic tissue is in constant remodeling, i.e. the replacement of old and damaged bone by new bone to maintain the structural integrity of the skeleton and bone volume; remodeling is also central to calcium and phosphorous metabolism, and also occurs, despite the alternative and perhaps ancient pathways, in acellular bone of teleost species (see details in section 2.2.1.; Shahar and Dean, 2013). Phases of bone remodelling are (Parra-Torres and Valdés-Flores, 2013) (**Figure 1.5**):

1) Activation phase - it initiates with the detection of a remodeling signal (micro-fractures, mechanical load or the release of factors such as insulin growth factor I (IGF1), tumor necrosis factor α (TNF-α) and parathyroid hormone (PTH) into the bone microenvironment) and continues with osteoclast differentiation trough an increase in RANKL expression;
2) Resorption phase - osteoclast attached to bone surface form a sealed lacuna that they acidify by secreting H^+ ions facilitating bone dissolution and thus promoting contact of the organic matrix with proteolytic enzymes that degrade it;
3) Reversal phase - it is characterized by osteoclast apoptosis and osteoblast recruitment and differentiation. It is also associated with the cleaning of the lacuna from bone matrix leftovers to facilitate osteoblast attachment;
4) Formation phase - growth factors stored in bone (e.g. fibroblast growth factors, transforming growth factors and bone morphogenetic proteins) are released into the lacuna and trigger osteoblast recruitment and the production of the osteoid that will get mineralized;

5) Termination phase - osteoblasts differentiate into osteocytes that remain embedded inside the mineralized matrix and secrete inhibitory factors that slow down the rate of bone formation.

The remodeling cycle ends when resorbed bone has been replaced by an equal quantity of newly formed bone (Parra-Torres and Valdés-Flores, 2013).

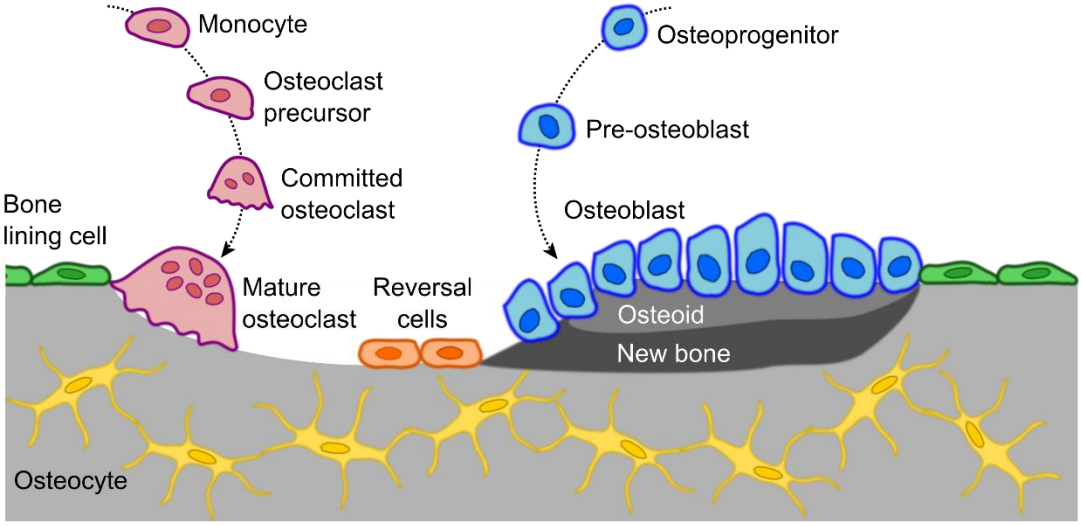


Figure 1.5. Schematic representation of bone cells involved in bone remodeling. Osteocytes (star-shaped yellow cells) are embedded within the mineralized bone matrix and connected through a complex network of cytoplasmic extensions inside lacunae and canaliculi. They are actively involved in bone turnover through the recruitment of bone forming cells (osteoblasts; blue) and bone resorbing cells (osteoclasts; pink). Bone remodeling can be influenced by a variety of factors, such as mechanic stress, structural damage or exposure to systemic or paracrine factors. Haematopoietic cells of the monocyte/macrophage lineage differentiate into mature osteoclast and resorb bone. During the reversal phase, osteoprogenitors are recruited to the site of resorption, differentiate and secrete the osteoid that will mineralize and form new bone. Adapted from Nicholls et al., 2012 by Vincent Laizé.

Mammalian *versus* teleost bone

Teleost fish and mammalian bones are very similar with respect to anatomic characteristics with much of the skull, axial and appendicular skeleton extraordinarily conserved, and also developmental events regarding bone formation that have been maintained throughout evolution from fish to human ((Hall, 2005; Javidan and Schilling, 2004). Bone cells have the same origin – mesenchymal for osteoblasts and hematopoietic for osteoclasts – and the same function – bone formation by osteoblasts and resorption by osteoclasts – in both the teleosts and mammals (Witten and Huysseune, 2009; Shahar and Dean, 2013). Endochondral

and intramembranous ossification are mechanisms of bone formation occurring in both mammalian and teleost skeleton (Hall, 2005; Shahar and Dean, 2013). The most remarkable difference between mammalian and telostean bone is maybe the presence of acellular bone (absence of osteocytes) in most advanced teleost fish (e.g. gilthead seabream), while bone of basal teleost – e.g. zebrafish – and primitive osteichthyans contains osteocytes (cellular bone) (Cohen et al., 2012; Kranenbarg et al., 2005; Meunier and Huysseune, 1992) Although they lack osteocytes – the cell type associated in mammals with bone remodeling and sensing of mechanical load – acellular bones are still metabolically active and capable of resorbing, remodeling and responding to mechanical stimuli (Dallas et al., 2013; Witten and Huysseune, 2009; Shahar and Dean, 2013). Calcium-phosphorus homeostasis is regulated via the local process of osteocytic osteolysis, but since calcium deficiency in fish is rare and rather unlikely to occur (calcium is not limiting in both seawater and fresh water) and since calcium deficiency imposed to fish artificially lead to the mobilization of calcium stored in scales (exoskeleton) rather than in bones of the endoskeleton (Takagi and Yamada, 1992), Shahar and Dean (2013) suggested that metabolic cost of maintaining osteocytes had led to an evolutionary pressure toward acellularity (**Figure 1.6**). In fact, bone resorption in fish is mainly triggered by phosphorus deficiency, which availability in both fresh water and seawater is relatively low (Roy et al., 2002). Shahar and Dean (2013) have further suggested that most of the important osteocytic functions occur in acellular bone through alternative pathways accomplished by non-osteocytic routes (possibly through osteoclast and osteoblast signaling pathways).The presence of mononucleated osteoclasts in most teleost fish species (in advanced teleosts small mononucleated osteoclast are the prevailing cell type) while multinucleated osteoclasts are exclusively found in mammals is another characteristic that distinguishes cellular and acellular bones. In fact, the lack of osteocytes in advanced teleost bone could be the cause of the modified morphology of osteoclasts, since osteocytes are thought to regulate the differentiation of osteoclasts and trigger the fusion of mononucleated to multinucleated osteoclasts. This is in agreement with an alternative mode of bone resorption, that in advanced teleost occur without generating typical resorption lacunae (reviewed in Apschner et al., 2011).

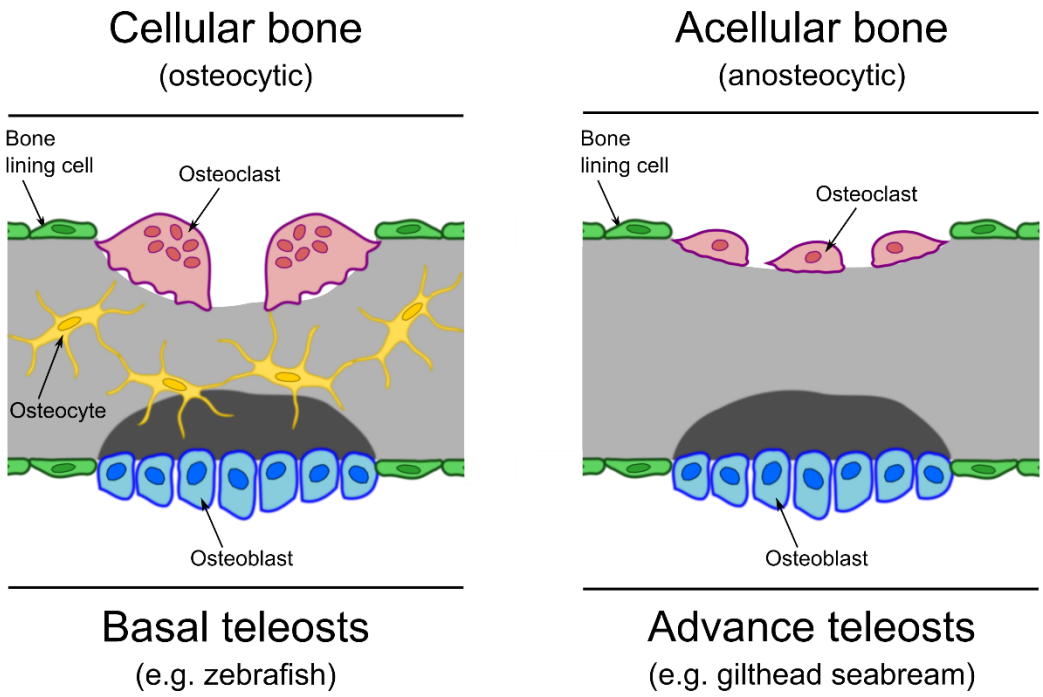


Figure 1.6. Main differences between osteocytic and anosteocytic bone. Osteocytes (star-shaped yellow cells), embedded in bone matrix are only present in osteocytic bone. In organisms that possess anosteocytic bone, osteoclasts (pink) are usually mononucleated and have a limited capacity of bone resorption, creating shallow lacunae, contrasting with the giant multinucleated cells found in osteocytic bone which produce deep resorption lacunae. Bone forming cells (osteoblasts) and bone lining cells are depicted in blue and green, respectively. Adapted from Witten and Huysseune, 2010 by Vincent Laizé.

1.3 Fish systems to study bone and cartilage formation and mineralization

The last decade has seen an increase in the development of transgenic fish lines (mainly zebrafish and medaka) and mutant lines to model human skeletal disorders (reviewed in Laizé et al., 2015). For example, osteogenesis imperfect, osteoporosis, hyper-ossification and skeletal overgrowth, idiopathic scoliosis, craniosynostosis are examples of human pathologies affecting skeleton that can be modeled by fish mutants (**Table 1.1**; reviewed in Laizé et al., 2015). Unveil the important mechanisms behind bone and cartilage tumor diseases has also been possible through the use of several zebrafish mutants (zebrafish *dackel*, zebrafish *boxer*, zebrafish *pinscher*; (**Table 1.1**; Clément et al., 2008).

Table 1.1. Example of fish models of human bone and skeletal disorders (adapted from Laizé et al., 2015).

Human bone/skeletal disorders (bone/skeletal phenotype)	**Fish model systems**	**Affected gene(s)**
Osteogenesis imperfecta (OI) (reduced bone density, bone fragility, skeletal deformities)	Zebrafish *chihuahua (chi)* mutant Zebrafish *frilly fins (frf)* mutant	*col1a1* *bmp1*
Osteoporosis (reduced bone mineral density)	*rankl-induced* medaka	*rankl*
Glucocorticoid-induced osteoporosis (GIOP) (reduced bone mineral density upon use of steroids)	Prednisolone-treated zebrafish	
Iron-induced osteoporosis (reduced bone mineral density upon iron overload)	Iron-overloaded zebrafish	
Raine syndrome (RNS) (increased ossification)	Zebrafish *fam2b* mutant	*fam20b*
Multiple osteochondromas (MO) (cartilaginous bone tumors leading to skeletal deformities)	Zebrafish *dackel (dak)* mutant Zebrafish *boxer (box)* mutant Zebrafish *pinscher (pic)* mutant	*ext2* *extl3* *papst1*
Mucolipidosis II (ML-II) (skeletal, craniofacial and joint abnormalities)	Zebrafish *gnptab* morphant	*gnptab*
Craniosynostosis (premature fusion of cranial sutures)	Zebrafish *dolphin (dol)* and *stocksteif (sst)* mutants	*cyp26b1*
Holospondyly (fusion of vertebral centra)	Zebrafish *stocksteif (sst)* mutant Retinoic acid-treated zebrafish	*cyp26b1*
Fibrodysplasia ossificans progressiva (FOP) (heterotopic endochondral ossification)	Zebrafish *lost-a-fin (laf)* mutant	*acvr1/alk8*
Idiopathic scoliosis (spinal deformity)	Guppy *curveback* mutant	Not determined
Arterial calcification of infancy (ectopic mineralization)	Zebrafish *dragonfin (dgf)* mutant	enpp1
Holoprosencephaly (HPE) (craniofacial defects)	Zebrafish *sonic-you (syu)* mutant	shh
Campomelic dysplasia (craniofacial defects)	Zebrafish *jellyfish (jef)* mutant	sox9a
Ehlers-Danlos syndrome (EDS) (craniofacial defects)	Zebrafish *b4galt7* mutant	b4galt7
DiGeorge syndrome (DGS) (craniofacial defects)	Zebrafish *van-gogh (vgo)* mutant	tbx1
Cranio-lenticulo-sutural dysplasia (CLSD) (craniofacial defects and short stature)	Zebrafish *crusher (cru)* mutant	sec23a
Osteopathy related to mineral homeostasis (failure to form mineralized bone)	Zebrafish *no bone (nob)* mutant	entpd5
Osteopathy related to abnormal ECM deposition (craniofacial defects)	Zebrafish *feelgood (fel)* mutant Zebrafish *man o'war (mow)* mutant Zebrafish bulldog (bul) mutant	creb3l2 uxs1 sec24d
Osteopathy related to delayed mineralization (delayed vertebrae calcification)	Zebrafish *bone calcification slow (bcs)* mutant	Not determined
Hyperossification (hyperossification and skeletal overgrowth)	Zebrafish *rapunzel (rpz)* mutant	rpz

Additionally, a vast number of transgenic lines had been developed allowing studies at cellular level, with elevated morphological detail: i*n vivo* labeled fluorescent proteins under the control of promoters related to bone (*osx* (*sp7*), *oc2*, *runx2*, *sox9*, *sox10*, *barx1*, *col2* and *col10;* DeLaurier et al., 2012; Hammond and Moro, 2012; Knopf et al., 2011; Nichols et al., 2013; Renn et al., 2013) can be seen *in loco*, to visualize bone and cartilage signaling during bone development *in vivo*. Furthermore, the use of fluorescent proteins to highlight particular structures in the skeleton without the need to sacrifice the fish or to reduce the number of

specimens needed for a particular observation is in agreement with the European guidelines aiming at limiting animal experimentation.

Because of their small size, transparency, rapid growth and availability, fish embryos/larvae, in particular those from zebrafish, are a valuable tool for high-throughput screening of molecule libraries. Larvae can be reared in 96-well plates and numerous molecules can therefore be tested at the same time improving and speeding up screenings.

Thus, looking at the effect of a molecule in the whole-organism is a desirable approach since it not only allows the identification of potential drawbacks but also the determination of therapeutic activity, range of action and general toxicity in the different body structures. Since in zebrafish the onset of bone formation and mineralization occurs as early as 2 day post-fertilization (dpf) and can be assessed easily through whole-mount bone-specific staining at these early stages (Gavaia et al., 2000; Walker and Kimmel, 2007) they are of great potential for screening of various novel osteogenic and/or mineralogenic drugs (Dong et al., 2012; Suzuki et al., 2000; Laizé et al., 2014). Moreover, these drug tests can be performed using the existing transgenic lines, where bone and cartilage cells are marked, something difficult in traditional mammalian models. Drugs and mechanisms affecting *de novo* bone formation can also be studied using the caudal fin regeneration system, where caudal fin is amputated 2 segments before the first branching of the rays and new bone is formed after 3 days of epimorphic regeneration. Caudal fin is a simple and accessible structure where *de novo* bone mineralization can be easily determined by alizarin red staining, imaging or morphometric (Laizé et al., 2014), thus becoming an excellent system for investigating underlying mechanisms of bone regeneration (Nakatani, Kawakami, et al., 2007).

As a complement to *in vivo* systems, *ex vivo* and *in vitro* approaches have also been established. As for regenerating fin rays, elasmoids scales of teleost fish (dermal bone elements) have been recently used to better understand mechanism of bone regeneration (Metz et al., 2012; De Vrieze et al., 2011) but also as an *ex vivo* disease model for osteoporosis studies (De Vrieze et al., 2014). They are useful to the study of cell-cell and cell-matrix interactions since they can be maintained *in vitro* as a bone unit (osteoclasts and osteoblasts cohabit on both sides of the mineralized matrix interacting similarly to *in vivo* conditions) in conditions resembling *in vivo* conditions. In addition, cell lines of fish origin capable of mineralizing their extracellular matrix *in vitro* (VSa13 – chondrocyte-like cells – and VSa16 - osteoblast-like cells) were established in 2004 from vertebrae of the gilthead seabream (Pombinho et al., 2004) opening a new variety of possibilities to unveil bone and cartilage mechanisms (**Figure 1.7**). More two mineralogenic cell lines have been developed after from gilthead seabream; one is

derived from the lower jaw (JSa1; Rafael et al., 2010) (**Figure 1.7**) and the second from the branchial arches (ABSa15; Tiago et al., 2014). Their ability to mineralized the extracellular matrix upon exposure to a mineralogenic cocktail composed of ascorbic acid, β-glycerophosphate and calcium chloride (Pombinho et al., 2004; Marques et al., 2007; Vijayakumar et al., 2013; Tiago et al., 2014), the relative rapidity of mineralization (onset and extent of mineralization is cell line-specific but mineral deposition is usually detected after 2-4 weeks of treatment; (Marques et al., 2007), and the simplicity in quantifying mineral deposition by alizarin red S or von Kossa staining, have fostered the use of mineralogenic cell lines in studies aiming at identifying pathways regulating cartilage and bone cell function and differentiation, but also mechanisms underlying extracellular matrix mineralization.

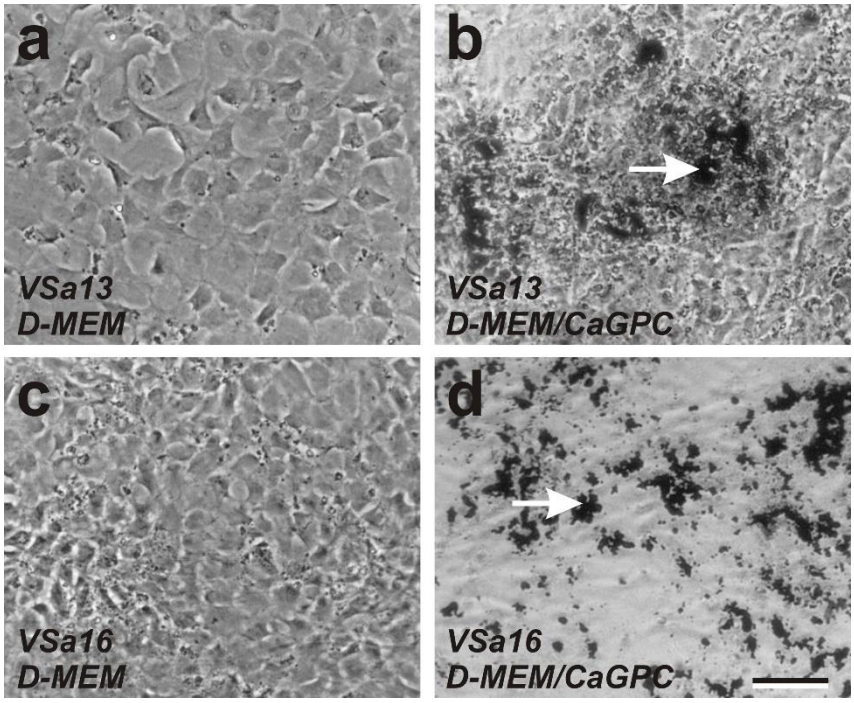


Figure 1.7. Micrograph of gilthead seabream VSa13 (**a**, **b**) and VSa16 (**c**, **d**) cell lines established from calcified vertebrae and cultured under control (**a**, **c**) and mineralizing conditions (**b**, **d**) then von Kossa-stained to reveal mineral deposition. Bar is 100 μm (from Pombinho et al., 2004).

Gilthead seabream cell lines were used to study the role of marker genes such as osteonectin (Laizé et al., 2005), osteopontin (Fonseca et al., 2007); matrix Gla protein (Conceição et al., 2008) or bone morphogenetic proteins (Rafael et al., 2006) but also to investigate the pathways involved in the proliferative and mineralogenic effects of vanadate (Tiago et al., 2008), retinoic acid (Fernández et al., 2014; Conceição et al., 2008) and polyunsaturated fatty acids (Viegas et al., 2012). Osteotoxicity of environmental pollutants and

their influence on the expression of bone-specific marker genes (V. Laizé, personal communication) and osteogenic activity of marine molecules purified from marine algae (M.L. Cancela, personal communication) are currently being tested using gilthead seabream mineralogenic cell lines. The unique position of gilthead seabream as a promising model of marine fish to study and identify genes and signaling pathways involved in mechanisms of tissue mineralization is certainly related to the availability of the several mineralogenic cell lines (and many other cells lines; see the Ficel database of fish cell lines available at bioskel.ccmar.ualg.pt) but also to the extended knowledge on skeletal and bone formation (Benjamin, 1990; Benjamin, JR Ralphs, et al., 1992; Faustino and Power, 2001; Faustino and Power, 1998; Faustino and Power, 1999; Pinto et al., 2003), the availability of various bone and cartilage marker genes (Pinto et al., 2001; Pinto et al., 2003; Laizé et al., 2005; Cancela et al., 1995; Simes et al., 2003) and transcriptomic and genomic data (Sarropoulou, Kotoulas, et al., 2005; Sarropoulou, Power, et al., 2005; Tiago et al., 2011), and the availability of a radiation hybrid panel (Senger et al., 2006).

1.4 New bone markers in fish

To better understand the complex mechanisms of bone formation and mineralization, it is primordial that genes central to signaling pathways, regulatory pathways and matrix formation are identified. Because *in vitro* cell systems derived from bone can mineralize their ECM and expressed various bone marker genes in a sequence that can be compared to the *in vivo* process of bone mineralization, they have been successfully use to discover new marker genes in mammals (Doi et al., 2002; Raouf and Seth, 2002; de Jong et al., 2004). Because bone cell types of fish origin are very similar to those in mammals regarding cell function and gene regulation (Wagner et al., 2003), they have already been successfully used to characterize orthologous genes related to bone as stated in section 2.3.2. and to identify novel genes involved in mechanisms of ECM mineralization. Gilthead seabream VSa13 and VSa16 cells were used in combination with suppression subtractive hybridization (SSH; Fonseca et al., 2007) and more recently microarray hybridization (Tiago et al., 2011) to identify genes differentially expressed during extracellular matrix mineralization and upon vanadate exposure, a transition metal with anti-mineralogenic activity. From the pool of genes differentially expressed in VSa16 cells during ECM mineralization and identified through SSH, 3 novel genes were of particular interest: while the role of *osteopontin* in mechanisms

underlying *in vitro* mineralization was further characterized (Fonseca et al., 2007), the role of *S100-like*, a calcium binding protein of the S100 protein family and *SDR-like*, a short-chain dehydrogenase/reductase, remained to be determined. Similarly, among the thousands of genes differentially expressed in VSa13 cells during ECM mineralization and identified through microarray hybridization, several were already know and one was of particular interest because remarkably down-regulated during mineralization and upon exposure to vanadate (Tiago et al., 2011): the *betaine homocysteine S-methyltransferase* (*bhmt*). While a role of S100 calcium binding protein in mechanisms of bone formation and tissue mineralization was conceivable (calcium ions play a central role in cell physiology as a signaling molecule (Ikura, 1996), but are also involved in the formation of hydroxyapatite crystals), there was no clear association of SDRs and BHMTs with bone homeostasis.

The objective of this work was to characterize the role of these 3 genes in mechanisms underlying bone and cartilage formation and homeostasis by collecting basic data on gene expression (levels and sites of gene expression) and transcriptional regulation (functional analysis of promoter activity) using gilthead seabream as experimental model. A secondary objective of this work was to get insights into the molecular evolution of the 3 gene families throughout vertebrate evolution. The same approaches and tools will be applied to the study of the 3 genes. Levels of expression during extracellular matrix mineralization will be determined through qPCR (northern hybridization in the case of S100) to confirm and extend previous results (Fonseca et al., 2007; Tiago et al., 2011). Patterns of gene expression during development and in adult tissues will also be determined by qPCRs and sites of gene expression will be inferred from *in situ* hybridization in developing embryo and in selected adult tissues. The 5' flanking region of each of the tree genes will be analysed using *in silico* tools for the presence of binding sites for transcription factors previously associated with the regulation of bone and/or cartilage cell differentiation and mineralization and promoter activity will be evaluated using luciferase reporter constructs to test the functionality of these sites. The taxonomic distribution and molecular phylogeny of the vertebrate orthologs of gilthead seabream genes will be inferred from genomic data collected from sequence databases or cloned within the scope of this work. The occurrence of paralogs and/or alternative spliced variants for *sdr-like*, *S100-like* and *bhmt3* will also be assessed.

CHAPTER 2

GILTHEAD SEABREAM S100 GENE

CHAPTER 2. GILTHEAD SEABREAM *S100-LIKE* GENE

2.1 Identification of a new cartilage-specific S100-like protein up-regulated during endo/perichondral mineralization in gilthead seabream

Joana Rosa
Vera G. Fonseca
Vincent Laizé
Paulo J. Gavaia
M. Leonor Cancela

2.1.1 Abstract

Calcium ions and calcium-binding proteins play a major role in many cellular processes, in particular skeletogenesis and bone formation. We report here the discovery of a novel S100 protein in fish and the analysis of its gene expression patterns. A 648-bp full-length cDNA encoding an 86-amino acid S100-like calcium-binding protein was identified through the suppression subtractive hybridization of a gilthead seabream (*Sparus aurata*) cDNA library constructed to identify genes associated with *in vitro* mineralization. Deduced protein lacks an identifiable signal peptide and exhibits two EF-hands motifs characteristic of the S100 proteins. Phylogenetic and bioinformatic analyses of S100 sequences suggested that gilthead seabream protein represents a novel and fish-specific member of the S100 protein family. Expression of S100-like gene was up-regulated during the *in vitro* mineralization of bone-derived cell lines and during seabream development from larvae throughout adulthood, reflecting the formation of the skeleton. The restriction of S100-like gene expression to the chondrocytes of cartilaginous tissues undergoing endo/perichondral mineralization in juvenile fish further

confirmed the mineralogenic role of the protein in fish and emphasized the potential of S100-like as a marker of mineralizing cartilage in developing fish.

2.1.2 Introduction

Calcium ion Ca^{2+} plays a major role in cell physiology, where it functions as a signal for many cellular processes. As a consequence, calcium levels (but also transport and sensing) are tightly controlled, primarily by calcium-binding proteins belonging to the EF-hand (two calcium-binding loops flanked by two alpha helices) protein superfamily (Ikura, 1996) S100 proteins, which are small, acidic, dimeric proteins, form the largest subfamily of EF-hand proteins (S Bhattacharya et al., 2004). They translate physiological changes in calcium levels into specific cellular responses through two EF-hands: the first one is C-terminal, canonical (present in all EF-hand proteins) and has a high affinity for calcium; the second one is N-terminal, non-canonical (specific to S100 proteins) and has a lower affinity for calcium (Heizmann et al., 2003; Marenholz et al., 2004; Marenholz et al., 2006). Although EF-hands of S100 proteins exhibit different affinities for calcium ions, both of them efficiently mediate calcium-dependent responses in intracellular and extracellular compartments (Ravasi et al., 2004). The S100 protein family has 20 members described in human so far; they are highly similar but exhibit distinctive patterns of cell and tissue distribution (Donato, 2003; Heizmann et al., 2003), being involved in various cellular activities such as signal transduction, cell differentiation, gene transcription, calcium homeostasis and cell cycle progression (Heizmann et al., 2002; Heizmann et al., 2003). Most studies aiming at understanding molecular function(s) and cellular role(s) of S100 proteins have focused on the mammalian members of this family, although a growing number of S100-like proteins are being discovered in non-mammalian species, fish in particular (Bobe and Goetz, 2000; Cao et al., 2003; Di Pietro and Santome, 2002; Hsiao et al., 2003; Kraemer et al., 2008). We report here the discovery of a novel S100 protein in the teleost fish gilthead seabream [*Sparus aurata* L.] and present expression patterns from embryo throughout adulthood and during *in vitro* mineralization determined by Northern and *in situ* hybridizations.

2.1.3 Materials and methods

2.1.3.1 Cell culture and extracellular matrix mineralization

Gilthead seabream VSa16 and VSa13 cell lines were cultured at 33ºC in a 10% CO_2 humidified atmosphere using Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum, 1% penicillin-streptomycin, 1% fungizone and 2 mM of L-glutamine (Pombinho et al., 2004). For mineralization experiment, cells were seeded in 6-well plates at 10^5 cells per well and cultured in DMEM until confluence. Extracellular matrix (ECM) mineralization was induced in confluent cell cultures by supplementing medium with 50 μg/ml of L-ascorbic acid, 4 mM of $CaCl_2$ and 10 mM of β-glycerophosphate. Culture medium was renewed every 3.5 days for 4 weeks until cells were fixed and mineral deposited within ECM was revealed by von Kossa staining (Pombinho et al., 2004). At appropriate times, total RNA from mineralizing and control cells (3 replicates per condition) was prepared as described below. All cell culture reagents were from Invitrogen (Carlsbad, CA).

2.1.3.2 Culture of larvae, juvenile and adult fish

Fish were raised as previously described (Rafael et al., 2006). When sampled for RNA preparation, larvae and juvenile fish were euthanized with 500 ppm of 2-phenoxyethanol (Sigma-Aldrich, St. Louis, MO) then washed twice in phosphate-buffered saline solution (PBS) and stored at -80ºC in 5 ml of Trizol reagent (Sigma-Aldrich). Adult tissues were removed from anesthetized (200 ppm of 2-phenoxyethanol) then decapitated adult fish, and stored at -80ºC in 5 ml of Trizol reagent.

2.1.3.3 RNA preparation

Total RNA was extracted from cultured cells according to Chomczynski and Sacchi (Chomczynski and Sacchi, 1987) and from larvae, juvenile and adult tissues using Trizol following manufacturer's instructions. Total RNA concentration was determined using GeneQuant apparatus (Pharmacia Biotech, Uppsala, Sweden) and its quality was assessed on formaldehyde-agarose gels.

2.1.3.4 Northern blot analysis

Ten micrograms of total RNA were fractionated on a formaldehyde-agarose gel, and transferred to a Hybond-XL nylon membrane by capillary blotting with 10× standard saline citrate solution (SSC; 1×SSC is 0.15 M NaCl and 15 mM sodium citrate, pH 7.0). Membranes were probed with [α-^{32}P]-dCTP radiolabeled *S. aurata S100-like* or ribosomal protein L27a (*rpl27a*) cDNAs using the Rediprime II kit (Amersham Biosciences, Carnaxide, Portugal) and purified from unincorporated radionucleotides using Microspin S-200 HR columns (Amersham Biosciences). Membranes were hybridized overnight at 42°C in ULTRAhyb solution (Ambion, Austin, TX) and washed 2×5 min in a low stringency solution (2×SSC, 0.1% SDS) and 2×15 min in a high stringency solution (0.1×SSC, 0.1% SDS) at 55°C. Membranes were then exposed to a Kodak XAR film (Amersham Biosciences). Signal intensity was estimated by densitometry methods using Quantity One software (Bio-Rad, Hercules, CA). Relative expression of *S100-like* gene was normalized with *rpl27a* signals.

2.1.3.5 Tissue preparation for *in situ* hybridization and histological analysis

Fish at 38 dph were fixed for 24 h at 4ºC in 4% paraformaldehyde (PFA; pH 7.4 with PBS), washed 3×10 min with PBS, then dehydrated through a PBS/methanol gradient and stored in 100% methanol at 4ºC. When appropriate, fish were passed through an increasing methanol/xylol series, embedded in paraffin then cross-sectioned. Sections (7-µm thick) were collected on TESPA (3-aminopropyltriethoxysilane, Sigma-Aldrich) coated slides, dried for 4 h at 37ºC and kept at 4ºC until used.

2.1.3.6 *In situ* hybridization

Sense and antisense RNA probes were generated from 1 µg of linearized plasmid containing *S100-like* complete cDNA (642 bp) using T7 or SP6 polymerases then labelled with digoxigenin-dUTP (DIG RNA labeling kit, Roche Diagnostics, Mannheim, Germany). Riboprobes were treated with RNase-free DNase, recovered by ethanol precipitation and their integrity was assessed through agarose gel electrophoresis. Preparation of sections and *in situ* hybridization were performed using a modified protocol previously described by Pinto et al. Pinto et al. (2001, 2003). Briefly, sections were hybridized at 68ºC for 12-16 h with 100-200 ng

of riboprobe, then incubated with a 1:2000 dilution of anti-digoxigenin-alkaline phosphatase antibody (Roche Diagnostics) and signal was revealed using NBT/BCIP (Sigma-Aldrich) as substrate.

2.1.3.7 Histology

Routine histological staining was performed using trichromic azan combination (Chroma-Waldeck, Münster, Germany) to identify tissues and cellular structures. In brief, tissues were incubated for 20 min with azan combination, differentiated for 20 s in isopropanol then washed in double-distilled water, dehydrated and mounted. Images were acquired using an Olympus IX81 microscope equipped with an Altra 20 camera. ISH images were captured using DIC contrast and azan sections using bright field. Skeletal tissues were classified according to Benjamin (1989) and Benjamin et al. (1992) for cartilage types, and Faustino and Power (2001) and Genten et al. (2009) for bony tissues.

2.1.3.8 Sequence collection and reconstruction

GenBank sequence database was searched for EST (expressed sequence tag) and WGS (whole genome shotgun) sequences showing similarity with seabream S100-like sequence using BLAST facilities at the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov). Species-specific sequences were aligned then assembled using ContigExpress module of Vector NTI Advance software (Invitrogen) to generate highly accurate consensus sequences. Virtual transcripts were deduced by comparative analysis from joined consensus sequences using stringent overlap criteria.

2.1.3.9 Sequence alignment and analysis

Collected sequences were aligned using T-Coffee (Notredame et al., 2000) with parameters set to default, to produce, after manual adjustment, a high quality alignment. Sequence logos – graphical displays, where the height of each letter is made proportional to its frequency – were created from T-Coffee multiple alignment using WebLogo (Schneider and Stephens, 1990). Signal peptide, *O*- and *N*-linked glycosylation sites were predicted using SignalP (Bendtsen et al., 2004), NetOGlyc and NetNGlyc (Julenius et al., 2005), respectively. Protein domains were identified using InterProScan (Zdobnov and Apweiler, 2001).

Percentage of protein identity was calculated using the Sequence Manipulation Suite available at www.bioinformatics.org. Neighbor-joining tree was built from T-Coffee multiple sequence alignment using MEGA3 (Foster et al., 2004).

2.1.4 Results

2.1.4.1 Reconstruction of gilthead seabream *S100-like* cDNA

The full-length cDNA of a gene recently shown to be up-regulated during the *in vitro* mineralization of *S. aurata* VSa16 pre-osteoblast cell line was reconstructed from various cDNA fragments collected from a subtractive library (Fonseca et al., 2007) (**Supplementary Figure 2.1** and GenBank accession No. AY787209). Comparison of the deduced protein sequence with sequences available in GenBank revealed its similarity to S100 calcium-binding proteins. This similarity was further confirmed by the presence of two EF-hands characteristic of S100 proteins, i.e. a N-terminal non-canonical EF-hand containing a 14-aa calcium-binding loop (residues 18-32) with two helix loops (residues 10-19 and 33-41), and a C-terminal canonical EF-hand (residues 59-71) containing a 12-aa calcium-binding loop with two helix loops (residues 51-58 and 72-80). While features typical of S100 proteins, e.g. hydrophobic residues and sites for serine and threonine phosphorylation, were identified in *S. aurata* S100-like protein (**Supplementary Figure 2.1**), no signal peptide or proteolytic cleavage sites were found.

2.1.4.2 Seabream S100-like protein is a novel member of S100 family

In order to better characterize the relationship of *S. aurata* protein with other vertebrate S100 proteins, a phylogenetic tree was built from the alignment of *S. aurata* S100-like sequence and 41 GenBank S100 reference sequences representing mammals (human), birds (chicken), amphibians (Xenopus) and fish (zebrafish, Japanese medaka, Japanese flounder, killifish, mouth-brooders and chilotes) (**Figure 2.1**). S100-like sequence did not cluster with any known S100 sequences (e.g. S100A1-15, S100B, G,H,Z and ictacalcin) but showed similarity to other fish S100-like sequences, suggesting that this protein might represent a new member of the S100 family with orthologs in other fish.

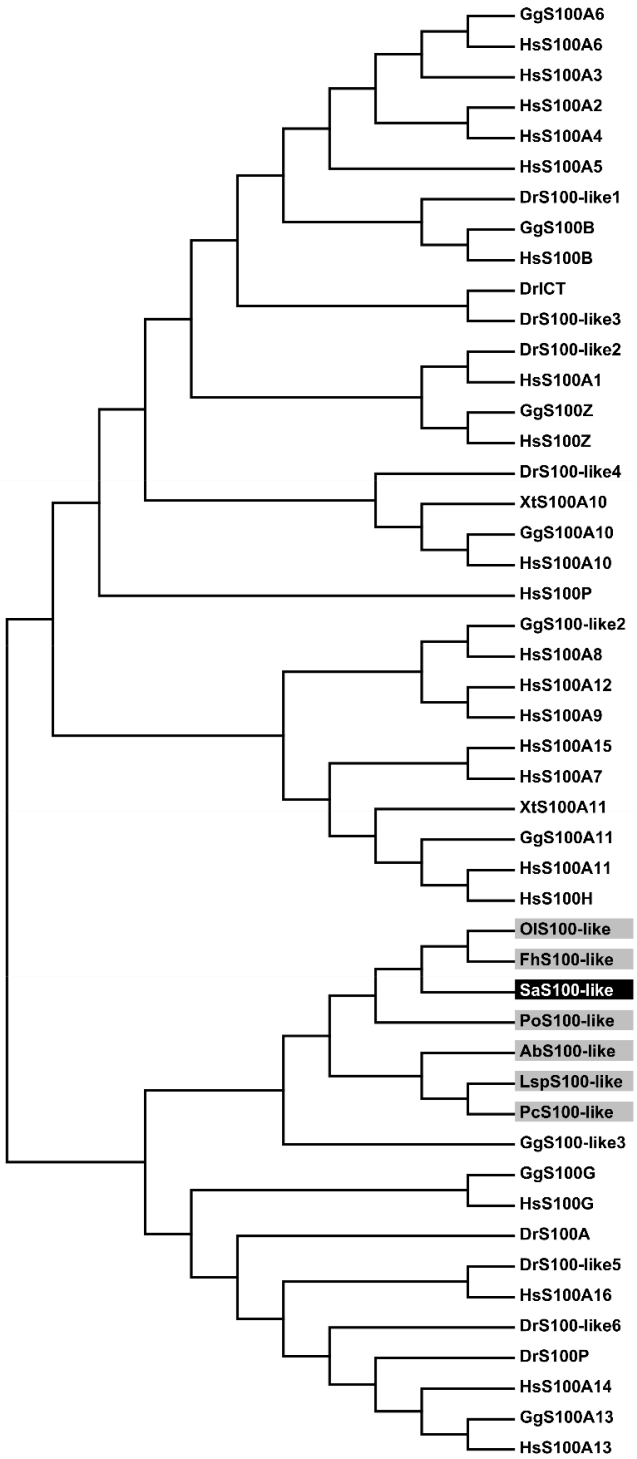


Figure 2.1. Unrooted phylogenetic tree of S100 protein family. Tree is based on the alignment of seabream S100-like sequence and 41 GenBank reference sequences representing mammals (human, Hs), birds (chicken, Gg), amphibians (xenopus, Xt) and fish (zebrafish, Dr) taxa. Neighbor-joining tree was built from T-Coffee multiple sequence alignment (Wallace et al., 2006) using Mega (Foster, 2004) and the Dayhoff distance matrix. Ol, *Oryzias latipes* (Japanese medaka); Fh, *Fundulus heteroclitus* (killifish); Sa, *Sparus aurata* (gilthead seabream); Po, *Paralichthys olivaceus* (Japanese flounder); Ab, *Astatotilapia burtoni* (Burton's mouthbrooder); Lsp, *Lipochromis sp.* (mouthbrooders); Pc, *Paralabidochromis chilotes* (chilotes).

2.1.4.3 Signature sequence of fish S100-like proteins

To reveal common features between *S. aurata* S100-like protein and proposed orthologs (all from fish; **Supplementary Figure 2.2**), peptide sequences were aligned then displayed as sequence logos (**Figure 2.2**), where conserved amino acid residues are identified by larger characters. The consensus/signature sequence of fish S100-like proteins revealed two highly conserved helix-loop-helix calcium-binding domains (EF hands), known to be involved in dimerization and binding to target proteins (Rety et al., 2000; Rintala-Dempsey et al., 2008). The non-canonical EF-hand is formed by the calcium-binding loop 1 flanked by helices I and II, whereas the canonical EF-hand is formed by calcium-binding loop 2 flanked by helices III and IV.

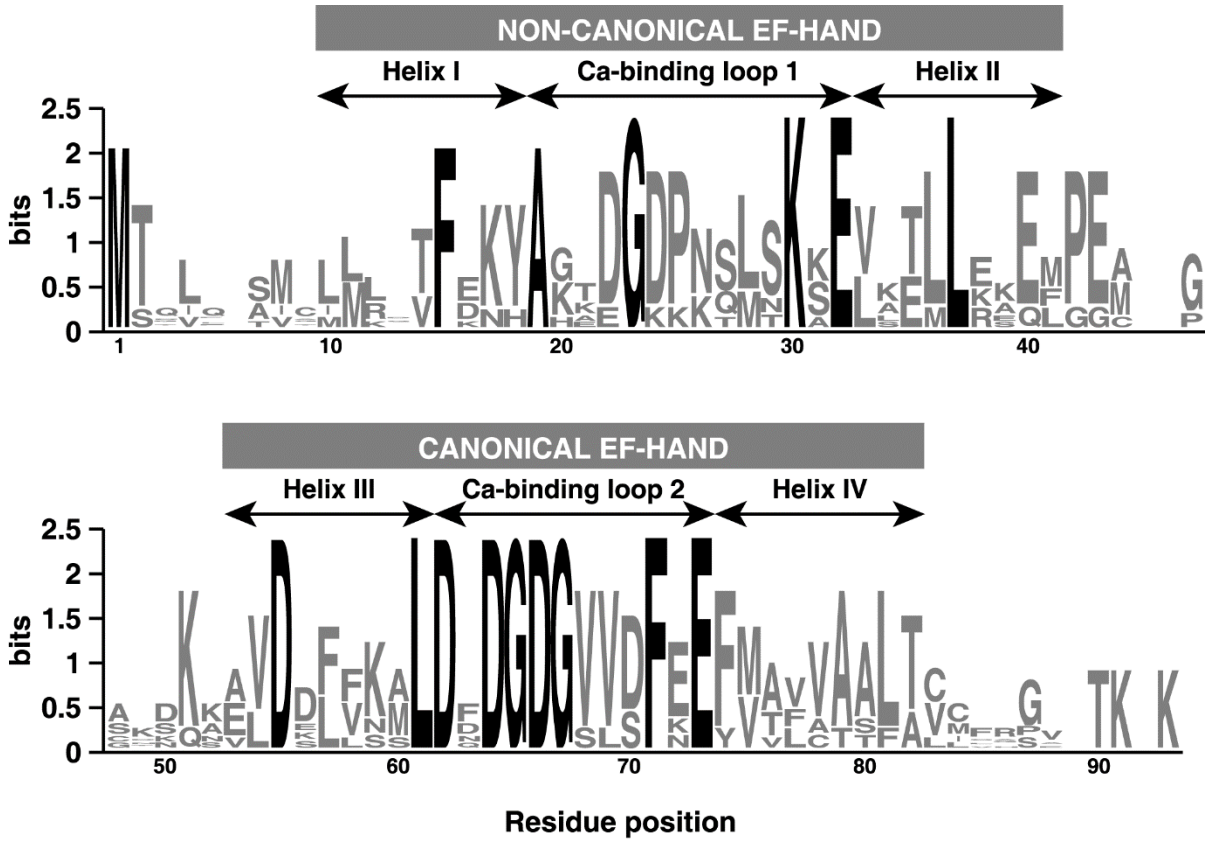


Figure 2.2. Fish S100-like sequence logos. The height of each letter is directly proportional to its frequency. 100% conserved residues are shown in black. Logos were created from T-Coffee multiple sequence alignment (Schneider and Stephens, 1990). C-terminal canonical and N-terminal non-canonical EF-hands (and respective secondary structures, e.g. helix and loop) are indicated above the sequence logos.

2.1.4.4 Expression of gilthead seabream, *S100-like* during *in vitro* mineralization

Levels of *S100-like* gene expression were determined by Northern hybridization (NH) using RNA from pre-osteoblast VSa16 and pre-chondrocyte VSa13 cell lines cultured under mineralizing conditions or left untreated for 4 weeks (**Figure 2.3**). While gene expression was basal in control cells, it progressively increased in mineralizing cells reaching highest levels at 4 weeks (approximately 4 times the levels in control cells). Up-regulation of S100-like gene expression during *in vitro* mineralization was later confirmed through microarray analysis with a fold stimulation of 4.4 and 8.1 in VSa16 and VSa13 cells, respectively (Tiago et al., 2011).

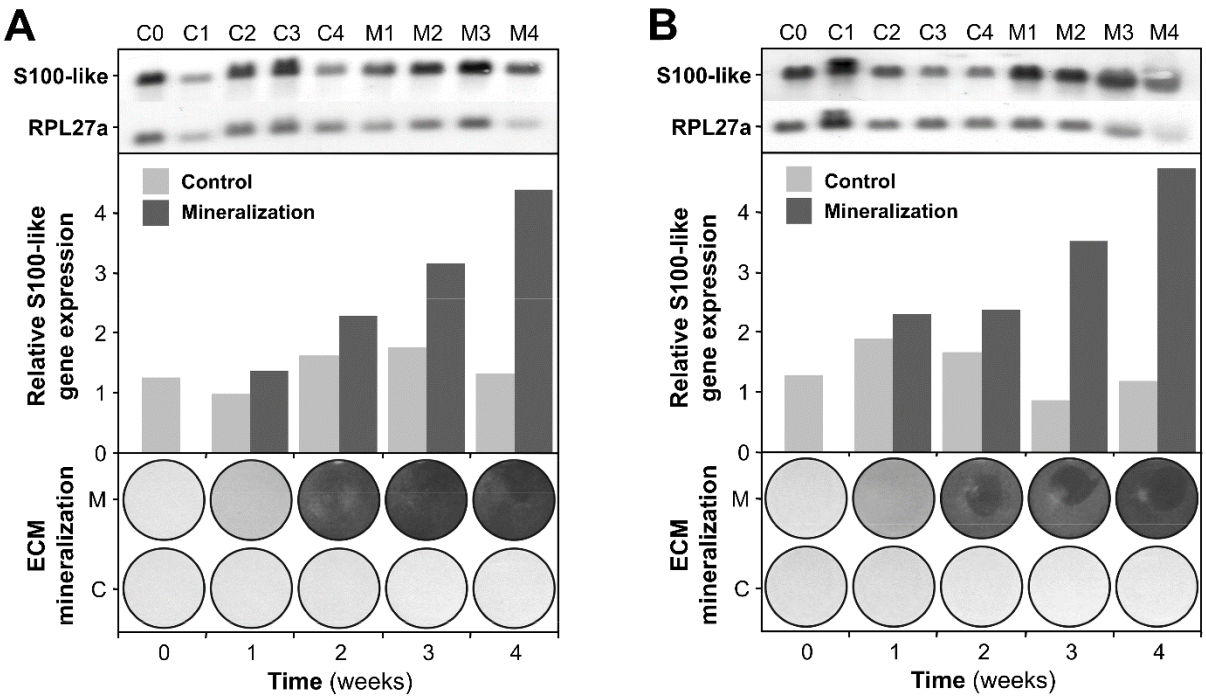


Figure 2.3. Relative S100-like gene expression and extracellular matrix mineralization in gilthead seabream VSa13 (**A**) and VSa16 (**B**) cells cultured for 4 weeks under mineralogenic conditions. Top panels, S100-like and RPL27a gene expression by Northern hybridization. Middle panels, relative S100-like gene expression determined by densitometry analysis and normalized with RPL27a. Bottom panels, pictures of von Kossa's stained cells cultured in a 6-well plate; mineral nodules are colored in black. C, control; M, mineralized.

2.1.4.5 Developmental patterns of gilthead seabream *S100-like* expression

A major objective of this study was to determine the spatial and temporal expression of S100-like gene and identify the cellular types responsible for its expression to get insights into the role of S100-like protein during *S. aurata* development from embryo throughout adulthood. Levels of S100-like gene expression were determined from Northern hybridization of total RNA prepared from embryo, larvae and juvenile fish. Transcript was first detected in gastrulated embryo at the onset of *de novo* transcription, at approximately 10 hours post fertilization (hpf) and is therefore not maternally inherited. Levels of expression remained basal during late embryonic development (until 20-24 hpf) then progressively increased throughout larval development reaching a maximum at 91 days post hatching (dph) in juvenile fish (**Figure 2.4)**.

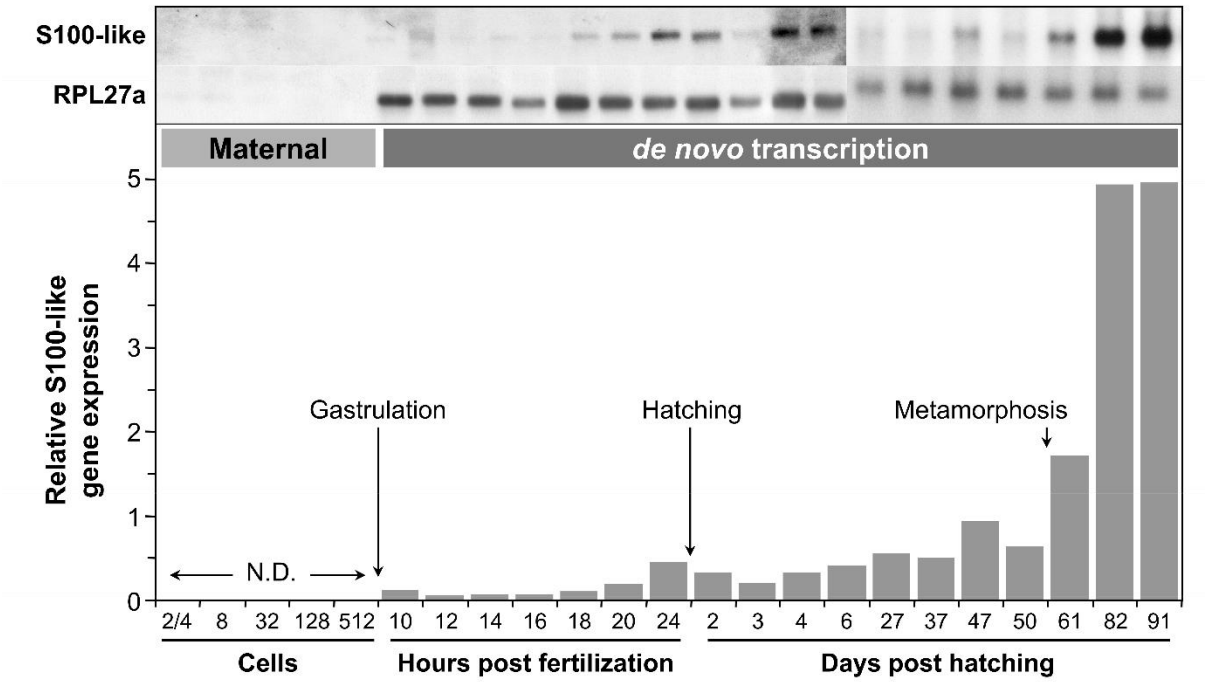


Figure 2.4. Relative S100-like gene expression during gilthead seabream development. Top panel, S100-like and RPL27a gene expression by Northern hybridization. Bottom panel, relative S100-like gene expression determined by densitometry analysis and normalized with RPL27a. N.D., not detected.

S100-like gene expression was also investigated in a broad variety of adult seabream tissues including soft and calcified tissues (**Figure 2.5**). S100-like was highly expressed in fin, branchial arches and vertebral column samples (calcified cartilage and bony tissues) suggesting

a possible involvement in cartilage development and in the regulation of tissue mineralization during endochondral ossification. Although to a lesser extent, S100-like expression was also detected in non-mineralized tissues (gall bladder, gonad, aorta, kidney and heart), suggesting a more diversified role for S100-like protein that may be linked to cellular regulation of calcium levels.

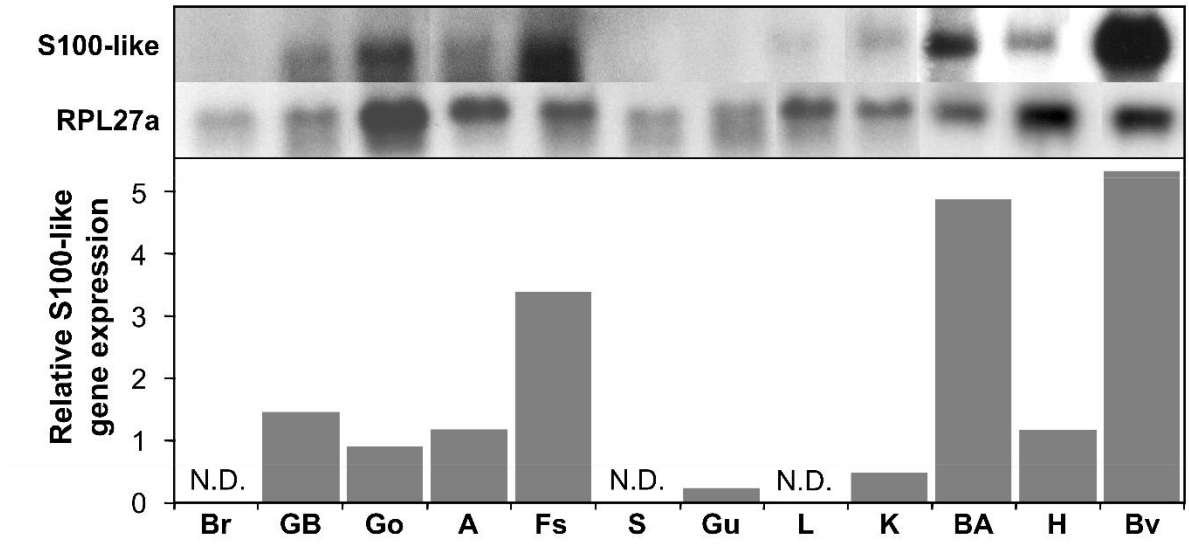


Figure 2.5. Relative S100-like gene expression in adult gilthead seabream tissues. Top panel, S100-like and RPL27a gene expression by Northern hybridization. Bottom panel, S100-like relative gene expression determined by densitometry analysis and normalized with RPL27a. N.D., not detected. Br, brain; GB, gall bladder; Go, gonad (ovary/testis); A, aorta; Fs, caudal fin-soft rays; S, spleen; Gu, gut; L, liver; K, kidney; BA, branchial arches; H, heart; Bv, bone-vertebra.

To confirm these data, sites of S100-like gene expression were determined by *in situ* hybridization (ISH) in a 38-dph juvenile *S. aurata* (**Figure 2.6**). Consecutive sections were stained with azan trichrome to help distinguish the different tissue and structures. A strong and specific signal was observed in most cartilaginous structure undergoing endochondral and perichondral ossification, namely skull (e, e´), palatoquadrate (f, f'), otic capsules (g, g', j, j'), trabecula communis (h, h'), meckel's cartilage (i, i'), trabecula (k, k'), branchial arches (zellknorpel cartilage; l, l') and neural arches (m, m'). Hypertrophic chondrocytes were strongly marked in most of these cartilaginous structures, while immature and proliferative chondrocytes less intensely positive. In the trabecula (k, k'), chondrocyte expression profile was inverted, with immature and proliferative chondrocyte marked, but not the hypertrophic ones.

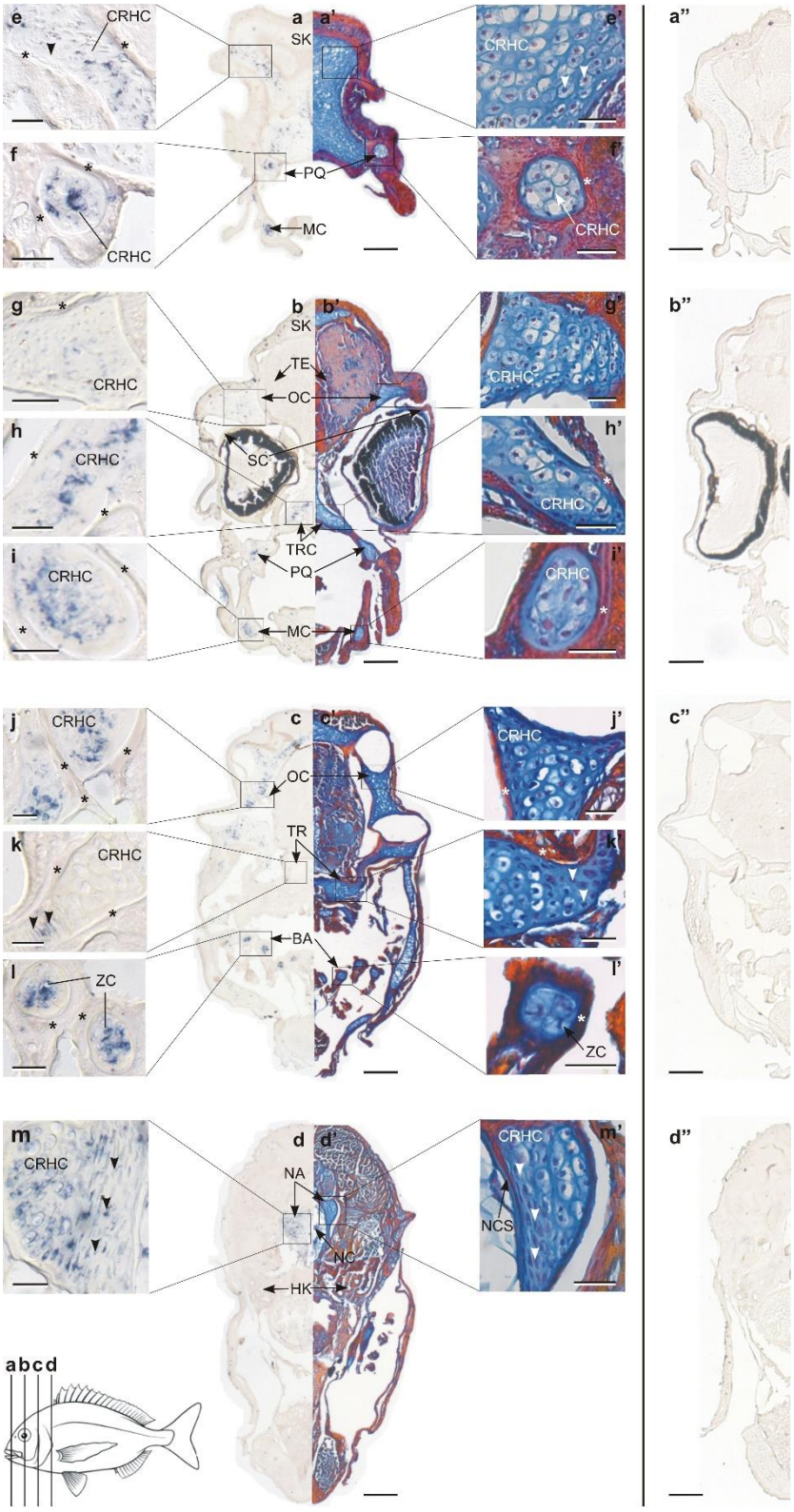


Figure 2.6. Sites of *S100-like* gene expression by *in situ* hybridization (ISH) in a 38-dph gilthead seabream. a-m, ISH using *S100-like* antisense riboprobe. a'-m' , azan staining. a''-d'' , ISH using *S100-like* sense riboprobe. BA, branchial arches; CRHC, cell-rich hyaline cartilage; EPC, epiphysial cartilage; HK, head kidney; MC, Meckel's cartilage; NA, neural arch; NC, notochord; NCS, notochord sheet; OC, otic capsules; PQ, palatoquadrate; SC, scleral cartilage; TE, telencephalon; TR, trabecula; TRC, trabecula communis; TTM, taenia tecti medialis; ZC, zellknorpel cartilage. Asterisks indicate bone.

No signal was observed in the unossified scleral cartilage (b, b') and pterygiophori of the dorsal fin (result not shown). Sites of gene expression by ISH were consistent with expression data collected by NH, e.g. strong ISH signal in the cartilage tissue of neural arches and zellknorpel cartilage, and high expression levels in vertebral column (possibly only the cartilage at the centre of the arches) and branchial arches, respectively. Similarly, the absence of expression in brain tissue was confirmed by ISH. Although detected by NH in aorta and heart, S100-like transcript was not detected in these tissues by ISH; different detection efficiency in the methods used to detect S100-like transcript (NH being more sensitive than ISH) or different developmental stages analyzed (juvenile *versus* adult specimens) may explain this discrepancy. Finally, gall bladder, gonads, caudal fin-soft rays, spleen, gut and liver were not present on sections used for ISH and gene expression detected by NH could not be further confirmed or localized.

2.1.5 Discussion and Conclusions

2.1.5.1 A fish-specific subfamily of S100 proteins

Although highly similar and with a structure typical of S100 calcium-binding proteins, *S. aurata* S100-like protein – and other fish orthologs identified *in silico* – could not be associated with any of the twenty S100 protein subfamilies and may represent a novel isoform. This isoform is probably fish-specific since no ortholog has been identified in other vertebrates and may have arisen from a gene duplication event that occurred in the ray-finned fish (Actinopterygii) lineage after branching from tetrapods. Although we cannot exclude that it arose from a more recent and/or lineage-specific (all fish species expressing S100-like gene are Percomorpha) gene duplication event, we propose that S100-like originated through the well supported whole genome duplication event that occurred 450 million years ago in the fish lineage (the 3R hypothesis) and is at the origin of many fish paralogous genes (Vanderpoele et al., 2004; Jaillon et al., 2004; Meyer and Van de Peer, 2005; Nakatani, Takeda, et al., 2007; Siegel et al., 2007). One could ask whether fish S100-like protein is still capable of binding calcium ions. Probably yes, since it contains key structural features of S100 proteins, i.e. two EF-hands (Shibani Bhattacharya et al., 2004; Fritz and Heizmann, 2004; Heizmann et al., 2002; Heizmann et al., 2003), responsible for calcium binding and characteristic hydrophobic residues in helices I and IV responsible for protein dimerization (Kraemer et al., 2008; Rintala-

Dempsey et al., 2008) and crucial for protein function (Fritz and Heizmann, 2004). However, this hypothesis would need to be confirmed by functional assays.

2.1.5.2 A role for S100-like protein in development and tissue ossification

Onset of S100-like gene expression occurred at hatching and transcript levels increased progressively thereafter, suggesting an increasing role for the protein during fish development. It has been proposed that calcium-binding proteins are mainly involved in calcium storage during vertebrate development, and that this storage mainly profit to skeletal growth and calcium-based crystal formation (Tuan and Scott, 1977). As they develop, vertebrates require larger amounts of calcium and more storage capacities, e.g. to build new bone and consolidate existing skeleton. Levels of seabream S100-like transcript could therefore increase as a consequence of skeletogenesis and associated demand in calcium ions. This hypothesis, speculative at this stage, was further supported by the strong expression of S100-like gene in juvenile cartilaginous tissues undergoing perichondral (e.g. Meckel's cartilage) and endochondral (e.g. trabecula) mineralization and in adult calcifying tissues (vertebrae and branchial arches), and the absence of expression in non-mineralizing cartilaginous tissues. While most S100 proteins are expressed at low levels (Fritz and Heizmann, 2004), gilthead seabream S100-like transcript was found to be rather abundant in mineralizing tissues, indicative of a key role for S100-like protein during osteogenesis, as already reported for other calcium-binding proteins in mammals (Balmain et al., 2003; Berdal et al., 1996; Faucheux et al., 1998). Up-regulation of S100-like gene expression in pre-chondrocyte and pre-osteoblast cells undergoing ECM mineralization provided additional support for this hypothesis. The progressive increase of S100-like transcripts throughout the mineralization period (last 3 weeks of treatment) further indicate a role in the regulation of extracellular matrix mineralization and/or crystal deposition rather than in the cell differentiation process (first week of treatment). An active role of S100-like protein in tissue mineralization implies the protein to be secreted within the extracellular matrix of bone cells. However, no identifiable sequence for a signal peptide, i.e. a signal that could mediate secretion, could be detected in seabream S100-like protein sequence. This may however not prevent S100-like protein secretion since other S100 proteins, also lacking a recognizable signal peptide, occur extracellularly through a mechanism presently poorly understood (Donato, 2001) involving cell surface receptors (Hofmann et al.,

1999; Leclerc et al., 2009). We hypothesize that seabream S100-like protein may be targeted to the extracellular matrix of seabream bone cells through a similar mechanism.

Altogether, results presented in this study are indicative of a novel fish-specific S100 calcium-binding protein with a role in development, possibly skeletogenesis, and tissue mineralization, possibly osteogenesis and/or chondrogenesis. Its role, still to be confirmed, could be to provide calcium ions necessary to the formation of calcium-based hydroxyapatite crystals.

2.1.6 Acknowledgments

This work was partially funded by grants POCTI/BCI/48748/2002 from the Portuguese Science and Technology Foundation (including funds from FEDER and OE) and GOCE-CT-2004-505403 from the European Commission under the 6th Framework Program (Marine Genomics Europe network of Excellence). VGF was partially supported by CCMAR funding. JR was supported by a doctoral grant (SFRH/BD/47433/2008) from the Portuguese Science and Technology Foundation (FCT). We thank the two anonymous reviewers whose suggestions have helped us improve our manuscript substantially.

2.2. Ets1 regulates the transcription of cartilage-specific S100 protein in gilthead seabream

Joana Rosa
M. Leonor Cancela
Vincent Laizé

2.2.1 Abstract

A novel S100 calcium-binding protein has been recently identified in teleost fish. Its expression is restricted *in vivo* to chondrocytes of cartilaginous tissues undergoing endo/perichondral mineralization and its function has been associated *in vitro* with mechanisms of extracellular matrix mineralization. To get more insights into this mineralogenic role, the transcriptional regulation of S100 gene was investigated using luciferase reporter constructs. The occurrence of several silencers was revealed through the analysis of the basal activity of promoter constructs in HEK-293 cells. Among those, a silencer located in the region -883/-768 had the capacity to shut down completely the activity of S100 gene promoter. The presence of several putative binding sites for Ets1, a transcription factor regulating the expression of several cartilage-related proteins, was predicted through *in silico* analysis. Analysis of luciferase activity in cells expressing zebrafish Ets1a revealed that regions -636/-513 and -82/+62 contain active Ets1 binding sites and decrease luciferase activity upon mutation of specific sites confirmed the effectiveness of Ets1 binding at positions -552/-539 and -517/-501. In conclusion, this work provided novel evidence for the transcriptional regulation of fish cartilage-specific S100 protein by Ets1.

2.2.2 Introduction

Calcium ion and S100 calcium-binding proteins are critical players in cell physiology. While the former is a universal, intracellular second messenger that functions as a signal for many cellular processes, the latter translate physiological changes in calcium levels into specific cellular responses (Donato, 2001). S100 proteins are small (10-12 kDa), highly

Homologous, acidic proteins that form non-covalent homo or heterodimers (Shibani Bhattacharya et al., 2004). Each S100 protein monomer contains two EF-hand structures with different affinities for calcium binding (Heizmann et al., 2002; Marenholz et al., 2004; Marenholz et al., 2006). This distinctive feature, together with subtle differences in amino acid sequence and in expression profiles, has resulted in specific and non-redundant functions of S100 proteins (Leśniak, 2011). These proteins have long been associated with chondrocytes and chondrogenesis in mammals (Stefansson et al., 1982; Ushigome et al., 1984; Chano et al., 1995; Li et al., 2002), where they would be involved in cartilage repair and calcification (Mohr et al., 1985). Consensus binding sites for transcription factors normally required for chondrocyte differentiation and cartilage development, e.g. specificity protein 1 (SP1), spleen focus forming virus (SFFV) proviral integration oncogene (PU.1) and activator protein 1 (AP-1) (Chadjichristos et al., 2003; Karreth et al., 2004; Ravasi et al., 2004), were identified in the promoter of several human and mouse S100 genes through *in silico* analysis (Ravasi et al., 2004; Gebhardt et al., 2006) but these predictions have not been functionally confirmed yet. We recently reported the discovery, in gilthead seabream *Sparus aurata*, of a novel fish-specific S100 calcium-binding protein, whose expression was restricted *in vivo* to chondrocytes of cartilaginous tissues undergoing endo/perichondral mineralization and associated *in vitro* with mechanisms of extracellular matrix mineralization (Fonseca et al., 2011). The presence of consensus binding sites for Ets1 (v-ets erythroblastosis virus E26 oncogene homolog 1), a transcription factor regulating chondrogenesis (Gao et al., 2010; Sugiura and Ito, 2010), predicted in the 5' flanking region of gilthead seabream S100 gene. We intended to evaluate, within the scope of this work, the functionality of these binding elements and have collected evidence towards a role of Ets1 in the control of gilthead seabream S100 gene transcription.

2.2.3 Materials and methods

2.2.3.1 Cloning of S100 gene promoter

The promoter region of gilthead seabream S100 gene was amplified from a GenomeWalker genomic DNA library (*Stu*I; Clontech) using Advantage 2 Polymerase mix and 0.2 µM of adaptor primer 1 and gene-specific primer S100_Rv1. Nested PCR was performed using a 1:50 dilution of the first PCR and 0.2 µM of adaptor primer 2 and gene-specific primer S100_Rv2. PCR products were separated on agarose gel, purified using GFX Gel Band

Purification kit (Amersham Biosciences) then cloned into pCRII-TOPO (Invitrogen) and sequenced. PCR primers are listed in **Supplementary Table 2.1**.

2.2.3.2 S100 gene promoter constructs

Promoter fragments were amplified by PCR using reverse primer S100_Rv3 in combination with forward primers S100_Fw1 (construct C1 -883/+62LUC), S100_Fw2 (construct C2 -768/+62LUC), S100_Fw3 (construct C3 -636/+62LUC), S100_Fw4 (construct C4 -513/+62LUC), S100_Fw5 (construct C5 -237/+62LUC), S100_Fw6 (construct C6 -142/+62LUC), and S100_Fw7 (construct C7 -82/+62LUC), then digested with appropriate endonucleases (*Xho*I/*Hin*dIII) and cloned into pGL4.10 vector (Promega) upstream of firefly luciferase reporter gene. PCR primers are listed in **Supplementary Table 2.1**.

2.2.3.3 Cell culture and DNA transfection

Human embryonic kidney (HEK-293) cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Life Technologies) supplemented with 10% fetal bovine serum (Sigma-Aldrich), 1% penicillin-streptomycin (Life Technologies), 1% L-glutamine (Life Technologies) and maintained in a 5% CO_2 humidified atmosphere at 37°C. The day before transfection cells were seeded in 24-well plates (Sarstedt) at 5×10^4 cells per well in order to achieve cultures at 50-60% confluence. DNA constructs and empty pGL4.10 vector (125 ng) were delivered into HEK-293 cells using X-tremeGENE HP transfection reagent (Roche). When appropriate, 25 ng of pCMX-Ets1 – the expression vector containing zebrafish *Danio rerio ets1a* gene (GenBank accession no. KF774190) under the control of CMV promoter – was cotransfected with selected constructs of S100 gene promoter. Cells were incubated for 48 h, lysed and luciferase activities measured using Dual-Luciferase Reporter Assay system (Promega) in a Synergy 4 microplate reader (BioTek). pRL-null vector (Promega), which expresses *Renilla* luciferase (Rluc) and lacks promoter and enhancer elements, was used in all transfections (50 ng) to normalize the firefly luciferase (FLuc) activity. Results are expressed as fold changes over the promoter-less value and represent the mean ± SD of at least four independent experiments performed in duplicates.

2.2.3.4 *In silico* analysis of S100 gene promoter

Potential *cis*-regulatory elements, i.e. transcription factor-binding sites, were predicted in the 5' flanking region of S100 gene using MatInspector (Cartharius et al., 2005) at www.genomatix.de, Transfac (Matys et al., 2003) at www.gene-regulation.com/index2, TFsearch (Heinemeyer et al., 1998) at www.cbrc.jp/research/db/TFSEARCH and Jaspar (Bryne et al., 2008) at asp.ii.uib.no:8090/cgi-bin/jaspar2010/jaspar_db.pl.

2.2.3.5 Mutagenesis of Ets1 binding sites

Sequence of putative Ets1 binding sites was mutated using QuikChange Lightning Site-Directed Mutagenesis kit (Agilent). Nucleotides G and A at position -547/-545 were substituted with nucleotides T and C, respectively using promoter-specific primers S100_Ets1_1_Fw and S100_Ets1_1_Rv. Nucleotides G and A at position -510/-508 were substituted with nucleotides T and C using promoter-specific primers S100_Ets1_2_Fw and S100_Ets1_2_Rv. All mutations were confirmed by sequencing. PCR primers are listed in **Supplementary Table 2.1**.

2.2.4 Results and Discussion

2.2.4.1 Identification of transcription factor binding elements in S100 gene promoter

The 5′-flanking region of gilthead seabream S100 gene was analyzed using MatInspector, Transfac, TFsearch and Jaspar online tools to identify putative binding motifs for transcription factors (TF) involved in the regulation of S100 gene expression. In addition to the canonical TATA box (TATAAA) located at -26/-20 bp, various consensus sequences for sites associated with transcription factor Ets1 were identified (**Figure 2.7**; motifs are numbered according to the transcription start site (+1) currently available in GenBank). Ets1 recognizes a purine-rich core motif 5'-GGA(A/T)-3', through the 'Ets' domain, a winged helix-turn-helix DNA-binding domain characteristic of Ets family members (Graves and Petersen, 1998; Oikawa and Yamada, 2003). These predictions included four Ets1 sites identified by the four different programs (scored 1) at positions -552/-539, -517/-501, -440/-425 and -158/-143, one site

predicted by three programs (scored 0.75) at position -121/-108, and eight sites predicted by only one program (scored 0.25) at positions -772/-757, -754/-739, -640/-625, -386/-371, -248/-233, -221/-200, -106/-87, -6/+9. A region located at -703/-637 and containing 16 repeats of AAAG motif was also identified in the positive strand.

2.2.4.2 Activity of S100 gene promoter in HEK-293 cells

In order to test the functionality of Ets1 binding elements predicted in gilthead seabream S100 gene promoter, a 944-pb fragment of genomic DNA containing the 5' flanking region of S100 gene was amplified by PCR using gene-specific primers (**Supplementary Table 2.1**) and cloned into pGL4 vector upstream the firefly luciferase reporter gene (construct C1, **Figure 2.7**).

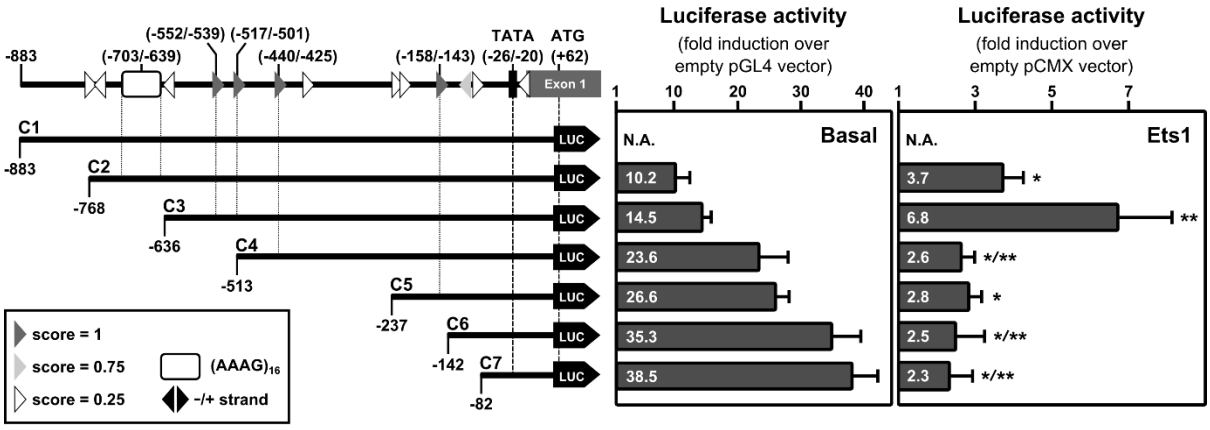


Figure 2.7. Transcriptional activity of gilthead seabream S100 gene promoter constructs in HEK-293 cells. Luciferase activity was measured either under basal conditions (and presented as fold changes over the empty pGL4 vector) or upon overexpression of zebrafish Ets1a transcription factor (and presented as fold changes over the empty pCMX expression vector). Promoter constructs C1-C7 are shown on the left. Nucleotide positions are numbered according to transcription start site (+1) currently available in GenBank. Putative Ets1 binding sites are represented as colour- and direction-coded triangles. Luciferase activity (FLuc/RLuc; $n \geq 5$) is presented as fold induction over the activity of promoterless pGL4.10 basic vector or empty pCMX-PL2 for co-transfections. N.A.: not active. Asterisks indicate values statistically different from C3 (*) and C2 (**) constructs (one-way ANOVA with Tukey's post-test; $P<0.05$).

Additional reporter vectors were subsequently constructed by truncating the 5' end of the initial promoter fragment (constructs C2-C7, **Figure 2.7**). Each construct was transiently transfected into HEK-293 cells and promoter activity was assessed as a function of luciferase

activity (ratio FLuc/RLuc) over the promoterless luciferase plasmid pGL4.10 basic. The suitability of skeletal cell lines – gilthead seabream mineralogenic ABSa15 and mouse chondrocyte-like ATDC5 cell lines – to host the different constructs was first tested but luciferase signals, in particular those of the *Renilla* luciferase expressed from the promoter-less pRL-null vector and used to normalize firefly luciferase signals, were weak. *Renilla* luciferase signals measured in HEK-293 cells were stronger and not subjected to TF regulation during co-transfection experiments. While luciferase activity was hardly detectable in HEK-293 cells transfected with construct C1 (-883/+62LUC), it was stimulated 10.2 fold upon transfection with construct C2 (-768/+62LUC). The absence of activity for C1 was further confirmed in other cell lines (i.e. gilthead seabream ABSa15 cell line, previously shown to express S100 transcript, and in mouse ATDC5; data not shown), suggesting the presence in the 115-pb fragment upstream position -768 of silencers capable of suppressing entirely the transcriptional activity of S100 gene promoter. Inactivity of C1 construct may also be related to the presence of a satellite sequence of 64-bp positioned at -703/-639 and constituted of 16 repeats of the tetranucleotide AAAG. AAAG repetitions have been associated *in vitro* with the repression of parathyroid hormone receptor 1 (PTHR1) gene promoter activity in human osteoblast-like Saos-2 and HOS cells (Minagawa et al., 2000; Minagawa et al., 2002); differences in repeat number were associated *in vivo* with differential promoter activity resulting in variations in adult height and bone mass density (Scillitani et al., 2006). Moreover, tetranucleotide repeats can form non-B DNA secondary structures under conditions of negative supercoiling of the DNA. $(AAAG)_n$, called also mirror repeats, which can form triplex DNA *in vitro* by folding back of a single strand into the major groove of repeat DNA (Slebos et al., 2002). Although the relevance of this satellite sequence for S100 gene promoter activity requires further studies, we propose that the 115-pb fragment present in C1, but not in C2, may interact with the satellite sequence and, either by repression or formation of secondary structures, prevent transcription. Sequential deletions of the promoter, ranging from -636/+62 (C3) to -82/+62 (C7), caused a gradual increase in promoter activity – from 14.5 to 38.5 folds, respectively – suggesting the presence of silencers in the region -768/-82, as already reported in the promoters of mammalian S100 genes (Leśniak et al., 2000; Leśniak, 2011).

2.2.4.3 Zebrafish Ets1a regulates the transcription of gilthead seabream S100 gene

In order to validate *in silico* predictions and characterize the regulatory effect of Ets1 on S100 gene transcription, pCMX-PL2 expression vector carrying the coding sequence of zebrafish Ets1a (pCMX-Ets1a) was co-transfected with the different reporter constructs (**Figure 2.7**). While C1 remained inactive in the presence of Ets1a, C2, which lacks the inhibitory sequence that inactivates C1, showed an induction of approximately 3.7 fold over control conditions (cells co-transfected with the empty pCMX-PL2 vector). Luciferase activity of C3 was increased 6.8 fold, suggesting the presence of functional Ets1 binding elements downstream of position -636, but also the occurrence of silencing elements (microsatellite?) in the region -768/-639 to explain the lower activity of C2. Moreover, binding elements predicted at positions -772/-757 and -754/-739 (scored 0.25) are most likely inactive, at least in the cell system used. Increase in luciferase activity upon expression of Ets1a was only 2.6 (C4), 2.8 (C5), 2.5 (C6) and 2.3 (C7) folds, suggesting the presence of active binding element(s) for Ets1 within the region -636/-513 but not within the region -513/-82. Similarly, the 2.3 fold increase in luciferase activity in C7 upon expression of Ets1a also suggests the presence of active binding element(s) for Ets1 within the region -82/+62. Three putative Ets1 binding sites were found within the region -636/-513, at positions -640/-625 (scored 0.25), -552/-539 (scored 1) and -517/-501 (scored 1). Because they were more likely to be active, sites with a score equal to 1 were mutated in C3, individually and in combination. Mutations of putative Ets1 binding sites did not affect basal activity of C3 construction (results not shown). Stimulation of luciferase activity by Ets1 in C3 construct mutated for binding sites at position -552/-539 (Mut 1) or -517/-501 (Mut 2) was reduced by approximately 2 and 2.5 fold, respectively, indicating that both elements are indeed binding sites for Ets1 in native promoter (**Figure 2.8**).

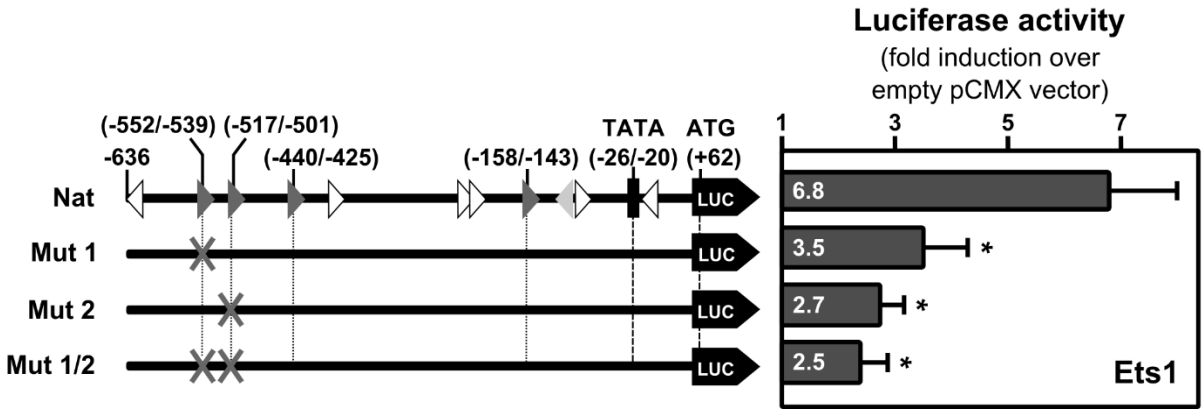


Figure 2.8. Transcriptional activity of native (Nat) and mutated (Mut) gilthead seabream S100 gene promoter (construct C3) in HEK-293 cells. Luciferase activity was measured upon overexpression of zebrafish Ets1a transcription factor and is presented as fold changes over the empty pCMX expression vector. Promoter constructs are shown on the left. Ets1 binding sites were mutated individually (Mut 1 and Mut 2) or in combination (Mut 1/2) and X indicates specific Ets1 binding core mutation. See Figure 2.7 legend for information on numbering and colour code. Luciferase activity (FLuc/RLuc; $n \geq 5$) is presented as fold induction over empty pCMX-PL2. Asterisks indicate values statistically different from wild type C3 constructs (one-way ANOVA with Tukey's post-test; $P<0.05$).

Mutation of the two binding sites (Mut 1/2) resulted in a slightly higher (2.7 folds) reduction of luciferase activity than for individual mutations, although differences were not statistically significant. In fact, it seems that mutating one site or the other, or both sites simultaneously, resulted in a similar reduction of luciferase activity, indicating that both elements need to be functional for an effective regulation of S100 transcription by Ets1. This could indicate that transcriptional regulation is achieved here through the formation of an Ets1 homodimer, as previously reported (Lamber et al., 2008; Babayeva et al., 2012). Although this should be further confirmed through additional site mutation, we propose that the remaining stimulation of luciferase activity by Ets1 in double mutated C3 or in native C7 may be achieved through the binding element at position -6/+9, although the vicinity of the transcription start site and its location after the TATA box may prevent its functionality, or through a binding element not predicted *in silico* and therefore still to be discovered. Altogether these results provided novel evidence for the transcriptional regulation of fish cartilage-specific S100 protein by Ets1. The role of Ets family of transcription factors in bone and cartilage development has been evidence through *in situ* hybridization (Kola et al., 1993; Maroulakou et al., 1994), where expression of Ets1 and Erg (ETS-related gene) was localized at sites of bone and cartilage formation during murine and avian development (Trojanowska, 2000). Ets1 has also been recently involved in the blocking of chondrocyte differentiation of cardiac neural

crest, being required to direct the proper migration and differentiation of cardiac neural crest in the formation of the interventricular septum: *Ets1*$^{-/-}$ mouse embryos develop an abnormal nodule of cartilage within the heart, suggesting that Ets1 may be a regulator of chondrogenesis (Gao et al., 2010; Sugiura and Ito, 2010). Various genes involved in chondrocyte differentiation have also been shown to be transcriptionally regulated by Ets1: *osteopontin*, expressed by hypertrophic chondrocytes (Sato et al., 1998), *collagen type II α-1*, expressed by differentiating chondrocytes (Peng et al., 2008), *integrin alpha-10*, expressed by articular chondrocytes (Wenke et al., 2006); *alkaline phosphatase*, expressed in maturing and hypertrophic chondrocytes (Qi et al., 2003) and *parathyroid hormone receptor 1*, expressed in pre-hypertrophic chondrocytes (Qi et al., 2003). Collectively, the available data and the new results presented here point toward a possible role of Ets1 as a specific regulator of chondrocyte function and/or cartilage formation.

2.2.5 Conclusions

This works provides evidence toward the transcriptional up-regulation of gilthead seabream S100 gene by Ets1, a transcription factor involved in the regulation of several other cartilage-related genes but never implicated in S100 gene regulation before. We were able to identify several DNA elements within S100 gene promoter capable of binding zebrafish Ets1a but also a region capable of suppressing entirely the transcription of S100 gene. Evidence was also collected towards a possible homodimerization of Ets1 in order to activate S100 transcription. The heterodimerization of Ets1 with other transcription factors such as the core-binding factor subunit α1 (Cbfa1), which was shown to enhance promoter activity of *osteopontin* in synergy with Ets1 (Sato et al., 1998), will need to be addressed in the future, in particular the role of the corresponding consensus binding site identified throughout our *in silico* analysis 10 pb downstream of Ets1 active binding site -552/-539**.** Elucidating S100/Ets1 complexes will certainly require further studies but the presence of both proteins in cartilaginous tissues, and the regulation of several cartilage-related proteins by Ets1, together with our novel data, provide some evidence towards a possible role of Ets1 as regulator of chondrocyte function and/or cartilage formation.

2.2.6 Acknowledgements

This work was co-funded by the European Regional Development Fund (ERDF) through COMPETE Program and by National Fund through the Portuguese Science and Technology Foundation (FCT) under PEst-C/MAR/LA0015/2011 project. Joana T. Rosa was supported by doctoral grant SFRH/BD/47433/2008 from the FCT.

CHAPTER 3

GILTHEAD SEABREAM SDR GENE

CHAPTER 3. GILTHEAD SEABREAM SDR-LIKE GENE

3.1 Identification of a short-chain dehydrogenase/ reductase in fish associated with tissue calcification and regulated by bone-related transcription factors

Joana Rosa
Cymon J. Cox
M. Leonor Cancela
Vincent Laizé

3.1.1 Abstract

Increase our knowledge on the biology of tissue calcification to better understand bone disorders such as osteoporosis is a major challenge for the next decades. Cell systems capable of *in vitro* calcification have been developed to study the mechanisms underlying bone formation. A gene coding for an unknown protein and strongly up-regulated during extracellular matrix (ECM) mineralization of a fish pre-osteoblast cell line was recently identified as a possible marker gene. Full-length cDNA was determined by PCR and *in silico* analysis of the deduced protein revealed the presence of domains typical of short-chain dehydrogenase/reductases (SDR). Closely related to carbonyl reductase 1, gilthead seabream Sdr-like protein belongs to a novel subfamily of SDR proteins with no orthologs in mammals. Analysis of gene expression by qPCR confirmed the strongly up-regulation of *sdr-like* expression during *in vitro* mineralization but also revealed high expression levels in calcified tissues. A possible role of Sdr-like in tissue mineralization and/or osteoblast differentiation was further evidenced through *in situ* hybridization and the localization of *sdr-like* transcript in bone opercula (pre-osteoblasts) and through the functional analysis of *sdr-like* promoter and the regulation of gene transcription by Runx2 and retinoic acid receptor, two regulators of

osteoblast differentiation and mineralization. Expression data also indicated a role of Sdr-like in gastrointestinal tract homeostasis and during gilthead seabream development at gastrulation and metamorphosis. This study reports a new subfamily of short-chain dehydrogenases/reductases in vertebrate and evidences, for the first time, a role for SDRs in tissue mineralization and/or osteoblast differentiation.

Keywords: Short-chain dehydrogenase/reductase; bone cell differentiation; Extracellular matrix mineralization; Gilthead seabream *Sparus aurata* (Teleostei); Suppression subtractive library.

3.1.2 Introduction

Bone diseases represent a major health problem, affecting millions of people worldwide. Still, no successful therapies have been developed and this situation is somehow related to the complexity of the mechanisms underlying bone formation and tissue mineralization in mammals. Non-mammalian models of tissue mineralization have been sought to collect new data toward the better understanding of these mechanisms and teleost fishes, which share with mammals a large number of important characteristics (e.g. organ systems, developmental mechanisms and physiological processes; (Fisher et al., 2003; Belloni et al., 1996; Nissen et al., 2006; Laizé et al., 2015), in particular regarding skeleton development and tissue mineralization, were found to be a suitable model to study mechanisms of vertebrate bone formation and homeostasis. In addition, the availability of various mutants with a bone phenotype or transgenic lines marking bone-related cells but also the availability of various *in vitro* and *in vivo* fish systems to study mechanisms of bone formation, have further validated teleost fish as a good model (Laizé et al., 2014 and references therein).

In an effort to develop resources to study tissue mineralization in fish, various *in vitro* cell systems representing different bone-related cell types have been developed from the vertebra of the gilthead seabream (Pombinho et al., 2004) and used to identified signaling pathways and genes controlling *in vitro* mineralization in fish (Pombinho et al., 2004; Rafael et al., 2006; Tiago et al., 2008; Conceição et al., 2008). A cDNA subtractive library prepared from osteoblast-like cell line VSa16 was used to identify genes differential expressed during extracellular matrix mineralization and likely to play a critical role in this process (Fonseca et al., 2007). Among those genes, *osteopontin/spp1* and *S100-like* transcripts were further studied

and shown to be involved in endo/perichondral ossification in gilthead seabream *Sparus aurata* (Fonseca et al., 2007; Fonseca et al., 2011). The second most up-regulated gene coded for an enzyme of the short-chain dehydrogenase reductase (SDR) superfamily. whose members participate in the metabolism of steroids, prostaglandins, glucocorticoids, lipids, amino acids, carbohydrates, retinoids, aliphatic alcohols, xenobiotics and as well in redox sensor mechanisms (reviewed in Oppermann et al., 2003). More than 47000 SDR genes have been identified throughout the tree of life (including about 82 genes in human) and over 300 crystal structures have been determined (Kallberg et al., 2010). While several members of the SDR superfamily have been associated to pathogenic processes (e.g. breast and prostate cancers), inflammation, and degenerative defects (e.g. the Alzheimer's disease, osteoporosis and diabetes) (Tomlinson et al., 2004; Vihko et al., 2005; Mindnich et al., 2004), many remain uncharacterized and could represent potential candidates for monogenic and multifactorial human diseases (Keller et al., 2006) and novel targets for drug development. The present study aims at characterizing the role of the gilthead seabream *sdr-like* gene during tissue mineralization by collecting basic data on levels and sites of gene expression and by investigating transcriptional regulation by mineralization-related transcription factors.

3.1.3 Materials and methods

3.1.3.1 cDNA and gene cloning

Full-length transcript and cDNA fragments were amplified by PCR using a 1:50 dilution of VSa16 Marathon cDNA library (Fonseca et al., 2007), 0.4 μM of gene-specific primers (**Supplementary Table 3.1** and **Supplementary Figure 3.1**) and Advantage cDNA polymerase mix (Clontech), according to manufacturer's instructions. Gene and 5'-flanking region were amplified by PCR using genomic DNA libraries (*Eco*RV, *Stu*I and *Sca*I, constructed using genomic DNA of a single individual and Clontech Universal GenomeWalker kit), 0.2 mM of adaptor primer 1 or 2 (AP1 or AP2, Clontech) and gene-specific primers (see **Supplementary Table 3.1**) and Advantage *Tth* Polymerase mix (Clontech), according to manufacturer's instructions. PCR products were size-separated by agarose-gel electrophoresis, purified using GFX Gel Band Purification kit (Amersham Biosciences) then cloned into pCRII-TOPO (Invitrogen) and sequenced.

3.1.3.2 RNA preparation

Total RNA was extracted from cell cultures and adult tissues using the method described by Chomczynski and Sacchi (1987) and from embryo, larvae and juvenile fish using TRIzol (Sigma-Aldrich) following manufacturer's instructions. Fish and cell lines were cultured and maintained as previously reported in Rafael et al. (2006). RNA quantity was determined using a NanoDrop 1000 spectrophotometer (Thermo Scientific) and five micrograms of each RNA sample were fractionated on 1% (w/v) agarose-formaldehyde gels to evaluate their integrity.

3.1.3.3 Analysis of gene expression by qPCR

Quantitative real-time PCR (qPCR) was performed using iCycler iQ system (Bio-Rad). Total RNA (1 μg) was treated with RQ1 RNase-free DNase (Promega) then reverse-transcribed at 37ºC for 1 h using M-MLV reverse transcriptase, RNase-out (Invitrogen) and reverse gene-specific primers SDRreal-RV and RPLreal-RV (qPCR primers are listed in **Supplementary Table 3.1**) specific for gilthead seabream *short-chain dehydrogenase reductase-like* (*sdr-like*) and *ribosomal protein L27a* (*rpl27a*) cDNAs, respectively. The reaction mixture containing 10 ng of reverse-transcribed RNA, 0.4 μM of SDRreal-FW/SDRreal-RV or RPLreal-FW/RPLreal-RV primers, and 1× iQ SYBR Green I mix (ABgene) was submitted to the following PCR conditions: an initial denaturation step at 95°C for 15 min then 50 cycles of amplification (each cycle is 30 s at 95°C, 45 s at 68°C). Levels of SDR-like gene expression were calculated using the $\Delta\Delta C_t$ method (Pfaffl, 2001) and normalized using expression level of *rpl27a* gene.

3.1.3.4 Promoter constructs

Promoter fragments were amplified by PCR from pCRII-TOPO vector carrying the full promoter sequence using reverse primer (SDR-RV-P2) in combination with forward primer SDR-FW-P1 (construct -386/+7LUC), SDR-FW-P2 (construct -253/+7LUC) or SDR-FW-P3 (construct -176/+7LUC), then digested with appropriate endonucleases (*Xho*I or *Hin*dIII) and cloned into pGL2-basic (Promega) upstream of luciferase reporter gene.

3.1.3.5 Transfection and luciferase assays

Xenopus laevis A6 cells (ATCC#CCL102) were cultured at 22ºC in 0.6× L15 medium supplemented with 5% fetal bovine serum and 1% antibiotics (all from Invitrogen) then transiently transfected using FuGene6 (Roche). Luciferase activity was assayed using Luciferase Assay System kit (Promega) in a TD-20/20 luminometer (Turner Designs). Relative light units were normalized to protein content using Coomassie Plus assay (Pierce). Expression vectors containing the cDNA of human retinoic acid receptor a (pCMX-hRARα; Umesono et al., 1991) or mouse Osf2 (pCMV-Osf2; Ducy and Karsenty, 1995) were used in this study. Medium of cells transfected with pCMX-hRARα was supplemented with 1 µM of all-*trans* retinoic acid (RA, in ethanol) 6 h after transfection. As a positive control of transfection, the vector containing mouse osteocalcin gene promoter (pII1.3luc; Ducy and Karsenty, 1995) was co-transfected with pCMV-Osf2 in A6 cells.

3.1.3.6 *In situ* hybridization

Gilthead seabream juveniles at 38 dph and individualized adult tissues were fixed for 24 h at 4ºC in 4% paraformaldehyde (PFA buffered at pH 7.4 with PBS), washed 3 x 10 min with PBS, then dehydrated through a PBS/methanol gradient and stored in 100% methanol at 4ºC. When appropriate, fish were passed through an increasing methanol/xylol series, embedded in paraffin then cross-sectioned. Sections (7-µm thick) were collected on TESPA (3-aminopropyltriethoxysilane, Sigma–Aldrich) coated slides, dried for 4 h at 37ºC and kept at 4ºC until used. Fragment of *sdr-like* cDNA (446 bp; complete coding sequence) cloned into pCRII-TOPO was linearized with appropriate endonucleases and 1 µg was used to synthesized dioxigenin (DIG)-labeled sense and antisense RNA probes using T7 or SP6 RNA polymerases (DIG RNA labeling kit, Roche Diagnostics). Size and integrity of the riboprobes were analyzed by agarose (1.2%) gel electrophoresis and *in situ* hybridization assay were performed as previously described (Fonseca et al., 2011).

3.1.3.7 Sequence collection, alignment and phylogenetic reconstruction

Sequences similar to gilthead seabream sdr-like were searched in GenBank (ncbi.nlm.nih.gov) and Ensembl (ensembl.org) public databases using on-site BLAST facilities. Sequences were clustered by species and aligned using the ContigExpress module of

Vector NTI (Invitrogen) to construct consensus sequences of sdr transcripts. Alignment of SDR protein sequences was created using TranslatorX V1.1 (Abascal et al., 2010) with parameters set to default. Phylogenetic analyses, using marginal likelihood (Lh = 16669.2659. 1 chains, 234,335 generations/samples, 34,335 burnin; Mean tree length = 11.0123), were conducted using Phylobayes Bayesian Markov chain Monte Carlo (MCMC) inferred from the data (CAT-GTR+G settings) (Lartillot et al., 2013).

3.1.4 Results

3.1.4.1 cDNA and gene cloning

Sequence of the full-length gilthead seabream *sdr-like* transcript was determined from the assembly of SSH fragments (Fonseca et al., 2007) and PCR fragments amplified from cDNA libraries using gene-specific primers (this study) using stringent overlapping criteria (sequence is displayed in **Supplementary Figure 3.1** and is available from GenBank sequence database using accession No. EU557022). Longest transcript is 854-bp long and codes for a protein of 256 amino acids (**Figure 3.1**).

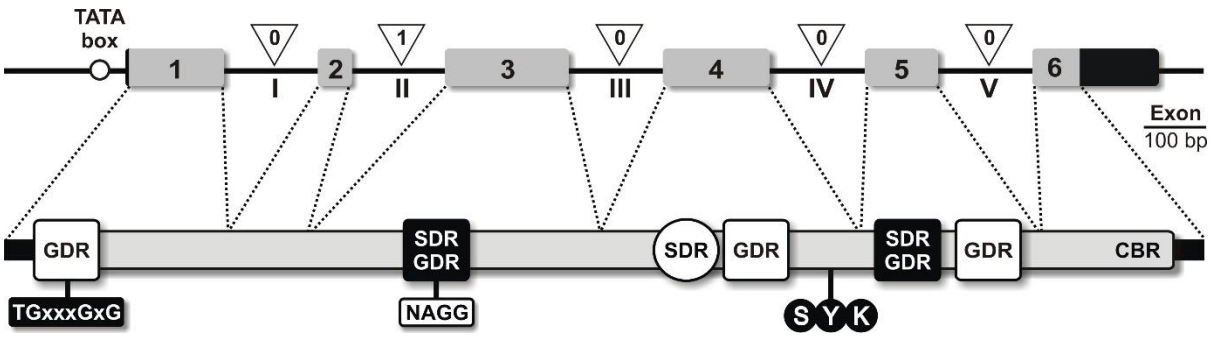


Figure 3.1. Schematic representation of gilthead seabream SDR-like gene and deduced protein structures. In gene structure, introns and exons are displayed as solid black lines and grey boxes, respectively. Phase of intron insertion is indicated in white triangles. In protein structure, domains and motifs identified using InterProScan are indicated as followed: $TGxxxGxG_{11-18}$, cofactor-binding site; $NNAG_{94-97}$; $N\text{-}S\text{-}Y\text{-}K_{119\text{-}158\text{-}177\text{-}181}$; SDR, short-chain dehydrogenase reductase; GDR, glucose/ribitol dehydrogenase; CBR, carbonyl reductase.

Corresponding gene and promoter region were cloned by PCR from genomic DNA libraries using gene-specific primers (designed in cDNA untranslated regions) and adapter primers 1 or 2. A 3386-bp gene sequence was reconstructed from the assembly of PCR fragments (sequence is displayed in **Supplementary Figure 3.2** and is available from

GenBank sequence database using accession No. EU557023) and gene structure (6 exons, 5 introns and a 386-bp promoter region) was determined using Spidey mRNA-to-genomic alignment tool available at NCBI (**Figure 3.1**).

3.1.4.2 Gilthead seabream sdr-like is a classical short-chain dehydrogenase/reductase

Protein sequence deduced from *sdr-like* cDNA/gene was submitted to (i) Blast at NCBI and HMMER at hmmer.janelia.org to identify similar sequences and get insights into protein function, conservation and evolution, and (ii) InterProScan at www.ebi.ac.uk to identify signatures for protein domains and collect evidences for protein function(s) and structure. Sequences exhibiting the highest similarity to gilthead seabream protein belong to the short-chain dehydrogenases/reductase family and the presence of domains and motifs typical for SDRs – (i) $TGxxxGxG_{11\text{-}18}$ and D_{66}; (ii) $NNAG_{94\text{-}97}$; (iii) $N\text{-}S\text{-}Y\text{-}K_{119\text{-}158\text{-}177\text{-}181}$ and (iv) $PGxxxT_{208\text{-}209\text{-}213}$ – were also identified (**Supplementary Figure 3.1** and **Figure 3.1**). Using facilities at the Short chain Dehydrogenases/Reductases database (www.sdr-enzymes.org) it was possible to classify gilthead seabream Sdr-like into the group of classical SDRs (C type), which comprises proteins consisting of approximately 250 amino acid residues. Signatures for domains specific of the carbonyl reductase (CBR), retinol dehydrogenase (RDH) and glucose/ribitol dehydrogenase (GRD) were identified, narrowing the affiliation of Sdr-like to 3 enzymatic groups (**Figure 3.1**). Finally, the absence of a classical signal peptide was proposed by SignalP (www.cbs.dtu.dk), suggesting that gilthead seabream Sdr-like remains intracellular. Among SDR protein family, the carbonyl reductase 1 (CBR1) was the member exhibiting the highest percentage of identity with gilthead seabream protein (around 30% for fish species and 23% for mammals). Due to the limited homology observed between SDR proteins – typically 15-30% identity in pairwise comparisons for members of the same family – the identification of orthologous sequences may not be trivial and a phylogenetic analysis of the most closely related sequences may be needed to establish the relationship of novel sequences with known members of this family. Gilthead seabream Sdr-like, sequences representative of major SDR groups and unannotated sequences showing high similarity with seabream sequence were aligned and their relationship represented in a phylogenetic tree (**Figure 3.2; Supplementary Sequences 3.1** and **Supplementary Figure 3.3**).

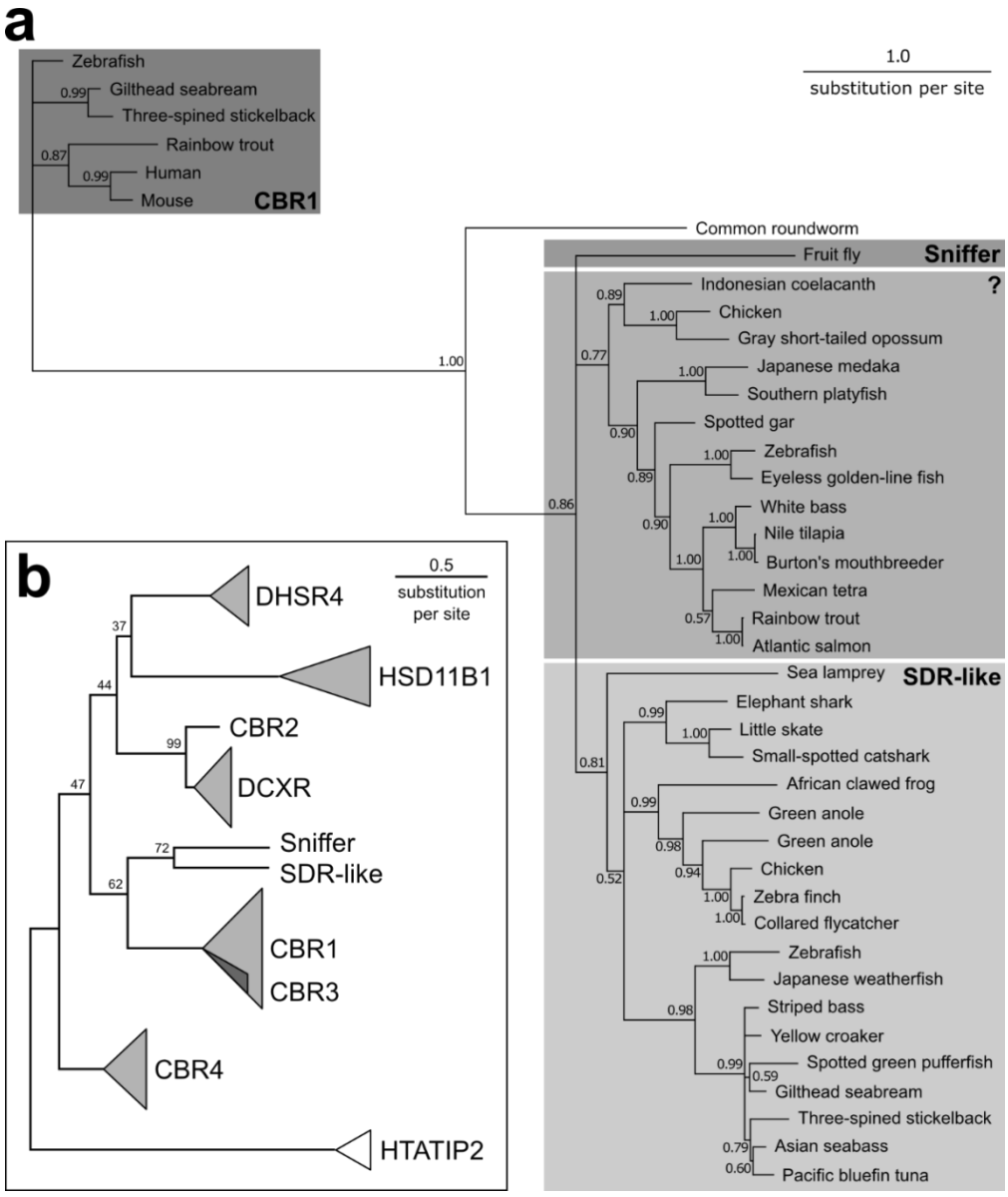


Figure 3.2. Molecular phylogenetic analysis of vertebrate *SDR* genes. Tree (marginal likelihood using Phylobayes Bayesian Markov chain Monte Carlo (MCMC) inferred from the data (CAT-GTR+G settings)) in panel **a** represents the phylogeny of unannotated sequences closely related to gilthead seabream *sdr-like* and their relationship with the fruit fly sniffer gene. Sequences of the carbonyl reductase 1 (CBR1) subfamily were used as an outgroup and a root for the tree. Tree (maximum likehood; 1000 bootstrap constructed using MEGA 6) in panel **b** represents a simplified phylogeny of the SDR subfamilies closely related to gilthead seabream *sdr-like* and fruit fly *sniffer*; Sequences of the HIV-1 Tat Interactive Protein 2 (HTATIP2) were used as an outgroup and a root for the tree (original tree in Supplementary.Figure 3.3). Numbers on panel **a** branches represent bootstrap values and posterior probabilities of the Bayesian analysis, respectively, and on panel **b** branches represent the percentage of trees in which the associated taxa clustered together.

While its close relationship with CBR1 was confirmed (**Figure 3.2b**), gilthead seabream protein still better clustered with unannotated SDR members apart from all known subfamilies (e.g. CBR1, CBR2, CBR3, Sniffer, DCXR), suggesting that these proteins may form a novel

group of SDRs that remain to be characterized (**Figure 3.2a**). Strikingly, unannotated sequences clustering with gilthead seabream Sdr-like were found in most vertebrate taxa (e.g. jawless fish, cartilaginous fish, ray-finned fish, amphibians and sauropsids) but none were from mammals, suggesting that *sdr-like* gene may have disappeared in the genome of the mammalian ancestor throughout evolution.

3.1.4.3 Expression of *sdr-like* gene is up-regulated during *in vitro* mineralization

Levels of *sdr-like* transcript were determined by qPCR in gilthead seabream cell lines derived from vertebra (VSa13 and VSa16), branchial arches (ABSa15) and caudal fin (CFSa1). While *sdr-like* expression was basal in all cells lines cultured under non-mineralizing conditions, it was up-regulated in cells capable of *in vitro* mineralization (i.e. VSa13, VSa16 and ABSa15; **Figure 3.3**), although following different patterns: a strong initial up-regulation (approximately 12 folds) followed by a progressive decrease in cells from the chondrocyte lineage (i.e. VSa13 and ABSa15) and a milder but sustained up-regulation (approximately 6 folds) in cells from the osteoblast lineage (i.e. VSa16).

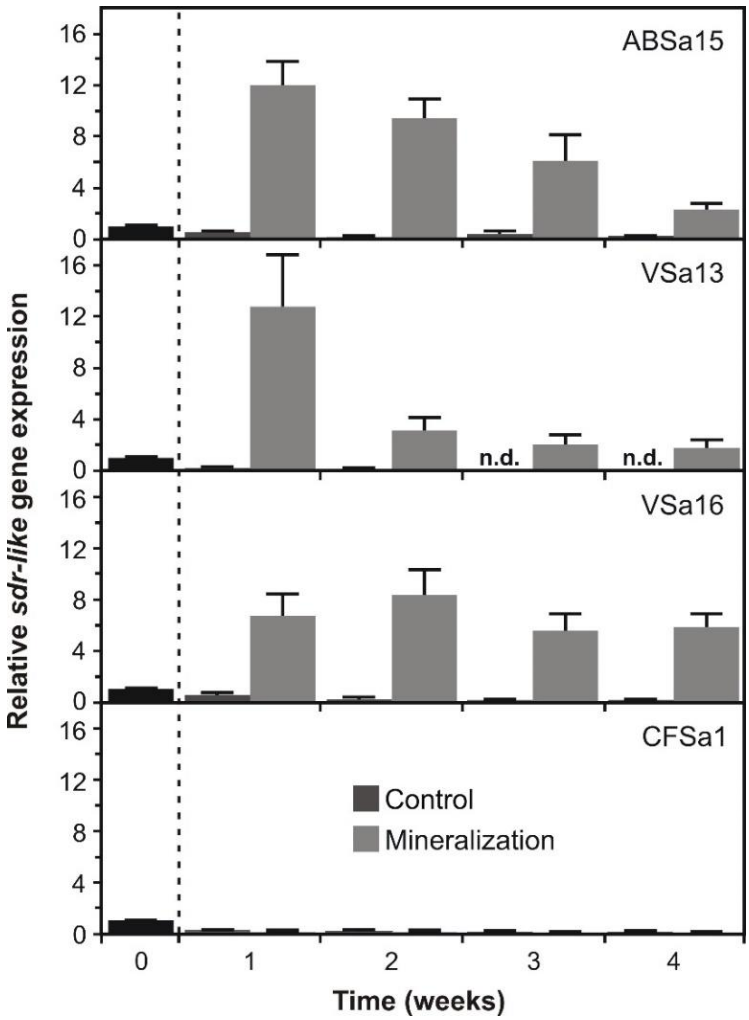


Figure 3.3. Relative expression of gilthead seabream *sdr-like* gene in ABSa15, VSa13, VSa16 and CFSa1 cell lines cultured for 4 weeks under control (dark grey bars) and mineralizing (light grey bars) conditions. Expression of *rpl27a* housekeeping gene was used to normalize levels of *sdr-like* gene expression. Time zero (black bar) was selected as the calibrator sample and its expression was set to 1 for all cell lines. Values are the mean of 4 independent qPCR experiments and are presented with their standard deviation. All values were statistically different from corresponding controls ($P < 0.05$; Student's *t*-test), except for CFSa1 cells.

No significant changes in *sdr-like* expression were observed in CFSa1 cells (not capable of mineralizing their extracellular matrix) cultured under control or mineralizing conditions. Altogether, these results suggest a role of Sdr-like in mechanisms of *in vitro* mineralization.

3.1.4.4 Spatiotemporal expression of *sdr-like* gene during development and in adult tissues

Levels of *sdr-like* transcript were also determined by qPCR throughout fish development and in adult fish tissues. Gene expression was basal during embryonic development until gastrulation then strongly up-regulated (*ca.* 110-260 folds over reference sample, i.e. unfertilized eggs) until hatching. Gene expression was again basal during larval development until the end of metamorphosis then again up-regulated throughout juvenile development, although to a lesser extent (*ca.* 50-120 folds; **Figure 3.4**).

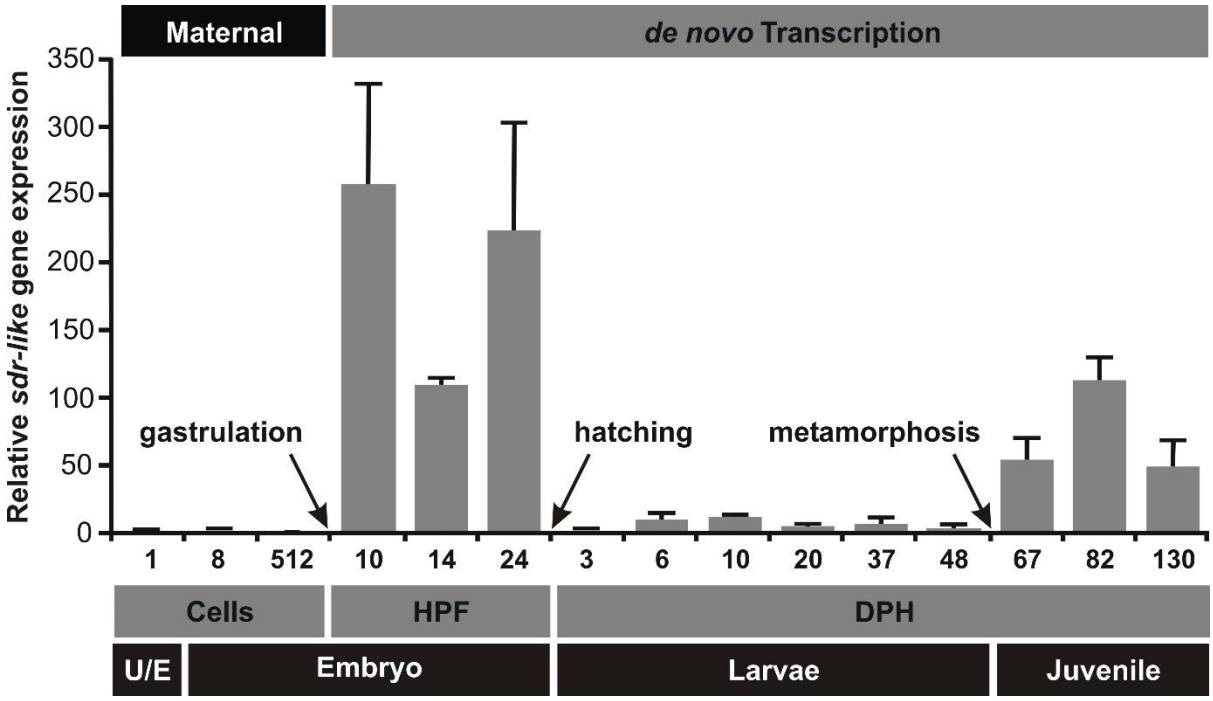


Figure 3.4. Relative expression of gilthead seabream *sdr-like* gene throughout development. Expression of RPL27a housekeeping gene was used to normalize levels of SDR-like gene expression. Unfertilized eggs (U/E) was selected as the calibrator sample and its expression was set to 1. Values are the mean of 4 independent qPCR experiments and are presented with their standard deviation. HPF, hours post-fertilization; DPH, days post-hatching.

In adult fish, *sdr-like* transcript was detected in all tissues tested. Highest levels of expression were found in soft tissues (organs and accessory organs of gastrointestinal tract, namely stomach, gall bladder, tongue and pancreas, *ca.* 180, 40, 10 and 8 folds over reference sample, i.e. liver, respectively), while intermediate levels were found in calcified tissues, e.g. branchial arches, teeth and operculum (*ca.* 16, 12 and 4 folds, respectively; **Figure 3.5**).

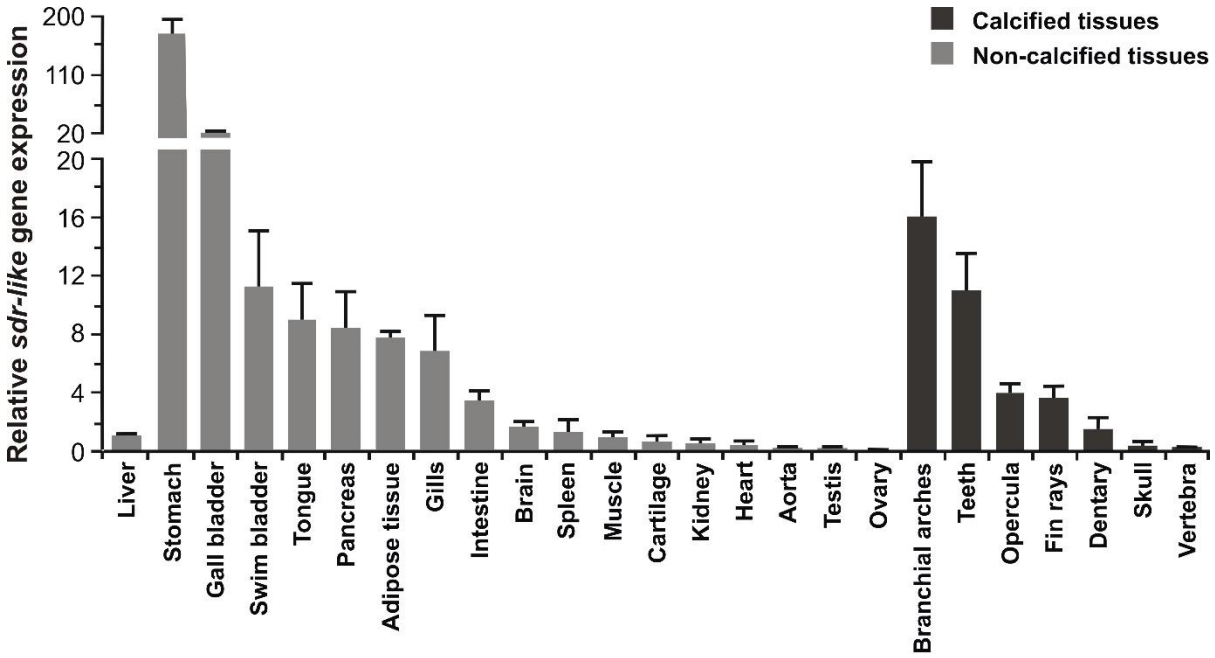


Figure 3.5. Relative expression of gilthead seabream *sdr-like* gene in calcified and non-calcified adult tissues. Expression of RPL27a housekeeping gene was used to normalize levels of SDR-like gene expression. Liver was selected as the calibrator sample and its expression was set to 1. Values are the mean of 5 independent qPCR experiments and are presented with their standard deviation. L, liver; St, stomach; GB, gall bladder; SB, swim bladder; To, tongue; P, pancreas; AT; adipose tissue; Gi, gills; I, intestine; Br, brain; Sp, spleen; M, muscle; Ca, cartilage; K, kidney; H, heart; Ao, aorta; Ts, testis; O, ovary; BA, branchial arches; Te, teeth; BO, Bone Operculum FR, fin rays; BD, Bone Dentary; BS, Bone Skull; BV, Bone Vertebra.

Sites of gene expression were determined by *in situ* hybridization (ISH) in the whole body of a 38-dph juvenile and in the stomach of and adult gilthead seabream (**Figure 3.6**). A specific and strong signal was observed in structures of the gastrointestinal tract for both 38-dph and adult fish. 38-dph fish exhibited strongly marked cells at the simple columnar epithelium (sce) of the oesophagus (**Figure 3.6a**), with apparently no expression in goblet cells (gc). In adult stomach (**Figure 3.6b**), cells from the submucosa (sm) were also positive, while no expression was observed in gastric glands (gg) or the mucosa (m). In bony structures, pre-osteoblastic cells lining the trabecular bone (**Figure 3.6c**) and the operculum (**Figure 3.6d**) also shown a positive signal, which once again is consistent with the gene expression detected in calcified tissues. No expression was detected in branchial arches or teeth as expected from qPCR results.

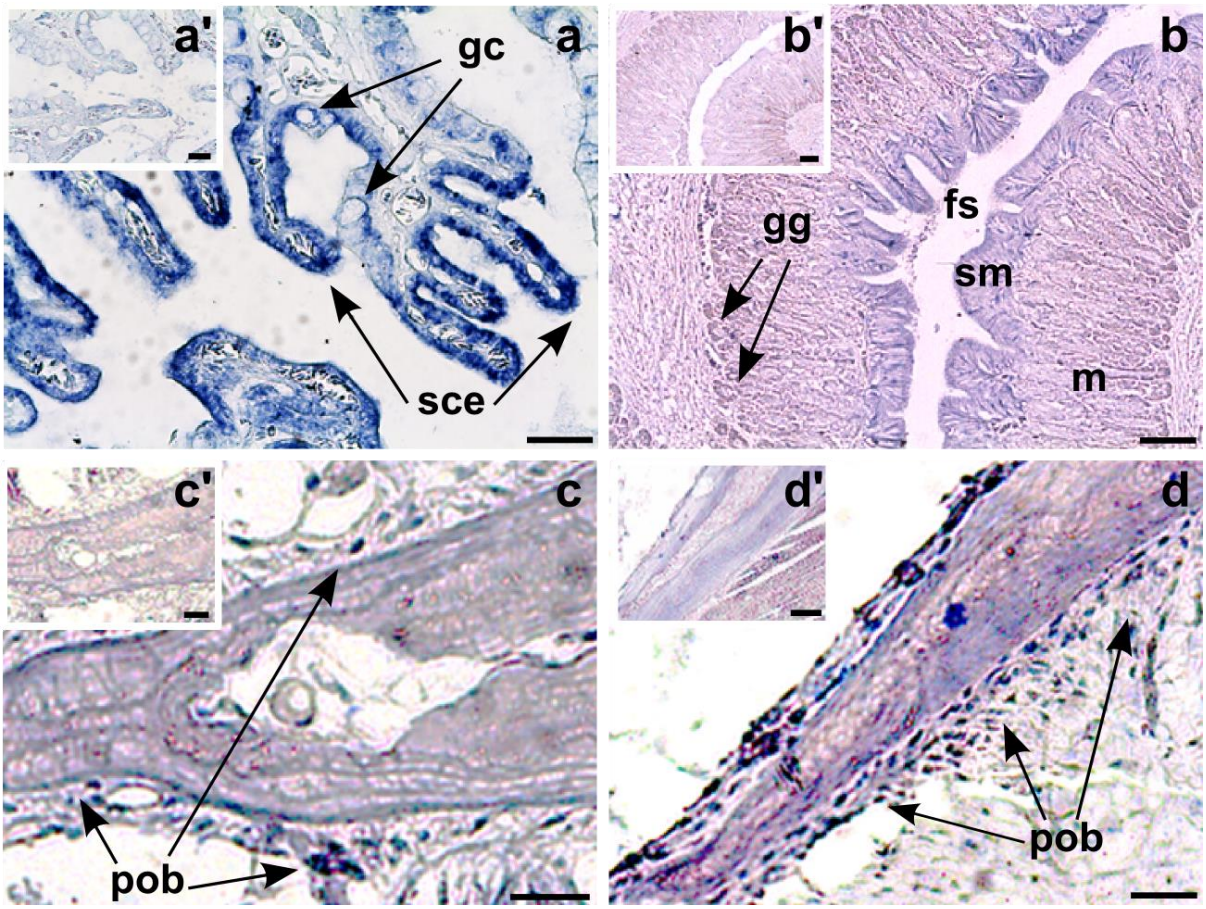


Figure 3.6. Sites of gilthead seabream *sdr-like* gene expression determined by *in situ* hybridization in 38-dph (**a**, **a'**, **c**, **c'**, **d** and **d'**) and adult (**b** and **b'**) gilthead seabream specimens. **a-d**, antisense riboprobe. **a'-d'**, sense riboprobe.; OP, operculum;; TC, trabecula cranii; fs, fundic stomach; gc, goblet cells; gg, gastric glands; m, mucosa; po, pre-osteoblasts; sce, simple columnar epithelium; sm, submucosa. Bar is 70 µm in **b** and **b'**, 40 µm in **a** and **a'**, and 20 µm in **c**, **c'**, **d** and **d'**.

However, in general, both qPCR and ISH data clearly confirmed that Sdr-like probably plays a central role in processes associated to tissue mineralization and bone formation but also suggested that it is equally important, if not more, in mechanisms not related to mineralization as shown by the high levels of expression in stomach.

3.1.4.5 Runx2 and retinoic acid regulates the transcription of *sdr-like* gene

In silico analysis of the 5'-flanking region of *sdr-like* gene using MatInspector identified, in addition to the canonical TATA box (TATAAAA$_{-26/-20}$), various consensus sequences of transcription factor binding sites previously shown to be associated with promoter activity in eukaryotic genes (**Supplementary Figure 3.3**). Two binding sites for the runt-related

transcription factor 2 (Runx2_1 at -276/-262 and Runx2_2 at -194/-180) and one binding site for retinoic acid receptor (RAR; -56/-32), two well-known regulators of osteoblast differentiation and mineralization and of various mineralization-related gene expression (Cancela and Price, 1992; Drissi et al., 2000; Lind et al., 2013), were predicted (matrix similarity > 0,85) in the 5'flanking region of gilthead seabream *sdr-like* gene. The functionality of these sites was investigated in A6 cells transiently co-transfected with constructs containing fragments of *sdr-like* promoter upstream the luciferase reporter gene and expression vectors carrying either mouse Runx2 (Osf2) or human RARα cDNA under the control of CMV promoter. Activity of *sdr-like* promoter was assessed as a function of luciferase activity (**Figure 3.7**) and shown to be stimulated approximately 5 times by Runx2 while inhibited approximately 12 times by RAR in the construct C1 containing the region -386/+1. Runx2 stimulation of promoter activity in constructs C2 (region -253/+1) and C3 (region -176/+1), containing either a single or no putative binding site for Runx2, respectively, was reduced by half or totally abolished, suggesting the functionality of both predicted Runx2-binding sites and that to a similar extent. Using the same promoter constructs, RARα inhibition of *sdr-like* promoter activity was apparently not affected in C2 and C3 (**Figure 3.7**), suggesting that RAR binding sites are probably all located within the promoter region -176/+1.

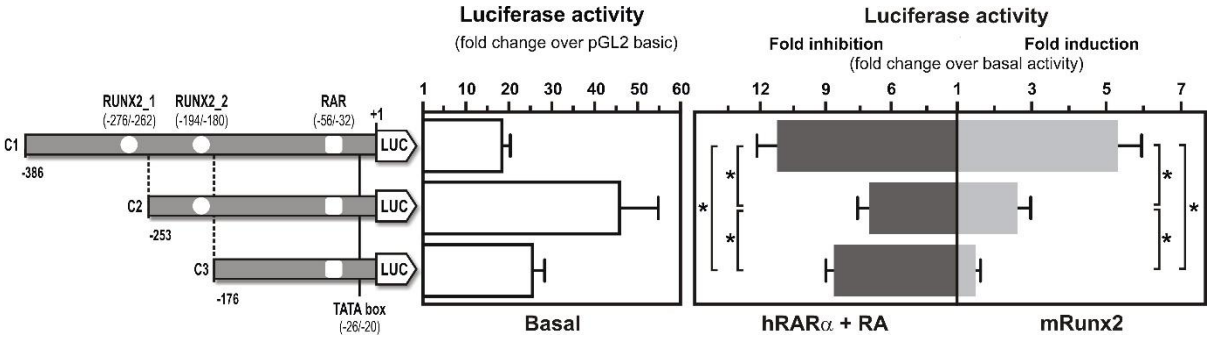


Figure 3.7. Transcriptional activity of gilthead seabream *sdr-like* gene promoter and its regulation by runt-related transcription factor 2 (Runx2) and retinoic acid receptor (RAR) in A6 cells. Promoter constructs (-386/+1Luc, -253/+1Luc and -176/+1Luc) used in transient transfections are presented in the left panel and putative binding sites for Runx2 (white circles) and RAR (white squares) are indicated. Basal activity (middle panel) and regulation by Runx2 and RAR (right panel) is inferred from luciferase (LUC) activity. Values are the mean of 4 independent transfections and are presented with their standard deviation. All values were statistically different from corresponding controls ($P < 0.05$; Student's *t*-test).

Although this should be demonstrated, e.g. through site-directed mutagenesis, we propose that RAR-binding site predicted at position -56/-32 may be responsible for the

regulation of *sdr-like* promoter activity by retinoic acid. Regulation of *sdr-like* transcription by Runx2 and RAR further evidenced the possible role of this short chain dehydrogenase/reductase in mechanisms of bone formation probably osteoblast differentiation.

3.1.5 Discussion and Conclusions

A powerful approach for studying the genetic nature of biological processes is to characterize genes that vary in expression during this process. Cell lines recently developed from vertebrae of the gilthead seabream and capable of *in vitro* mineralization were previously used to construct a cDNA subtractive library and identify genes up-regulated during extracellular matrix mineralization (Fonseca et al., 2007). Expanding from this initial work, we have already demonstrated the role of osteopontin and S100 calcium binding proteins in the mechanisms underlying endochondral ossification (Fonseca et al., 2007; Fonseca et al., 2011). Following the same approach, we collected here basic data on the expression and regulation of a gene coding for a short-chain dehydrogenase/reductase and evidenced its role in mechanisms of osteoblast differentiation and mineralization.

3.1.5.1 Gilthead seabream short-chain dehydrogenase/reductase is part of a novel SDR subfamily

SDRs form one of the largest and oldest protein superfamilies characterized so far, with thousands of members clustered in hundreds of subfamilies and present in all the kingdoms of life, from primitive bacteria to higher eukaryotes (Kallberg et al., 2010). Probably because they diverged long time ago, members of the different subfamilies identified so far only share a limited similarity (typically 15–30% identity; Joernvall et al., 1995), making difficult their identification. However, various canonical features have been maintained throughout evolution (Oppermann et al., 2003) and several domains and motifs typical for SDRs were identified in Sdr-like. The comparative analysis of protein sequence has revealed high similarity between Sdr-like and type C short-chain dehydrogenases/reductases, which are NAD(P)(H)-dependent oxidoreductase of 250 aa in length (Kallberg et al., 2010) with conserved regions containing several motifs: (i) $TGxxxGxG_{12\text{-}19}$ and D_{60} in N-terminal region for coenzyme binding and maintenance of central β-sheet, (ii) $NNAG_{86\text{-}89}$ important for the stabilization of the 3D structure; (iii) a catalytically active site composed by a tetrade $N\text{-}S\text{-}Y\text{-}K_{111\text{-}138\text{-}151\text{-}155}$ and N_{179} and (iv) in C-terminal the motif $PGxxxT_{183\text{-}184\text{-}188}$ crucial for the determination of reaction

direction (numbering refers to 3β/17β-HSD with Protein Data Bank accession number 1hxh; see **Supplementary Figure 3.1** and **Figure 3.1** for Sdr-like numbering). Although they shared some similarity, gilthead seabream Sdr-like and other closely related proteins identified among unannotated sequences clearly clustered apart from the SDRs of the CBR1 subfamily (and from any member of the other 313 SDR subfamilies) and probably represents a novel subfamily that remain to be characterized. A classification system was developed based upon hidden Markov models (HMMs) in order to facilitate SDR proteins identification in subfamilies, and until now 31900 of ~ 47000 retrieved SDRs (Persson et al., 2009; Kallberg et al., 2010) were already addressed into a specific subfamily. However approximately 9700 SDRs form clusters that are too small (fewer than 20 members with maximum 80% sequence identity) for them to be reliably identified with an HMM, which could be the case of the Sdr-like protein and orthologs. No ortholog of Sdr-like was identified in mammals and, although we cannot excluded that it still has to be discovered in mammalian genomes, the most probable hypothesis for its absence in sequence databases is that *sdr-like* gene was lost during the course of evolution in the genome of the mammalian ancestor after branching from other tetrapods. This hypothesis is in agreement with the existence of SDR subfamilies that are unique to one group of species, or more but not to all. To date HMM classification comprises: 178 families with only bacterial members; 41 with only eukaryotic members (1 exclusive for insect, 7 limited to plants, ~15 for fungi and ~16 with multiple groups including mammalians); 63 families with members among both bacteria and eukaryotes; 18 families with members among archaea and bacteria; 14 families with members from all domains (from these classification 37 families include the 82 human members) (Kallberg et al., 2010). *In silico* analyses presented here clearly indicated the SDR nature of gilthead seabream protein but failed to provide insights into its function – e.g. enzyme activity and substrate specificity – since none of the closely related proteins have been functionally characterized so far.

3.1.5.2 Role of Sdr-like in mineralization and fish development

Basic data on levels and sites of gene expression and on transcriptional regulation usually give precious insights on what could be the physiological role of a novel protein. Expression and regulatory data presented here have clearly evidenced a role of Sdr-like in mechanisms of tissue mineralization, possibly by regulating the differentiation of cells involved in this process, i.e. the osteoblasts. Both the up-regulation of *sdr-like* expression in mineralizing cell cultures and the high transcript levels in calcified tissues suggested a role in mechanisms of tissue

mineralization; the localization of *sdr-like* transcript in pre-osteoblasts lining bone structures, the early up-regulation of gene expression *in vitro* (cell differentiation occurs during the first week upon exposure to osteogenic cocktail; (Pombinho et al., 2004)) and the regulation of its transcription by factors controlling bone cell differentiation (Drissi et al., 2000; Conceição et al., 2008; Lind et al., 2013) further evidenced a possible role of Sdr-like in osteoblast differentiation. Despite the central role of SDR proteins in a variety of pathways and metabolisms related to development and homeostasis, only two other members of this large protein superfamily have been associated with bone function so far: (i) the hydroxysteroid (17β) dehydrogenase 2 (HSD17B2) catalyzes the conversion of 17β-hydroxy forms of sex steroids into 17β-keto forms, protecting tissues against excessive amounts of sex steroids (Shen et al., 2008); (2) the hydroxysteroid (11β) dehydrogenase (HSD11B) converts the cortisol (biologically active form) into the cortisone (biologically inactive form), preventing the over-stimulation of the mineralocorticoid receptor by glucocorticoids (reviewed in Gathercole et al., 2013). Excess of both sex steroids and glucocorticoids is detrimental to skeletal development and impact bone function (Shen et al., 2008; Rauch et al., 2010; Weinstein et al., 1998) and it is therefore expected that enzymes that regulate circulating levels of both molecules are central to bone metabolism. Immunohistochemistry and *in situ* hybridization also revealed that HSD11B expression was primarily localized in osteoblasts, and to a lesser extent also in osteoclasts of human adult bone (Cooper et al., 2000), and was dynamically regulated across osteoblast differentiation (Eijken et al., 2005). This kind of feedback regulatory mechanism, where the production of a protein is negatively regulated by the molecule that is metabolizes, has also been reported for SDR proteins related to retinoic acid metabolism. Human retSDR1, a regulator of vitamin A metabolism involved in the local storage of retinol in neuroblastoma cells is induced by RA in a wide array of cell lines derived from different human tissues (Cerignoli et al., 2002) and African clawed frog Dhrs3 was reported to function as a RA-induced feedback inhibitor of RA synthesis during embryonic development (Kam et al., 2010). Although none of the SDRs involved in RA metabolism have been directly associated to bone function, mouse retinol dehydrogenase 10 (RDH10) was shown to be essential for embryonic limb, craniofacial, and organ development on a RA-dependent manner (Sandell et al., 2007). RDH10 was also shown to regulate the expression of bone morphogenetic protein 2, a signalling protein central to mechanisms underlying chondrocyte proliferation and differentiation (Minina et al., 2001). Although this is speculative and would need to be validated through experimental and functional data, we propose, based on evidences associating retinoic acid to the regulation of *sdr-like* transcription and reporting both the

presence of *sdr-like* transcript and RA molecules at same stages of osteoblastogenesis, but also based on the occurrence in gilthead seabream Sdr-like of signatures for retinol binding domains (retinoic acid is produced from retinol by retinol dehydrogenases), that gilthead seabream Sdr-like could be a regulator of circulating RA levels to prevent the overstimulation of RAR and consequent effects on the differentiation of primary osteoblast and bone matrix synthesis in mature osteoblasts (reviewed in Imai et al., 2013).

3.1.5.3 Role of Sdr-like in processes not related to bone function

Because SDRs can metabolize steroids, prostaglandins, retinoids, glucocorticois, alcohols and xenobiotics in a large set of tissues through various pathways, one single enzyme could be expressed and implicated in more than one system/process/tissue (Oppermann et al., 2003). Our expression data clearly evidenced the role of Sdr-like in physiological processes not related to tissue mineralization. Expression of *sdr-like* gene at gastrulation, when mineralization process is not started yet, could be related to cell differentiation and proliferation, processes that are actively taking place and many signalling pathways are activated (Stern, 2004). Again, it will be useful to better understand tissue- or cell-type specific regulation of *sdr-like* by RA, since many of the known functions of RA, besides is role in bone metabolism, take place during development, at early embryogenesis, starting at the gastrula stage (Maden et al., 1998; Ribes et al., 2009), and it is feasible to hypothesize that higher levels of *sdr-like* expression at this particular stage could be mediated by RA. It is known, for example, that during embryonic development African clawed frog Dhrs3 is highly active at the onset of gastrulation (Kam et al., 2010) and dependent of RA regulation. *In silico* analysis of *sdr-like* promoter also revealed putative binding sites for transcription factors with key roles in differentiation, proliferation and cell growth (S**upplementary Figure 3.3**): 1) S8 is expressed at site of epithelial-mesenchymal interactions, including those in craniofacial tissues (Karg et al., 1997) and is activated by RA in embryonic stem cells, neural crest cell lines, and also during spontaneous differentiation in embryonic stem cells (Jong et al., 1993); 2) C/EBPs (CCAAT/enhacer binding proteins) are a family of factors that regulate cell growth and differentiation and function in a variety of tissues (Fan et al., 2011); 3) the LIM-homeodomain transcription factor (Lmx1b) is expressed in dorsal mesenchymal cells, where it is required autonomously in skeletal progenitors to direct specify dorsal pattern (Mcmahon et al., 2009); 4) Nkx2-5 one of the earliest markers of the cardiac lineage as it is abundantly expressed in the cardiac progenitors that form the cardiac crescent (Behrens et al., 2012); and 5) Barx2 that

differentially control the expression of target genes during embryonic development (Jones et al., 1997). Although it will require experimental validation, *in silico* data suggests that *sdr-like* expression could be regulated by transcription factors involved in embryonic development and therefore that Sdr-like may play a central role in the early development of gilthead seabream. Still, high expression of *sdr-like* at gastrulation stage could also be associated with the formation of the gastrointestinal tract, a process that initiates from the endoderm precisely at this stage in zebrafish (Kimmel et al., 1995). qPCR and *in situ* hybridization data revealed the presence of high levels of *sdr-like* transcripts in the stomach and gastrointestinal tract from juvenile throughout adult fish and we hypothesize a possible role of this novel SDR in mechanism associated with gastrointestinal tract development and metabolic functions. Interestingly, human HSD11B and mouse HSD17B2, previously associated with bone formation (see above), are also expressed in gastrointestinal tissues (Chapman et al., 2013; Mustonen et al., 1998), indicating a possible accumulation of functions, as it could be the case for gilthead seabream Sdr-like. The high levels of *sdr-like* expression at the end of metamorphosis and during juvenile growth may also be related with the gastrointestinal system, in addition to bone remodeling processes. At day 50, the fish diet was changed from freshly hatched *Artemia* nauplii to artificial food, requiring new adaptive metabolic processes and the differential expression of gastrointestinal system-associated genes in order to metabolize the new nutritional intake (Infante and Cahu, 2001; Micale et al., 2008). Interestingly α1-fetoprotein transcription factor (FTF), also identified as a putative transcription factor regulating Sdr-like expression, besides being crucial for embryogenesis is also a central regulator of lipid metabolism, having a central role in developmental, nutrition, and metabolic functions from early embryogenesis through (Paré et al., 2004; Xua et al., 2010), further strengthening our hypothesis.

In conclusion, we demonstrated the existence of a novel subfamily of short-chain dehydrogenase/reductase – SDR-like – with members distributed in all vertebrate taxa but absent in mammals. In gilthead seabream, Sdr-like has been associated to mechanisms of osteoblast differentiation and mineralization but also to gastrointestinal tract function and possibly to embryonic development. We hypothesized that Sdr-like function could be related to retinoic acid metabolism and that gene transcription could be regulated b RA in a feedback regulatory loop. Future studies should aim at collecting functional data toward validating these basic expression data and get insights into the osteogenic activity of Sdr-like proteins

3.1.6 Acknowledgments

Acknowledgment: JR was supported by a doctoral grant (SFRH/BD/47433/2008) from the Portuguese Foundation for Science and Technology (FCT). This work was partially funded by CCMAR and EU-FP6 MGE network of excellence (GOCE-CT-2004-505403). pCMV-Osf2 and pII1.3luc vectors were kindly provided by Dr. Gerard Karsenty, Baylor College of Medicine, Houston, TX. pCMX-hRARa vector was kindly provided by Dr. Roland Schüle, University of Freiburg, Germany. The help of Vera G. Fonseca (SSH fragment cloning) and Natércia Conceição (functional promoter analysis) is gratefully acknowledged.

CHAPTER 4

GILTHEAD SEABREAM BHMT3 GENE

CHAPTER 4. GILTHEAD SEABREAM BHMT3 GENE

4.1 Central role of betaine-homocysteine S-methyltransferase 3 in chondral ossification and evidence for sub-functionalization in fish

Joana Rosa
Daniel M. Tiago
Cátia L. Marques
Vijayakumar Parameswaran
Luís Fonseca
M. Leonor Cancela
Vincent Laizé

4.1.1 Abstract

To better understand the complex mechanisms of bone formation and mineralization, it is fundamental that genes central to signaling and regulatory pathways, and to matrix formation are identified. In this regard, cell systems capable of *in vitro* calcification have been used to analyze genes differentially expressed during extracellular matrix mineralization and the gene *bhmt3*, coding for a betaine-homocysteine S-methyltransferase, was shown to be strongly down-regulated in mineralizing gilthead seabream cells. Gene expression patterns determined by qPCR in adult tissues and throughout gilthead seabream development revealed the role of Bhmt3 in mechanisms underlying tissue mineralization and *in situ* hybridization indicated that gene expression was limited to cartilaginous structures undergoing endochondral ossification. Activity of *bhmt3* promoter was regulated by transcription factors involved in bone and cartilage development (ETS1, SP1 and MEF2c), further demonstrating the central role of Bhmt3 in chondrogenesis and/or osteogenesis. Additional *bhmt* isoforms identified in gilthead

seabream and in neoteleost fish species revealed the explosive diversity of *bhmt* genes in this lineage. Their appearance was associated with multiple gene duplication events that occurred throughout evolution and their expression split in several soft tissues suggested that neoteleostean Bhmt may have undergone several steps of subfunctionalization when compared to mammalian proteins. Data reported here provides new insights into the evolution of betaine-homocysteine methyltransferases in vertebrates and evidences a novel function in bone and cartilage development for Bhmt3 that could represent a suitable marker of chondral ossification

Keywords: Gilthead seabream *Sparus aurata*; betaine-homocysteine S-methyltransferase; bone formation; *in vitro* mineralization; molecular evolution; gene expression; transcriptional regulation; taxonomic distribution

4.1.2 Introduction

Vertebrate skeleton is a multifunctional organ that provides support and protection for internal organs, storage and balance for minerals (mainly calcium and phosphorus) and, in mammals, plays hematopoietic functions (Karsenty and Wagner, 2002; Pirraco et al., 2010). It is mainly composed of bone and cartilage and undergoes constant ossification and remodelling trough complex cellular and molecular mechanisms (Blair et al., 2002). Although human and mouse genetics have largely expanded our understanding of skeletal development and bone formation during the last decades, current knowledge on mechanisms underlying physiological and pathological processes is still insufficient to develop successful therapies for human skeletal and bone diseases. Thus, the use of alternative non-mammalian animal models, e.g. zebrafish, *Xenopus* and chicken, to study skeletal development and bone mineralization has been sought. Fish species – mainly the zebrafish and Japanese medaka – have received a particular attention from the scientific community. Their similarity with mammals regarding biochemical, developmental, physiological mechanisms, in particular those underlying skeletal development and bone mineralization, as well as the presence in fish genome of orthologs for most mammalian genes and a number of technical advantages (e.g. large progeny, external reproduction fast growth and translucent larvae) make fish models a suitable complement or alternative to mammalian models (Fisher et al., 2003; Belloni et al., 1996; Nissen et al., 2006; Laizé et al., 2015). Several *in vitro* cell systems of fish origin have been developed to complement *in vivo* systems, and various bone-derived cell lines capable of extracellular matrix

mineralization and representing different bone cell types and fish species have been established in the last decade (Pombinho et al., 2004; Rafael et al., 2010; Tiago et al., 2014; Vijayakumar et al., 2013). Mineralogenic cell lines developed from gilthead seabream vertebra (VSa13 and VSa16) were used to identify genes differential expressed during *in vitro* mineralization and possibly involved in mechanisms of bone cell differentiation and endochondral ossification in fish (Fonseca et al., 2011; Fonseca et al., 2007; Tiago et al., 2011). In a recent report by Tiago et al. (2011), expression of a betaine-homocysteine S-methyltransferase (BHMT) gene was strongly down regulated (>10 folds) during extracellular matrix mineralization and was proposed to play a central role in the mechanisms underlying the mineralogenic capacity of these cells. BHMT is a cytosolic and zinc-dependent enzyme (McKeever et al., 1991) that catalyses the remethylation of homocysteine (Hcy) into methionine (Met) using betaine as methyl donor, (Finkelstein et al., 1972). Primarily produced in liver and kidney (Sunden et al., 1997), BHMT is involved in the branch-point metabolism of Hcy, along with methionine synthase (MS). Hcy metabolism has been the subject of a particular attention since hyper-homocysteinemia (HHcy; high levels of Hcy in plasma) is associated in human with vascular diseases, renal insufficiency, non-insulin dependent diabetes, adverse pregnancy outcomes, Alzheimer's disease and osteoporosis (Mizrahi et al., 2002; Herrmann et al., 2005; Wijekoon et al., 2005; Heil et al., 2000; Nelen et al., 2000; McGregor et al., 2001). In a bone context, HHCy was reported to hamper bone formation by inhibiting osteoblast differentiation and/or activity, and to trigger bone resorption by stimulating osteoclast formation/differentiation and activity (Herrmann et al., 2008; Kriebitzsch et al., 2011). On the basis of existing data, we hypothesize that BHMT could play a role in mineralization through its ability to convert homocysteine into methionine and therefore limit its anti-osteogenic/pro-resorptive properties. The aim of the present study was to characterize sites and levels of BHMT gene expression throughout gilthead seabream development, in adult tissues and during *in vitro* mineralization, and to identify transcriptional factors responsible for the regulation of its expression. Taxonomic distribution and molecular phylogeny of BHMT proteins in vertebrates will also be assessed.

4.1.3 Materials and methods

4.1.3.1 Larvae, juvenile and adult fish culture

Eggs collected from natural spawning of gilthead seabream *Sparus aurata* (IPMA, Olhão, Portugal) were placed at 16°C in a closed recirculating system of 32-ppm seawater under a 12:12-h light-dark photoperiod. After mouth opening, larvae were fed with rotifers (enriched with cultured marine microalgae *Tetraselmis suecica* and *Isochrysis galbana)* until 20 days post-hatching (DPH) and with freshly hatched *Artemia* nauplii from 20 to 50 DPH. Juvenile and adult fish were reared at 16–20°C in 100-L seawater (32-ppm) tanks with a 12:12-h light-dark photoperiod, aeration of 100 ml air/min, and renewal flow of 1 tank/day; they were fed with artificial food (Sorgal).

4.1.3.2 Cell culture and extracellular matrix mineralization

Gilthead seabream bone-derived cell lines VSa13 and VSa16 were cultured as previously described by Pombinho et al. (2004). Extracellular matrix (ECM) mineralization was induced in confluent cultures by supplementing medium with 50 μg/ml of L-ascorbic acid, 10 mM of β-glycerophosphate and 4 mM of $CaCl_2$ (all from Sigma-Aldrich) for 4 weeks. Culture medium was renewed twice a week. Mineral deposition was revealed through von Kossa staining (Pombinho et al., 2004).

4.1.3.3 RNA preparation

When sampled for RNA preparation, larvae and juvenile fish were given a lethal anesthesia of 500 ppm of 2-phenoxyethanol (Sigma-Aldrich then washed twice in phosphate-buffered saline solution (PBS) and stored at -80ºC in 5 ml of TRIzol reagent (Sigma–Aldrich). Adult fishes were anesthetized (200 ppm of 2-phenoxyethanol) and sacrificed according to guidelines on animal experimentation available in Portugal. Tissues were sampled and washed twice in PBS and stored at -80ºC in 10 volumes of TRIzol reagent. Total RNA was extracted from cell cultures (three replicates per condition) using the method described by Chomczynski and Sacchi (1987) and from samples stored in TRIzol following manufacturer's instructions. RNA quantity was determined using a NanoDrop 1000 spectrophotometer (Thermo Scientific)

and five micrograms of each RNA sample were fractionated on 1% (w/v) agarose-formaldehyde gels for quality control.

4.1.3.4 *bhmt3* cDNA and gene cloning

bhmt3 cDNA was amplified by PCR from a Marathon cDNA library (Clontech) prepared from mRNA of adult gilthead seabream liver, kidney and testis using gene-specific primers (**Supplementary Table 4.1**) designed in available ESTs (GenBank accession No. FP336052, AM965012, FM151264, AM969436 and CX735005), adapter primers AP1 and AP2 and Advantage cDNA polymerase mix (Clontech). *bhmt3* gene was amplified by PCR from genomic DNA prepared from tissues of a single individual using DNeasy Blood & Tissue kit (QIAGEN), Adavantage *ThT* Polymerase mix (Clontech) and gene-specific primers (**Supplementary Table 4.1**). All PCR products were size-separated by agarose-gel electrophoresis, purified using illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare), cloned into pCR2.1-TOPO vector (Invitrogen) and sequenced on both strands.

4.1.3.5 Analysis of gene expression levels by quantitative real-time PCR

Quantitative real-time PCR (qPCR) was performed using iCycler iQ system (Bio-Rad). Total RNA (1 μg) was treated with RQ1 RNase-free DNase (Promega) then reverse-transcribed at 37ºC for 1 h using M-MLV reverse transcriptase, RNase-out (Invitrogen) and oligo-dT reverse primer. The reaction mixture containing cDNA (1μl of 1:10 dilution of reverse transcription reaction), 0.4 μM of gene-specific primers (**Supplementary Table 4.1**), and 1× iQ SYBR Green I mix (ABgene) was submitted to the following PCR conditions: an initial denaturation step at 95°C for 15 min then 50 cycles of amplification (each cycle is 30 s at 95°C, 45 s at 68°C). Levels of gene expression were calculated using the ΔΔCt method (Pfaffl, 2001) and normalized using level of *rpl27a* housekeeping gene expression.

4.1.3.6 *In situ* hybridization

Adult tissues and juveniles at 38 DPH were fixed for 24 h at 4ºC in 4% paraformaldehyde (PFA; pH 7.4 in PBS), washed 3 × 10 min with PBS, then dehydrated through a PBS/methanol gradient and stored in 100% methanol at 4ºC. When appropriate, tissues were passed through an increasing methanol/xylene series, embedded in paraffin then cross-sectioned. Sections (7-

µm thick) were collected on TESPA (3-aminopropyltriethoxysilane, Sigma–Aldrich) coated slides, dried for 4 h at 37ºC and kept at 4ºC until further processed. Sense and antisense RNA probes were generated from 1 µg of linearized pCR2.1-TOPO plasmid carrying partial BHMT3 cDNA (343 bp from 1135 bp to 1477 bp) using T7 or SP6 RNA polymerases then labeled with digoxigenin-dUTP (DIG RNA labeling kit, Roche Diagnostics). Riboprobes were treated with RNase-free DNase, ethanol precipitated and their integrity was assessed through agarose gel electrophoresis. Preparation of sections and *in situ* hybridization (ISH) were performed using the protocol previously described by Fonseca et al. (2011).

4.1.3.7 Histology

Tissues and cellular structures were identified through routine histological staining with trichromic azan combination (Chroma-Waldeck) following the protocol described in Fonseca et al. (2011). Bright field images were acquired using an Olympus IX81 microscope equipped with an Altra 20 camera. Skeletal tissues were classified according to Benjamin (1990) and Benjamin et al. (1992a) for cartilage types, and Faustino and Power (2001) and Genten et al. (2009) for bony tissues.

4.1.3.8 *bhmt3* promoter constructs

The 5'flanking region of gilthead seabream *bhmt3* gene was amplified from *Eco*RV and *Stu*I GenomeWalker DNA libraries using Advantage 2 Polymerase mix and 0.2 µM of adaptor primer 1 (AP1) and gene-specific primer SauBHMT3_Rv2_promoter. Nested PCR was performed using a 1:50 dilution of the first PCR and 0.2 µM of adaptor primer 2 (AP2) and gene-specific primer SauBHMT3_Rv1_promoter. PCR products were separated on agarose gel, purified using illustra GFX PCR DNA and Gel Band Purification kit then cloned into pCR2.1-TOPO and sequenced. Promoter region 898/+69 was then subcloned into pGL4.10 vector (Promega) upstream of firefly luciferase reporter gene using gene-specific primers SauBHMT3_promo_Rv_pGL4_HindIII and SauBHMT3_promo_Fw_pGL4_XhoI containing cleavage sites for endonucleases *Xho*I and *Hin*dIII. PCR primers are listed in **Supplementary Table 4.1**.

4.1.3.9 Cell culture and DNA transfection

Human embryonic kidney (HEK-293) cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Life Technologies) supplemented with 10% fetal bovine serum (Sigma-Aldrich), 1% penicillin-streptomycin (Life Technologies), 1% L-glutamine (Life Technologies) and maintained in a 5% CO_2 humidified atmosphere at 37°C. Gilthead seabream ABSa15 cells (European Collection of Cell Cultures No. 13112201) were cultured in DMEM supplemented with 10% fetal bovine serum and maintained in a 10% CO_2 humidified atmosphere at 33ºC (Tiago et al., 2014). The day before their transfection, cells were seeded in 24-well plates at a density of 5×10^4 cells per well (HEK-293) or in 12-well plates at a density of 8×10^4 cells per well (ABSa15) to achieve cultures at 50-60% confluence. Promoter construct and empty pGL4.10 vector were delivered into HEK-293 (125 ng) or ABSa15 (500 ng) cells using X-tremeGENE HP transfection reagent (Roche Diagnostics). When appropriate, 25 ng of pCMX.PL2 or pcDNA3 vectors expressing zebrafish Ets1 (v-ets avian erythroblastosis virus E26 oncogene homolog 1; KF774190), zebrafish Mef2C (myocyte enhancer factor 2C; BC059188 and EU825718 for a and b forms, respectively), human AP2 (activating enhancer binding protein 2 alpha; BC017754) and mouse SP1 (specific protein 1 transcription factor; AF022363) under the control of CMV promoter were co-transfected with *bhmt3* promoter construct. At 48 h post-transfection, cells were lysed and luciferase activity was measured using Dual-Luciferase Reporter Assay system (Promega) in a Synergy 4 microplate reader (BioTek). pRL-null vector (50 ng; Promega), which expresses Renilla luciferase (Rluc) and lacks promoter and enhancer elements, was used to normalize firefly luciferase (FLuc) activity in HEK-293 cells. pRL-SV40 vector (150 ng; Promega) was used to normalize FLuc activity in ABSa15 transfections. Normalized luciferase activity (FLuc/RLuc) is presented as fold change over control value and is the mean ± SD of at least four independent experiments performed in duplicates.

4.1.3.10 Sequence collection and molecular phylogeny

GenBank (ncbi.nlm.nih.gov) and Ensembl (ensembl.org) sequence databases were searched for nucleotide sequences showing similarity to gilthead seabream *bhmt* transcripts or genes using on-site BLAST facilities. Species-specific raw sequences (e.g. ESTs and WGSs) were aligned then assembled using ContigExpress module of Vector NTI Advance software (Invitrogen) to generate highly accurate consensus sequences. Virtual transcripts were

predicted from gene sequences through comparative homology analysis of previously annotated genes using stringent overlap criteria. Collected sequences with a complete coding sequence were aligned using TranslatorX V1.1 (Abascal et al., 2010) with parameters set to default and molecular phylogeny (maximum likelihood tree; 1000 bootstrap replicates) was constructed from sequence alignment using MEGA version 6 (Tamura et al., 2013).

4.1.4 Results

4.1.4.1 Gilthead seabream Bhmt3, a novel member of the BHMT protein family

The sequence of gilthead seabream *bhmt3* transcript (GenBank accession No. GU119905) was determined by PCR and RACE-PCR using gene-specific primers designed according to available EST sequences (**Supplementary Table 4.1**). The longest transcript amplified was 1527-bp long and coded for a 400 amino acids protein with an estimated molecular mass of 44 kDa (**Figure 4.1**). The sequence of the corresponding gene (GenBank accession No. GU119906) was determined by PCR using gene-specific primers designed according to cDNA sequence (**Supplementary Table 4.1**). It is 5076-bp long and composed of 8 coding exons and 7 introns (**Figure 4.1**). Gilthead seabream Bhmt3 protein deduced from cDNA and gene sequences is highly similar to human BHMTs, sharing 75.9% and 62.8% identity with BHMT1 (406 aa, NP_001704; note that BHMT1 is referred in many papers as BHMT) and BHMT2 (363 aa, NP_057877), respectively. Homocysteine domain of gilthead seabream Bhmt3 (from K^{8} to L^{309} according to the InterPro collection of protein signature databases) represents the main feature of the protein and includes binding sites for homocysteine, betaine and Zn^{2+}. Residues essential to BHMT protein function in mammals were also found in gilthead seabream Bhmt3: (i) $DGGFV^{23-27}$ and E^{154}, possibly involved in the binding to homocysteine (González et al., 2004), (ii) W^{41} and Y^{74} could interact with betaine carboxylates, (iii) Y^{155}, F^{256} and F^{262} may protect betaine methyl groups by creating a hydrophobic environment (González et al., 2004), and (iv) C^{212} and $CC^{294-295}$ are known to be critical to Zn^{2+} binding (Evans et al., 2002; González et al., 2003) and participate in motifs typical of thiol-seleniol methyltransferases, i.e. $GINCH^{209-213}$ and $VRYIGGCCGFEPYH^{289-301}$ (Heil et al., 2000). Two domains typically conserved among BHMTs were also found in gilthead seabream Bhmt3: a dimerization arm (aa 314-349) important for the interaction with

opposite monomer (in L7 loop), and a hook region (aa 350-364) central to tetramer formation (Szegedi and Garrow, 2004). *In silico* analysis of protein primary structure indicated the presence in gilthead seabream Bhmt3 of residues central to the remethylation of homocysteine into methionine, possibly using betaine as a methyl donor, suggesting that it represents a novel member of BHMT protein family.

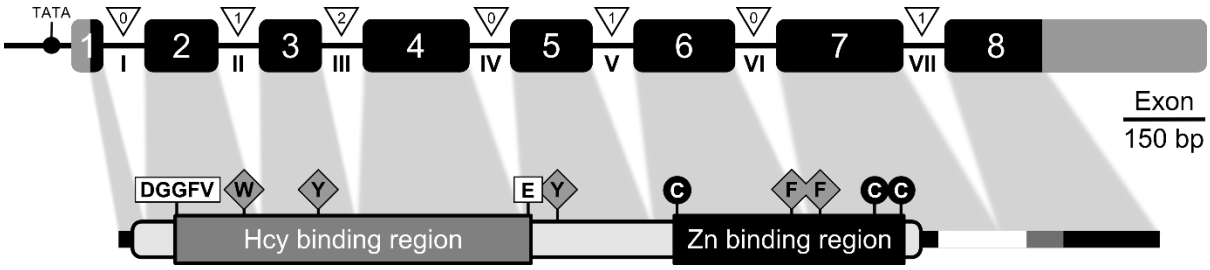


Figure 4.1. Schematic representation of gilthead seabream betaine-homocysteine S-methyltransferase 3 gene and protein structures. In gene structure exons are displayed as *black boxes* (coding sequences) or *grey boxes* (non-coding sequences) and indicated with *arabic numbers*. Introns are displayed as *solid black lines* and indicated with *roman numbers*. Phase intron insertion is indicated in *white triangles*. In the protein structure, DGGFV$^{23-27}$ and E^{154} indicate the fingerprint sequence and residue that are signals for homocysteine binding; W^{41} and Y^{74} indicate residues known to interact with betaine carboxylates and Y^{155}, F^{256} and F^{262} residues that create an hydrophobic environment protecting the betaine methyl groups; C^{212} residue in the well conserved GINCH sequence and CC$^{294-295}$ in the conserved region RYIGGCCGFEPYH are cysteine residues essential for zinc binding; TATA, vertebrate TATA binding protein factor. *Light grey box* indicates homocysteine domain, *white line* indicates the dimerization arm and *grey line* the hook region. *Dark grey box* and *black box* indicates Hcy and zinc binding regions, respectively.

4.1.4.2 *bhmt3* expression is reduced upon extracellular matrix mineralization

Levels of *bhmt3* expression were determined by qPCR in control and mineralizing VSa16 and VSa13 cell cultures (extracellular matrix mineralization was confirmed by von Kossa staining; **Figure 4.2**). A strong decrease of *bhmt3* expression (146 folds in VSa13 cells and 127 folds in VSa16 cells) was observed in cells undergoing ECM mineralization *versus* control cells (**Figure 4.2**), suggesting a central role in this process. Interestingly, while levels of expression in control VSa13 cells was similar to that in vertebrae, it was much lower in control VSa16 cells (approximately 310 folds), indicating that background levels is cell type specific and higher in cells of the chondrocytic lineage than in cells of the osteoblastic lineage.

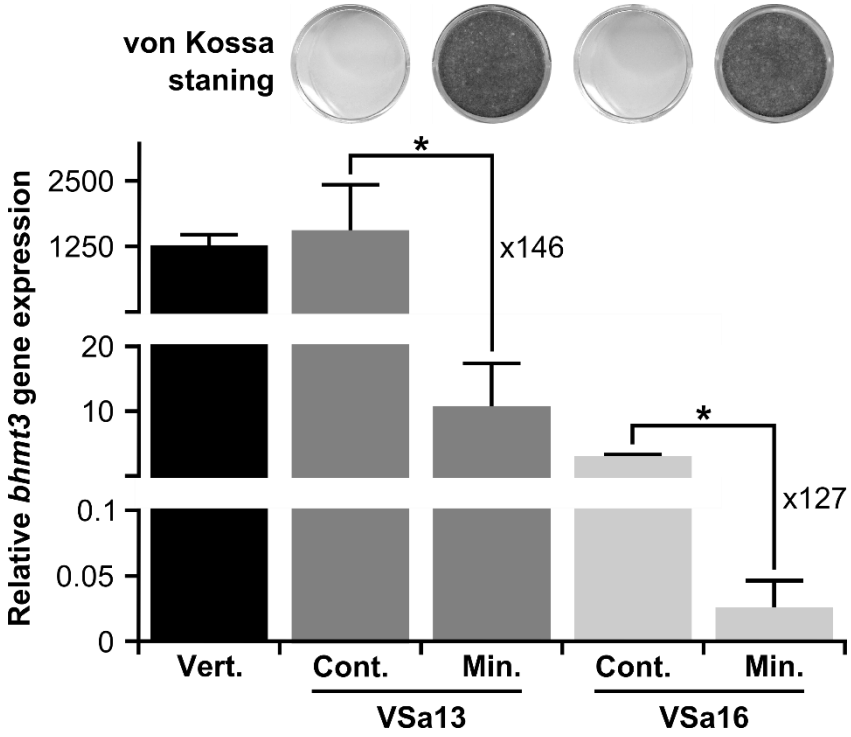


Figure 4.2. Relative expression of *bhmt3* gene upon mineralization of the extracellular matrix of gilthead seabream VSa13 and VSa16 cell lines. Transcript levels were determined by qPCR from three replicates and normalized using *rpl27a* housekeeping gene. *Vert.*, vertebra; *Cont.*, control cell culture; *Min.*, mineralized cell culture. *Asterisks* indicate values statistically different control conditions (Student's *t* test; $P < 0.05$)

4.1.4.3 *bhmt3* expression is restricted to bone and cartilage tissues

Spatial and temporal expression of *bhmt3* was determined during gilthead seabream development – from embryonic stages throughout adulthood – and in a selection of adult tissues. *bhmt3* transcript was first detected at blastula stage (512 cells) before the onset of *de novo* transcription, indicating that it may be maternally inherited (**Figure 4.3a**). Intermediate and high levels of gene expression were observed from 20 DPH to 48 DPH and at 130 DPH, respectively. At mouth opening (5 DPH) and during larval to juvenile transition (70 DPH) gene expression decreased to low or almost undetectable levels, respectively, suggesting that Bhmt3 function may not be needed or is undesired at these particular stages. In adult tissues, *bhmt3* transcript was clearly associated to calcified tissues, with levels of expression in these tissues (with the exception of the spiny rays) higher than those observed in soft tissues (**Figure 4.3b**). Expression of *bhmt3* in vertebra – the calcified tissue with highest levels of expression – was 34, 64 and 68 times higher than expression in kidney, brain and liver, which are the soft tissues exhibiting highest BHMT gene expression in mammals. Pattern of expression in adult tissues

further support the hypothesis that Bhmt3 has a central role in mechanisms underlying skeletal tissue homeostasis.

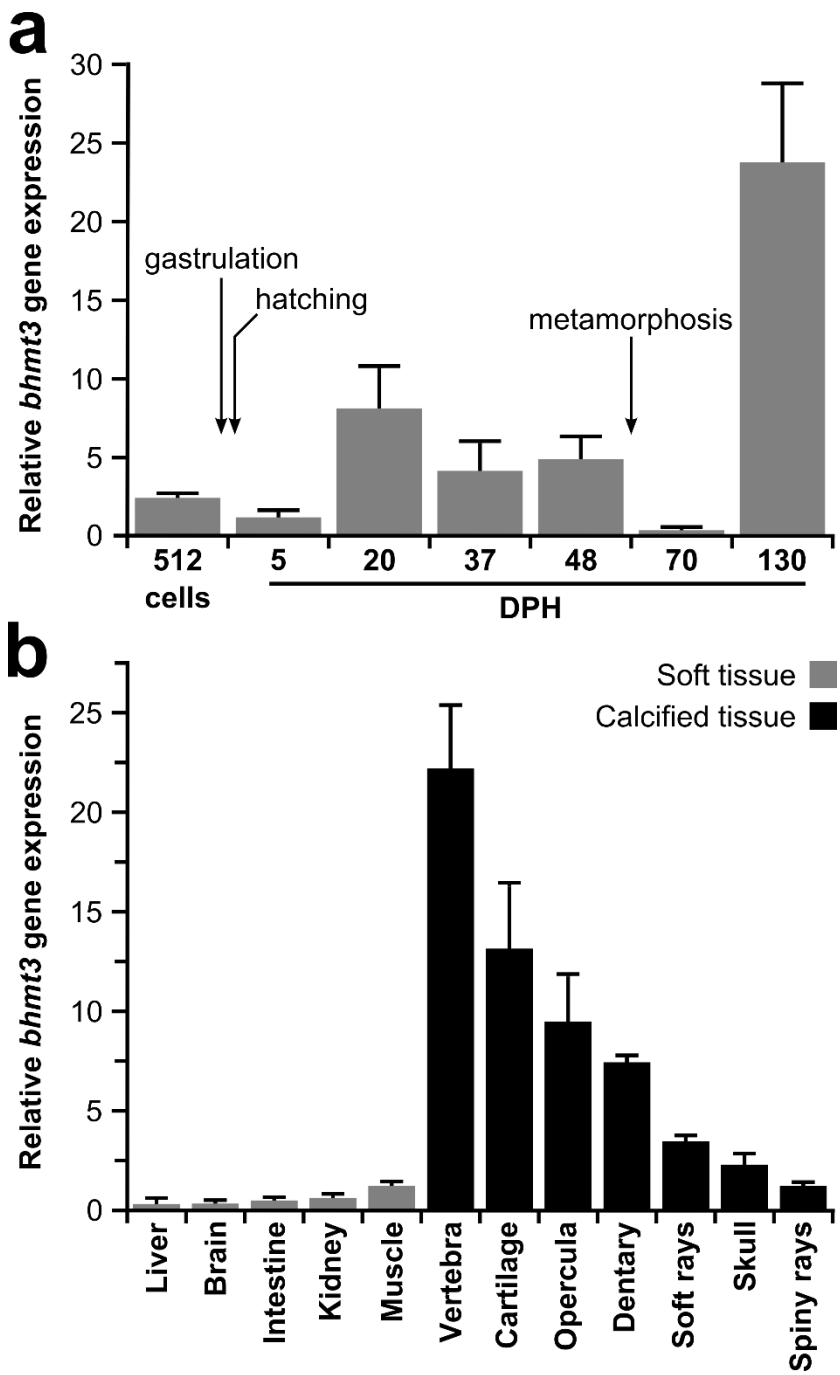


Figure 4.3. Relative expression of *bhmt3* gene throughout gilthead seabream development (**a**) and in adult tissues (**b**). Transcript levels were determined by qPCR from three replicates and normalized using *rpl27a* housekeeping gene. *DPH*, days post-hatching

To confirm these data, and unveil the cellular component expressing *bhmt3*, *in situ* hybridization (ISH) was performed on sections of 38-DPH juvenile specimens and adult

vertebra and branchial arches (representing bone and cartilage, respectively) and consecutive sections were stained with azan trichrome to identify tissue structures (**Figure 4.4**).

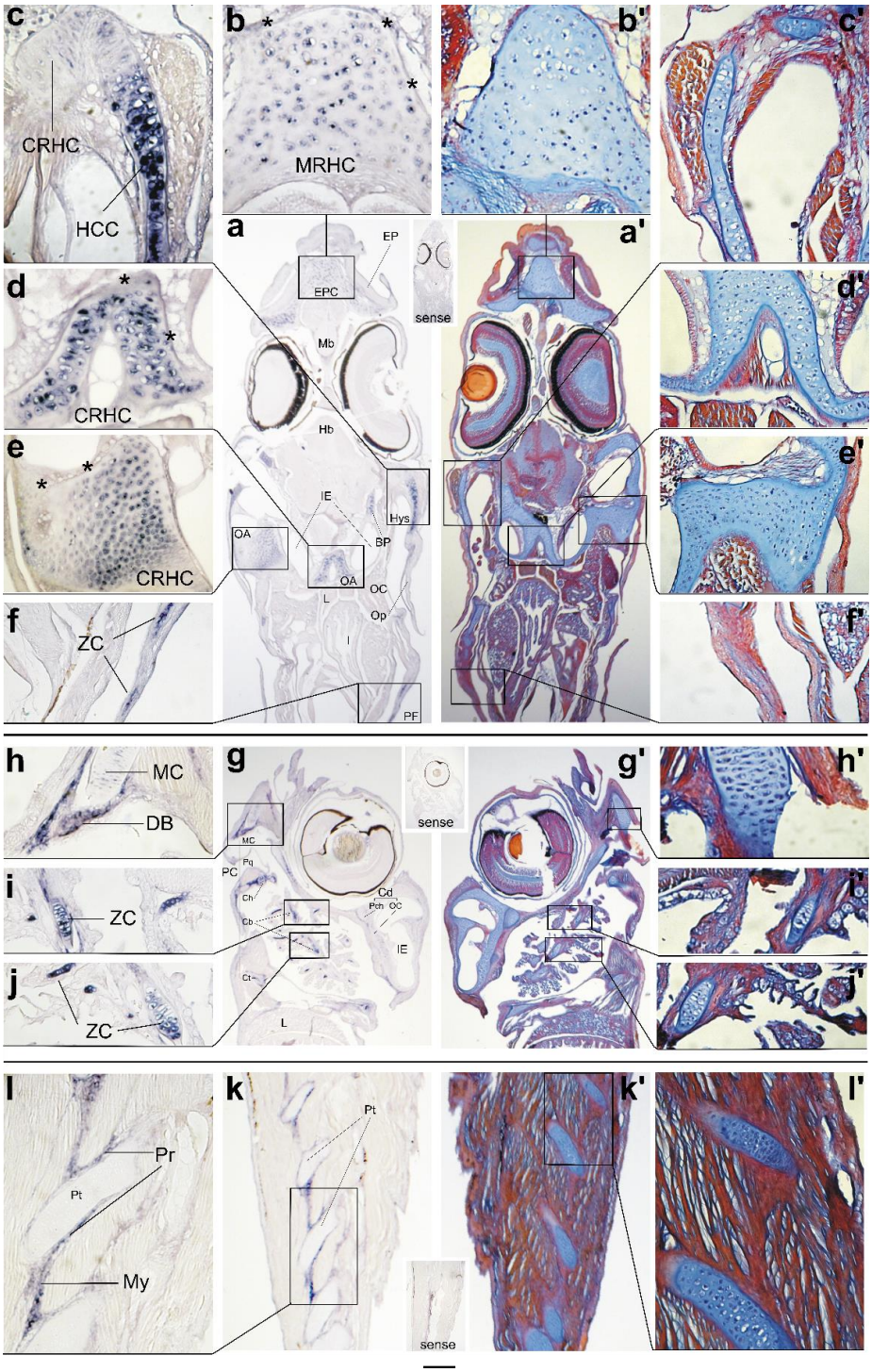


Figure 4.4. Sites of *bhmt3* gene expression by *in situ* hybridization (ISH) in a 38-DPH gilthead seabream. **a-l**, ISH using BHMT3 antisense riboprobe. **a'-l'**, azan staining. *Sense*, ISH using *bhmt3* sense riboprobe. *Top panel*: **a-f** and **a'-f'** coronal (dorso-ventral) and mirror of coronal sections, respectively. *Middle and bottom panel*: **g-l** and **g'-l'**, transversal and mirror of transversal sections, respectively. *BP*, basal plate; *Cd*, chondrocranium, *CRHC*, cell-rich hyaline cartilage; *Cb*, ceratobranchial; *Ct*, cleithrum; *Ch*, ceratohyal; *DB*, dentary bone; *EP*, ethmoid plate; *EPC*, epiphysial

cartilage; *Hb*, hindbrain; *HCC*, hyaline cell cartilage; *Hys*, hyosimplectic cartilage; *IE*, inner ear; *I*, intestine; *L*, liver; *Mb*, midbrain; *MC*, Meckel's cartilage; *My*, myotome; *MRHC*, matrix-rich hyaline cartilage; *NA*, neural arch; *OA*, occipital arch; *OC*, otic capsules; *Op*, operculum; *PC*, paryngeal cartilage; *Pch*, parachordal; *PF*, pectoral fin; *Pq*, palatoquadrate; *Pr*, perichondrium; *Pt*, pterygiophores; *ZC*, zellknorpel cartilage. *Asterisks* indicate bone collar. Bar represents: 1) 300 µm in sense images; 2) 100 µm in images a, a', g and g'; 3) 75 µm in images k and k'; and 4) 25 µm in images b-f, b'-f', h-j, h'-j', l and l'.

A strong and specific signal from *bhmt3* riboprobe was observed in several cartilaginous structures undergoing endochondral ossification at juvenile stage: (1) the ethmoid cartilage (EC), a typical matrix-rich cell cartilage (MRHC) of the ethmoid plate (EP), located in the anterior region of chondrocranium cartilage (Cd); (2) the hyosimplectic cartilage (Hys) that articulates posteriorly with the operculum (Op) and contains hyaline cell cartilage (HCC) and cell-rich hyaline cartilage (CRHC); (3) the occipital arch cartilage (OA), another element of the chondrocranium cartilage, also rich in cells (CRHC); (4) the parachordal cartilage (Pc), located in the basal region of the chondrocranium cartilage and later replaced by the prootic bones; and (5) the otic capsules cartilage (OC), a bilaterally paired cartilage encasing the semicircular canals of the inner ear (IE). Structures formed through perichondral ossification also exhibited intense and specific signal from *bhmt3* riboprobe: (1) the pharyngeal cartilage (PC), which is part of the splanchnocranium cartilage and contains structures such as the Meckel's cartilage (MC), (2) the palatoquadrate cartilage (Pq), (3) the ceratohyal cartilage (Ch), (4) the ceratobranchial cartilage (Cb), (5) the perichondrium (Pr) surrounding pterygiophores (Pt) of the dorsal fin and (6) the zellknorpel cartilage (ZC) of pectoral fin (PF). The only bony structure already formed showing signal for *bhmt3* riboprobe was the dentary bone (DB) located in the vicinity of MC. In all the cartilaginous structures analyzed, both hypertrophic and mature chondrocytes – active participants of the ossification process – were strongly labeled, while immature and proliferative chondrocytes showed a residual positive signal. No signal was observed in the uncalcified scleral cartilage and in pterygiophores of the pectoral fin. Adult vertebra and branchial arches (**Figure 4.5**) were also analyzed by ISH to determine whether sites of *bhmt3* expression were maintained throughout adulthood. Osteoblasts (Ob) present in the trabecular bone (TB) of vertebral bodies were strongly labeled, as well as chondrocytes present in the zellknorpel cartilage (ZC) of branchial arches (BA). Expression data collected through ISH and qPCR were highly consistent, with bone and cartilaginous tissues being the sites of higher *bhmt3* expression and with liver, kidney, intestine and brain lacking *bhmt3* expression. Myotomes (My), groups of muscles surrounding the

pterygiophores, were the only soft tissue positive for *bhmt3* expression by ISH in agreement with muscle being, the soft tissue that showed higher expression levels by qPCR.

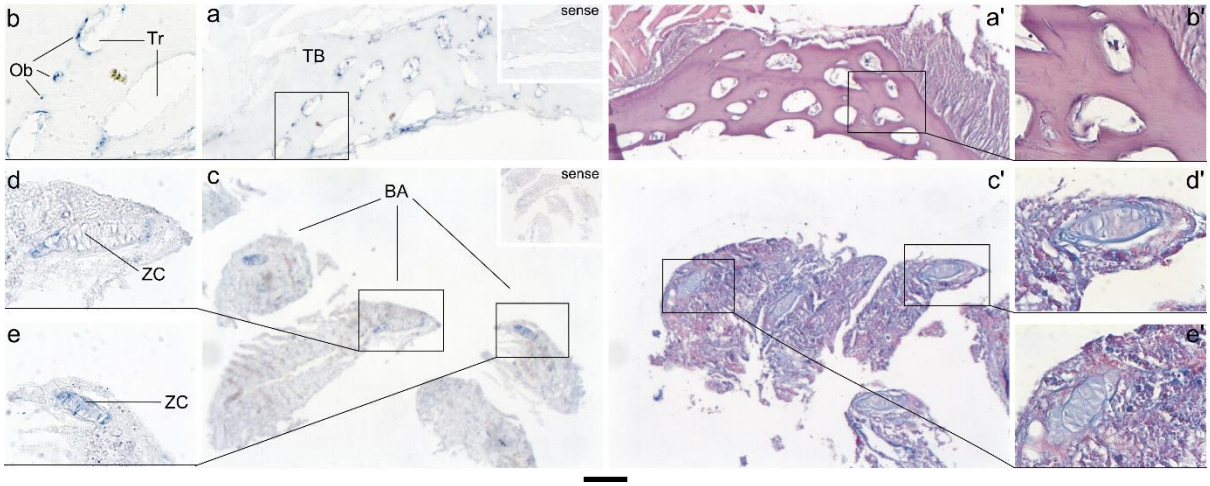


Figure 4.5. Sites of *bhmt3* gene expression by *in situ* hybridization (ISH) in adult gilthead seabream vertebra (**a-b**, **a'-b'**) and branchial arches (**c-e**, **c'-e'**). **a-e**, ISH using BHMT3 antisense riboprobe. **a'-e'**, azan staining. *Sense*, ISH using BHMT3 sense riboprobe. *BA*, branchial arches; *Ob*, osteoblast; *TB*, trabecular bone; *Tr*, trabecula; *ZC*, zellknorpel cartilage. Bar represents: 1) 60 µm in sense images; 2) 20 µm in **a-a'** and **c-c'** images; and 3) 7 µm in **b-b'**, **d-d'** and **e-e'**.

4.1.4.4 Expression of *bhmt4, 5, 6* and *7* are restricted to soft tissues

Restriction of gilthead seabream *bhmt3* expression to calcified tissues was intriguing given expression patterns reported in mammals for BHMTs, i.e. mostly in liver, and to a lesser extent in kidney, brain and muscle. Four additional *bhmt* genes have been identified in gilthead seabream: *bhmt4* (DQ470488), *bhmt5* (GU597054), *bhmt6* (GU197551) and *bhmt7* (GU197552). To test the hypothesis that one of them, or more, may be expressed in non-calcified tissues, patterns of *bhmt* expression in adult gilthead seabream tissues were determined by qPCR (**Figure 4.6**). While *bhmt4* and *bhmt5* were mainly expressed in liver and to a lesser extent in kidney, *bhmt6* and *bhmt7* transcripts were found mostly in kidney and intestine, respectively, and to a lesser extent in intestine and brain. Interestingly, *bhmt4* and *bhmt5* on one hand and *bhmt6* and *bhmt7* on the other hand shared similar patterns of expression and high sequence similarity, suggesting that they may share a closer evolutionary relationship, an hypothesis that will be evaluated through the molecular phylogenetic analysis of vertebrate BHMT sequences (see below). Although a state-of-the-art analysis should be performed to be able to precisely compare levels of *bhmt* expression and although more tissues

should be tested, *bhmt4*, *5*, *6* and *7* were apparently not or weakly expressed in calcified tissues while *bhmt3* expression was detected mostly in calcified tissues.

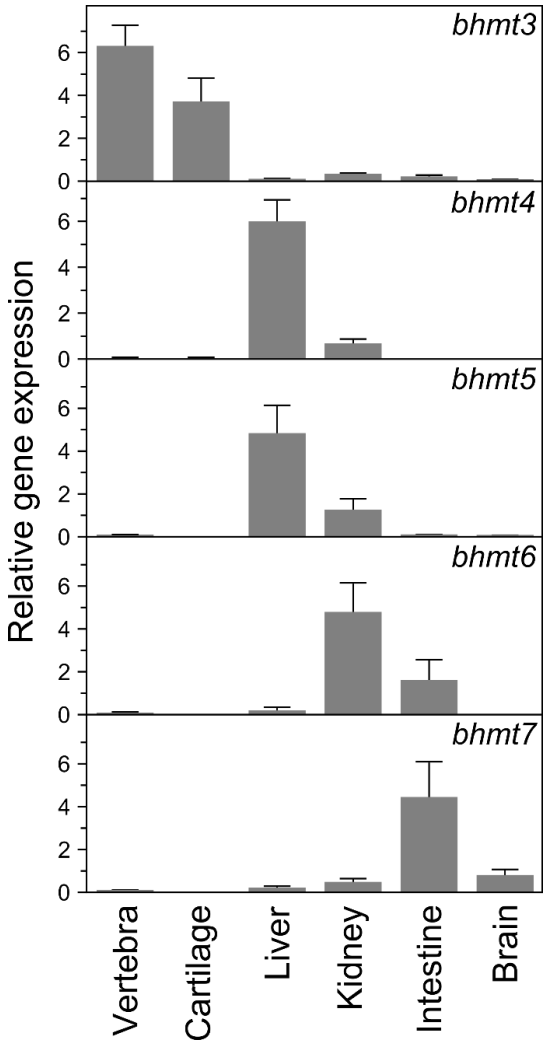


Figure 4.6. Relative expression of gilthead seabream *bhmt* genes in adult vertebra, cartilage, liver, kidney intestine and brain. Transcript levels were determined by qPCR from three replicates and normalized using *rpl27a* housekeeping gene.

4.1.4.5 SP1, ETS1 and MEF2C are regulators of *bhmt3* transcription

Absence of overlap in expression patterns suggested a tight control of the mechanisms regulating *bhmt3* expression and limiting its presence to calcified tissues; regulation of *bhmt3* expression by bone and cartilage-related transcription factors was therefore evaluated. The 5′-flanking region of gilthead seabream *bhmt3* gene was analyzed using MatInspector, Transfac, TFsearch and Jaspar online tools to identify *cis*-regulatory elements that may be involved in transcriptional regulation. All the identified motifs are numbered according to the transcription

start site (+1), determined from the longest transcript. In addition to the canonical TATA box (TATAAA) located at -45/-31 bp, various transcription factor (TF) binding sites were predicted with a high score; some of these TFs have been associated with the regulation of bone and cartilage functions: MEF2 (3 sites: -615/-592, -481/-458, -418/-395), SP1 (2 sites: -76/-70, -8/-2), ETS1 (2 sites; -139/-118, -116/-95) and AP2 (1 site: -117/-102) (**Figure 4.7a**).

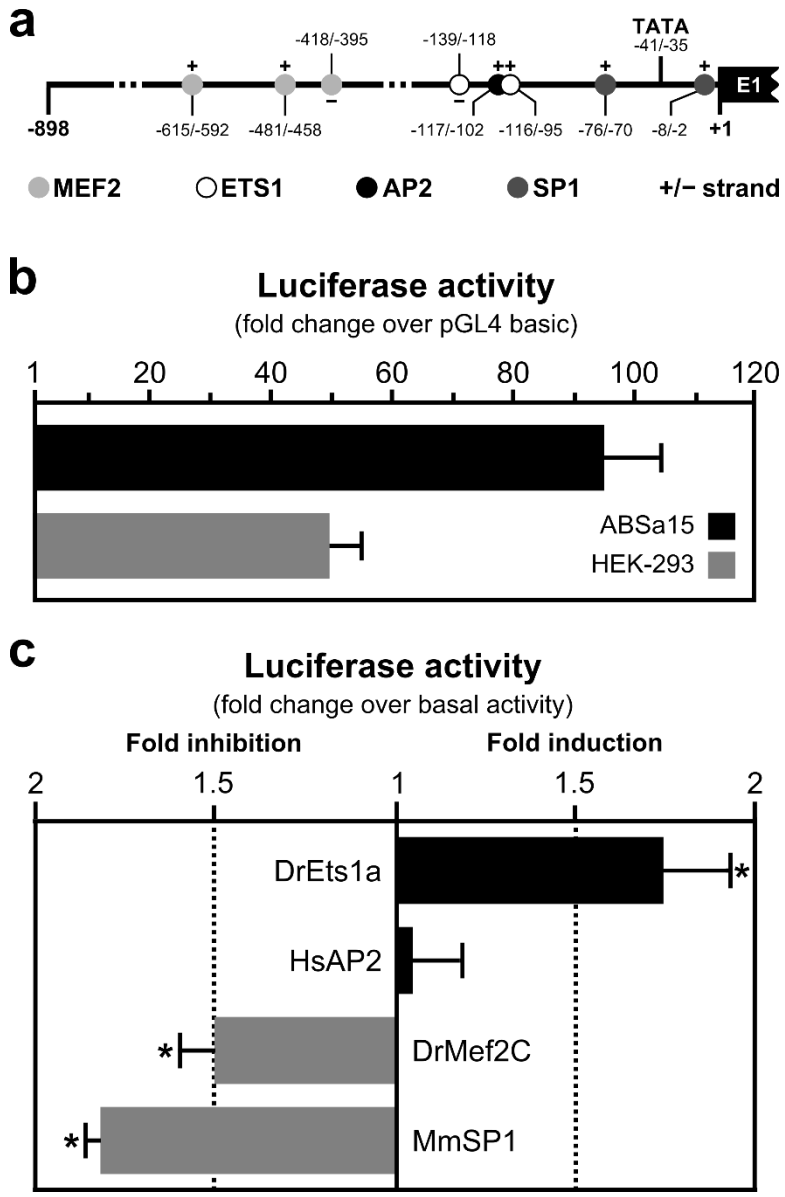


Figure 4.7. (**a**) *In silico* analysis of gilthead seabream *bhmt3* promoter and localization of putative binding sites for MEF2, ETS1, AP2 and SP1 transcription factors. (**b**) Transcriptional activity of gilthead seabream *bhmt3* promoter in ABSa15 and HEK-293 cells. Luciferase activity (FLuc/RLuc; n = 5) was measured under basal conditions and is presented as fold changes over the empty pGL4 vector. (**c**) Transcriptional activity of gilthead seabream *bhmt3* promoter in HEK-293 cells upon overexpression of zebrafish Ets1a, human AP2, zebrafish Mef2C or mouse SP1. Luciferase activity (FLuc/RLuc; n = 5) is presented as fold changes over the empty pCMX expression vector. *Asterisks* indicate values statistically different control conditions (one-way ANOVA with Tukey's post-test; $P < 0.05$)

Ability of these factors to regulate *bhmt3* transcription was assessed using a construct where promoter region was cloned upstream of firefly luciferase gene. Promoter activity was evaluated in two cellular hosts – the gilthead seabream ABSa15 cell line (capable of *in vitro* mineralization but hard-to-transfect cell line) and the human embryonic kidney HEK-293 cell line (high efficiency of transfection but not mineralogenic) – and normalized using values of the Renilla luciferase expressed by the promoter-less pRL-null vector. Basal activity of *bhmt3* promoter was measured in both cell lines and shown to be 2 times higher in ABSa15 cells than in HEK-293 cells (**Figure 4.7b**), possibly because of the mineralogenic and/or piscine nature of the ABSa15 cells and therefore the presence of specific transcription factors essential to promoter activation. Although *bhmt3* promoter was less active in these cells under basal conditions, HEK-293 cells were chosen to host the co-transfection assays because they revealed a higher and more constant rate of transfection (critical for co-transfection assays where 3 different vectors need to be efficiently delivered intracellularly) and it resulted therefore in a lower variability in luciferase values. Expression vectors carrying the coding sequence of zebrafish *ets1a* (pCMX-ETS1, zebrafish *mef2C* (pCMX-MEF2C), human *AP2* (pcDNA3-AP2) and mouse *Sp1* (pcDNA3-SP1) were independently co-transfected with *bhmt3* promoter construct. While AP2 failed to alter promoter activity, expression of SP1, Ets1 and Mef2C resulted in a mild but significant effect on luciferase activity, i.e. 1.7 fold stimulation by Ets1 and 1.5 and 1.8 fold repression by Mef2C and SP1, respectively (**Figure 4.7c**).

4.1.4.6 Explosive diversity of *bhmt* gene family in Neoteleostei

Sequences highly similar to gilthead seabream Bhmt3, 4, 5, 6 and 7 were retrieved from GenBank and Ensembl databases using on-site BLAST tools. Sequences were clustered by species and assembled to construct consensus sequences, which BHMT nature was confirmed through homology-based search in GenBank non-redundant protein sequence database. A tree was constructed using collected data, and the presence of the different BHMT genes in particular taxonomic group is indicated (**Figure 4.8**).

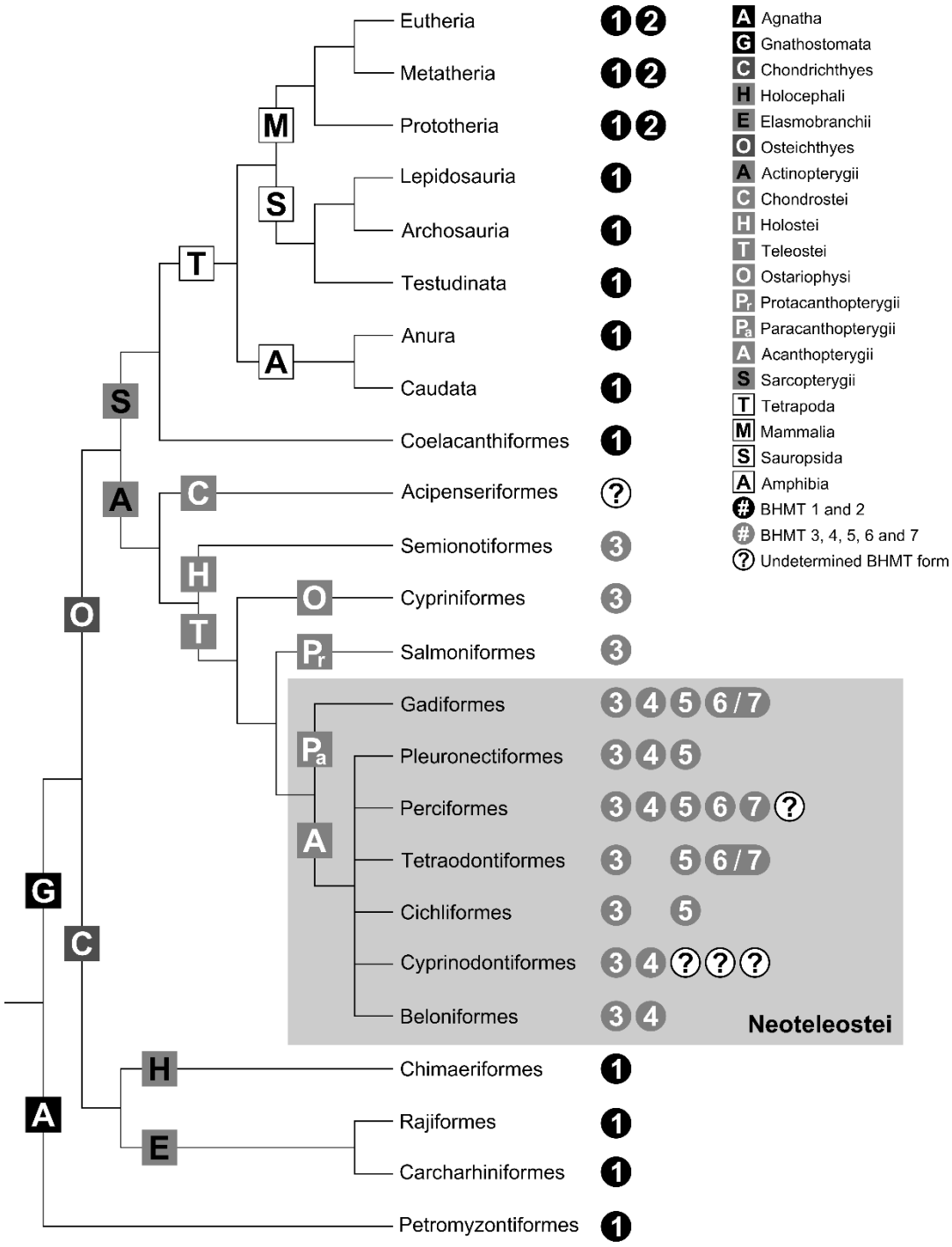


Figure 4.8. Taxonomic distribution of BHMTs (simplified from the Tree of Life at tolweb.org) throughout vertebrate taxonomy. Presence/absence of BHMT genes were inferred from sequence data collected from NCBI and Ensembl sequence databases.

While a single gene was identified in the genome of most vertebrate species (i.e. Sauropsida, Amphibia, Coelacanthiformes, Chondrostei, Holostei, Ostariophysi, Protacanthopterygii, Chondrichthyes and Agnatha), two genes were found in the genome of mammalian species (i.e. Eutheria, Metatheria and Prototheria), and from 2 to 6 genes were

observed in the genome of neoteleostean species (i.e. Gadiformes, Pleuronectiformes, Perciformes, Tetraodontiformes, Cichliformes, Cyprinodontiformes and Beloniformes). The occurrence of different number of genes in vertebrate genomes suggested gene duplication/loss events throughout evolution, in particular in the Neoteleostei.

To better understand the mechanisms underlying the loss or gain of *BHMT* genes, e.g. whether these events happened at genome, chromosome or gene level and whether there are lineage-specific, genomic regions flanking *BHMT* gene loci were analyzed for gene content and distribution in a subset of species representing several vertebrate taxonomic groups (**Figure 4.9**). In this regard, a remarkable conservation of the genomic region surrounding *BHMT* locus in Ostariophysi (zebrafish), Holostei (spotted gar) and Tetrapoda (Western clawed frog, green anole, chicken, mouse and human) was observed. In mammals, where two genes have been evidenced, *BHMT* loci were side-by-side arranged in a tandem repeat with identical orientations. None of the genes flanking mammalian *BHMT* locus have been duplicated suggesting that duplication event was locus-specific. The scenario is somewhat more complicated in Neoteleostei, where 2-6 genes were identified. In the genome of these species, *bhmt3* gene is flanked by the same genes surrounding *BHMT* locus in non-Neoteleostei species, suggesting that they may be orthologous. In the Atlantic cod, *bhmt4* and *bhmt5* were found together on the same chromosome, while only one locus (either 4 or 5) was observed in other neoteleostean genomes. In the three-spined stickleback (and possibly in the Atlantic cod if scaffolds 1288 and 1979 are on the same chromosome), *bhmt4* and *bhmt6/7* (either *bhmt6* or *bhmt7*) were found on the same chromosome, although in regions separated by various genes and several Mb of sequence. Although genomic information is scarce and possibly incomplete, it appears that some *bhmt* forms were not found in the genome of some neotelostean species at places identified in other species, e.g. *bhm6/7* was absent in the Japanese medaka and pufferfish while genes flanking the locus in the Atlantic cod and the three-spined stickleback were identified. Although we cannot exclude that these genes may be present at different locations or are yet to be discovered, species or linage-specific gene duplication/loss event may be at the origin of the complex distribution of *bhmt* genes in Neoteleostei. Blast analysis of the genomic regions in Tetrapoda, Holostei and Ostariophysi genomes that contain genes typically flanking *bhmt4/5* and *bhmt6/7* in Neoteleostei did not reveal un/mis-annotated *BHMT* genes or remnants of *BHMT* genes (e.g. pseudogenes), favoring the hypothesis of multiple and lineage-specific gene duplication events in the genomes of neoteleostean species.

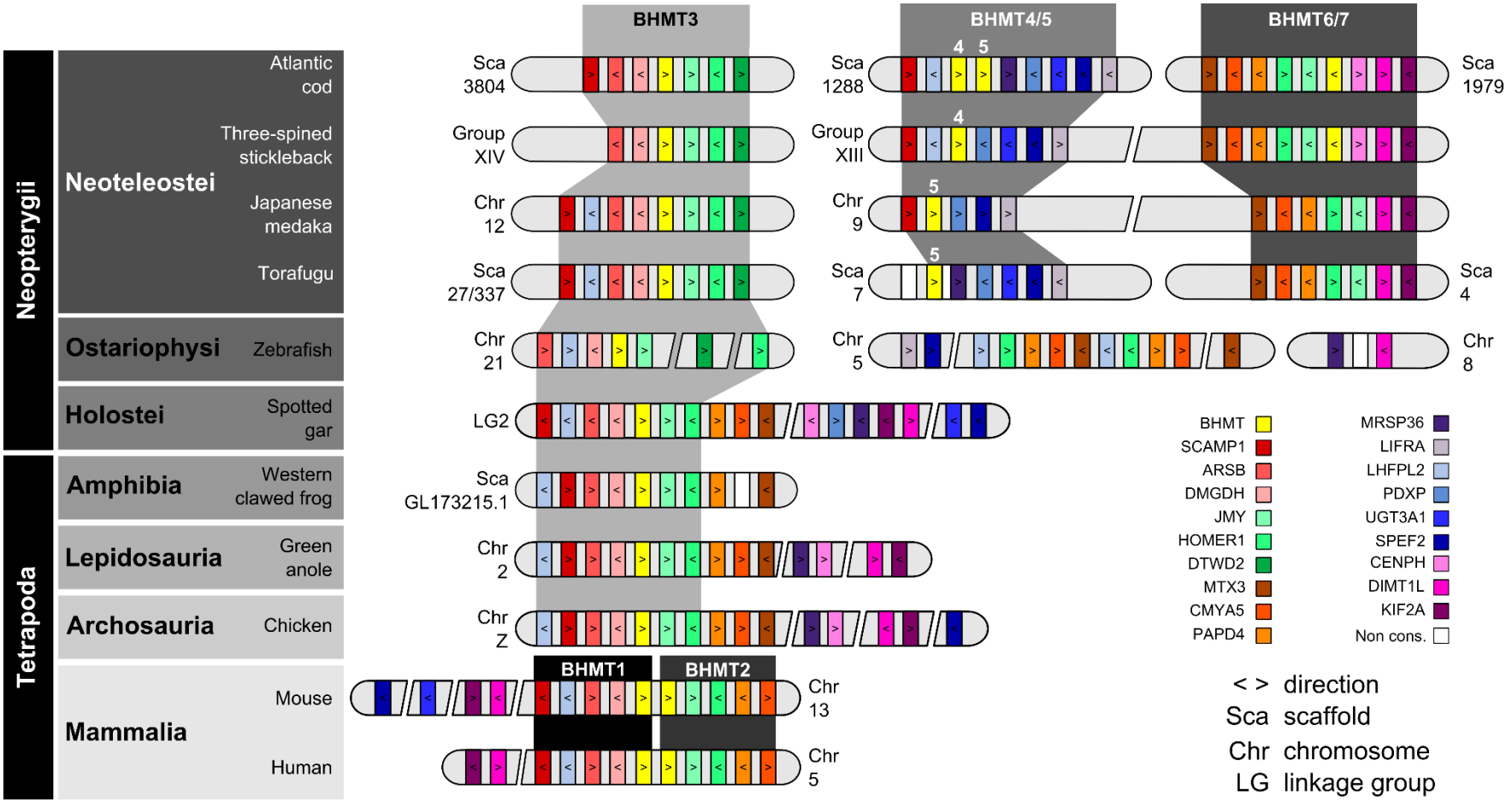


Figure 4. 9. Schematic representation of the genomic region flanking vertebrate *BHMT* genes using data from Ensembl project. Genes present in the vicinity of *BHMT* loci (*yellow*) and their orientation are indicated in colored boxes.

4.1.4.7 Evolution of BHMT proteins

To get insights into the evolutionary relationship of vertebrate *BHMT* genes, a phylogenetic tree was constructed from a subset of 59 complete coding sequences retrieved from annotated sequences or reconstructed from EST or WGS sequences available in Genbank and Ensembl databases (**Figure 4.10** and **Supplementary Sequences 4.1**), and representing the main vertebrate taxa (i.e. cartilaginous fish, ray-finned fish, lobe-finned fish, amphibians, sauropsids and mammals). Two well-separated clades were observed, one comprising sequences from cartilaginous fish, lobe-finned fish, amphibians, sauropsids and mammals and another one comprising sequences from ray-finned fishes. The two *BHMT* genes identified in mammals clustered together and not with any other sequences and may represent paralogs that probably originated from a duplication event that occurred in the mammalian lineage after branching from the sauropsids. Interestingly, sequences of cartilaginous fish species were more closely related to those of tetrapods than to those of ray-finned fish indicating a higher degree of divergence in the latter sequences, which could be related to gene duplication events that occurred in this lineage and consequent redundancy of BHMT function in these species. Molecular phylogeny of *BHMT* genes also indicated that ray-finned fish sequences clustered in three main groups: *bhmt3* genes (group 1), *bhmt4* and *bhmt5* genes (group 2), and *bhmt6* and *bhmt7* genes (group 3). The close relationship between *bhmt4* and *5*, and between *bhmt6* and *7* could indicate recent duplication events, and the similarity between *bhmt3* and *bhmt4/5* sequences may point out toward a common ancestral gene, while *bhmt6/7* sequences would have originated through a more ancient duplication event that occurred early in the Neoteleostei lineage. The single gene observed in non-Neoteleostei species of ray-finned fish (e.g. Ostariophysi and Protacanthopterygii) clustered together with *bhmt3* indicating that *bhmt3* may be orthologous to other vertebrate genes, while *bhmt4*, *5*, *6* and *7* would be paralogs. *bhmt* gene from **t**he spotted gar (Holostei) clustered with other ray-finned fish sequences but apart from any other group identified previously. It would be orthologous to non-Neoteleostei genes and its position close to divergent Neoteleostei sequences, and not in-between cartilaginous fish and lobe-finned fish sequences, could indicate that rate of divergence of ray-finned fish sequences increased after tetrapod branching.

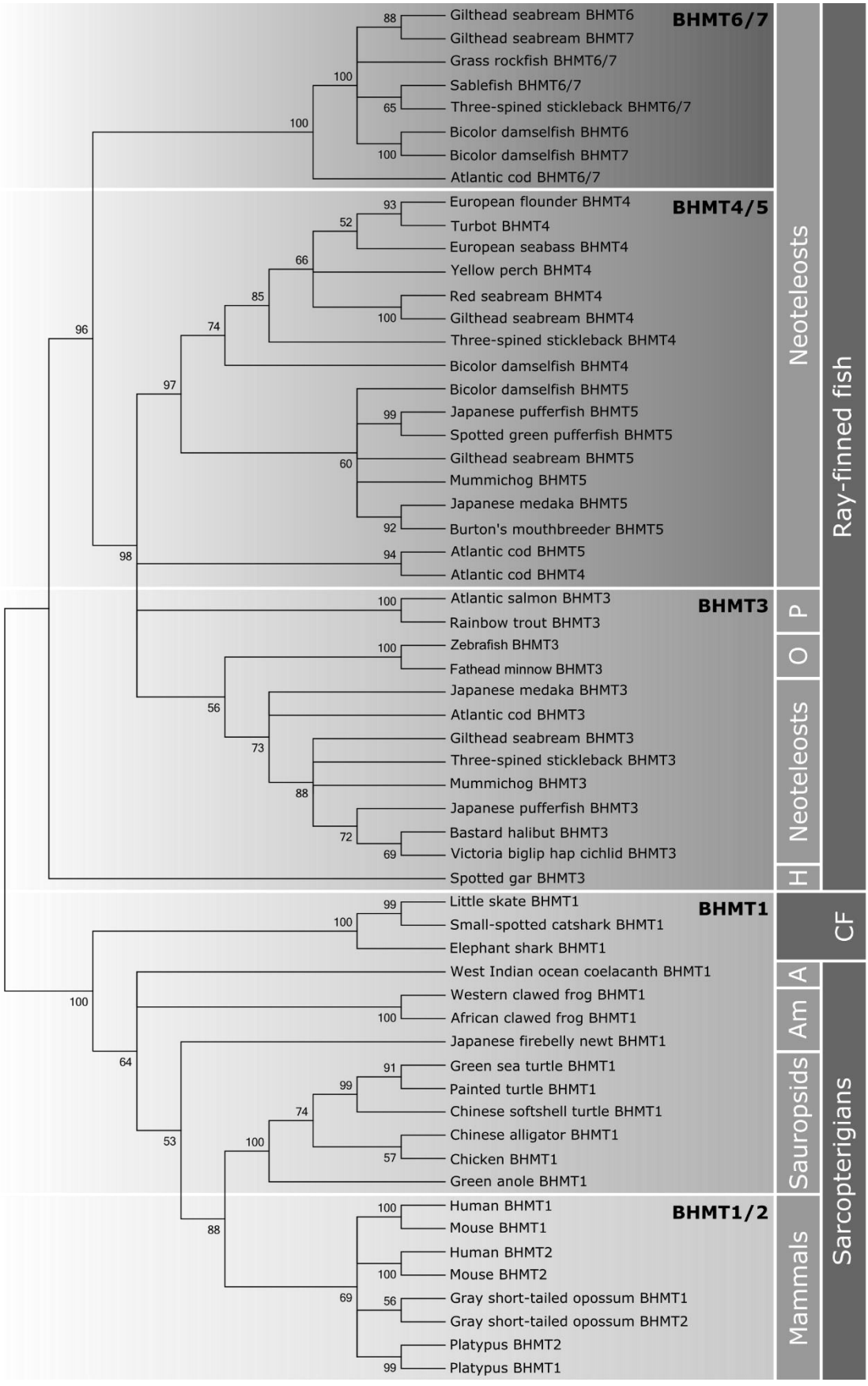


Figure 4.10. Molecular phylogenetic analysis of vertebrate *BHMT* genes by Maximum Likelihood method. The evolutionary history was inferred by using the Maximum Likelihood method based on the

Tamura-Nei model (Tamura and Nei, 1993). The tree (unrooted) with the highest log likelihood (-23757.7792) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The analysis involved 59 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 864 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

4.1.5 Discussion and Conclusions

The comparative analysis of gene and protein structures has revealed a high similarity between gilthead seabream Bhmt3 and vertebrate BHMTs. Same number and similar length for the exon, presence of amino acids and protein domains required for BHMT enzymatic activity and binding to substrate, homocysteine and Zn^{2+} but also for dimer and tetramer formation. In mammals, the substrate used as methyl donor is betaine (BHMT1) or S-methylmethionine (BHMT2) (Ganu et al., 2012; Szegedi et al., 2008; Chadwick et al., 2000). Although this remains to be tested in functional assays, we propose, given the highest similarity of Bhmt3 with BHMT1 (75.9% sequence identity *versus* 62.8% for BHMT2) and the presence in the N-terminal region of residues important for betaine recognition, which are absent in BHMT2 (e.g. Tyr^{77} and $RGNYVLEKI^{86–94}$ in human and rat proteins and Tyr^{74} and $RGNKQRF^{83-89}$ in gilthead seabream protein; Szegedi et al., 2008) that Bhmt3 may also use betaine as substrate. The partial conservation of these residues in gilthead seabream protein may indicate that the affinity for betaine is different or the usage of alternative substrate(s). *In vitro* studies (Szegedi et al., 2008; Bose et al., 2002) reported the ability of BHMT1 and BHMT2 to use various sulfonium compounds as substrates although with different catalytic efficiencies, while the use of betaine is limited to BHMT1. The possibility that BHMTs may use different substrates according to tissue localization was also reported (Pajares and Pérez-Sala, 2006; Ganu et al., 2011) and the restricted expression of *bhmt3* in calcified tissues could indicate the use of a different substrate instead of betaine.

The comparative analysis of *bhmt3* expression data with those available in other vertebrate species – mostly mammals – revealed several novel features specific for gilthead seabream protein. In mammals, *BHMT* transcripts were found mainly in liver and kidney (Chadwick et al., 2000; Garrow, 1996; McKeever et al., 1991) and to a lesser extent in pancreas (sheep; (Park and Garrow, 1999; Lu et al., 2001; Finkelstein and Martin, 1984; Finkelstein et al., 1986; Uthus et al., 2002)), lens (Rhesus monkeys; (Rao et al., 1998)) and hepatoma cells (human; (Skiba et al., 1982)), and have been associated to methionine metabolism and the

regulation of osmotic stress (processes carried out mainly in liver and kidney). Gilthead seabream *bhmt3* is not expressed in these soft tissues – it is mostly found in calcified tissues – and appeared to be involved in mechanisms of chondral ossification, a function that, to the best of our knowledge, has never been evidenced in any other vertebrate species. Whether this novel role is specific to gilthead seabream protein – and possibly to Neoteleost since zebrafish *bhmt* is expressed in bone, fin and gills according to UniGene Dr.75610 EST profiles at www.ncbi.nlm.nih.gov – or is found in other vertebrate species should be determined in future studies. Recent data reporting a role of zebrafish Bhmt in the modulation of Shh signaling to control β-cell development ((Yang et al., 2011)) further supports additional roles for Bhmts in fish. Data collected for other gilthead seabream *bhmt* genes revealed a different pattern of expression in adult tissues: none were expressed in calcified tissues and expression in soft tissues was distributed between the different forms (*bhmt4* and *bhmt5* mostly in liver, *bhmt6* mostly in kidney and *bhmt7* mostly in intestine). Although this hypothesis would need the support of some functional data, we propose that the broad distribution of *bhmt* expression in soft tissues of non-neoteleostean species was split among the different paralogs in gilthead seabream after gene duplication events that occurred in the neoteleostean lineage. This hypothesis is supported by UniGene Dr.75610 EST profiles of the single zebrafish *bhmt* gene, which transcript is detected in a variety of soft tissues (e.g. liver, intestine and kidney) and calcified tissues (bone, fin and gills). At protein level, Bhmt6 and Bhmt7 were found to be shorter in their C-terminal region than Bhmt4 and Bhmt5 (365 and 370 aa *versus* 401 and 397 aa respectively), indicating that molecular function – possibly in relation to substrate specificity – may have evolved in a way similar to what has been described in mammals for BHMT1 and BHMT2. It has been proposed that *BHMT2* appeared in the ancestral genome of the mammals through gene duplication and that protein got shorter throughout evolution (Ganu et al., 2011), losing residues important for betaine recognition, i.e. Tyr^{77} and amino acids at positions 86-94 and apparently modifying substrate specificity (Szegedi et al., 2008). We propose that the diversification of Bhmts in gilthead seabream, and possibly in other neoteleosteans, was followed by a mechanism of subfunctionalization where sites of protein production/accumulation and substrate specificity would have been distributed among the paralogs. Whether BHMT1 and/or BHMT2 are also expressed in mammalian tissues undergoing peri/endochondral ossification should also be determined to get insights into a possible new role of BHMTs or into the specificity of fish Bhmt3.

In gilthead seabream, *bhmt3* transcript was first detected at blastula stage (512 cells) before the onset of *de novo* transcription, indicating that it may be maternally inherited and

therefore play a role in early embryonic development, most probably in mechanisms not related to ossification since the first calcified structures are only observed after 20 days post-fertilization (DPF) in gilthead seabream (Pinto et al., 2003). Intermediate to higher levels of gene expression were observed from 20 DPH to 48 DPH and at 130 DPH, respectively, when processes such as bone mineralization (around 32 to 35 DPH) (Pinto et al., 2009; Pinto et al., 2003) and bone remodeling (130 DPH) (REF) are actively taking place. Presence of *bhmt3* transcript in structures undergoing chondral ossification in 38 DPH juvenile and in structure fully mineralized and subjected to remodeling in adult fish totally corroborated qPCR results. Although we cannot exclude that *bhmt3* expression at these particular stages could be associated with processes not related to bone formation, we propose – based on its specificity expression in calcified tissues or structures that will get calcified – that *bhmt3* is probably central to mechanisms of ossification during juvenile and adult development. Low expression observed at mouth opening (5 DPH) and during larval-to-juvenile transition (70 DPH) suggests that Bhmt3 function is not required at all stages of fish development. In that aspect, it would be interesting to determine the expression of the four *bhmt* paralogs throughout gilthead seabream development and their specific localization by ISH to get useful insights into their physiological roles and toward a better understanding of the distribution of Bhmt function during development Very little is known about the contribution of mammalian BHMTs to development and the scarce data only refer the presence of BHMT after 10 days of gestation in mouse (Fisher, 2002) and throughout adulthood in several mammalian species (McKeever et al., 1991). EST profiles of mouse (UniGene Mm.329582) and human (UniGene Hs.80756) *BHMT* gene indicate expression at several embryonic stages (cleavage, morula, blastocyst and organogenesis) and at fetal stage (between week 8 and birth), respectively, but a role during mammalian development should be further studied. Altogether, expression data related to *bhmt3* (e.g. regulation upon extracellular matrix mineralization, expression restricted to calcified tissues and particular developmental stages) suggested a tight regulation of gene expression. Although gilthead seabream ABSa15 cell line represented a suitable cellular host (same species and capable of mineralization; Marques et al., 2007) to test the activity of *bhmt3* promoter, the ability of selected transcription factors to regulate gene transcription was tested in the HEK-293 cell line (human origin and not mineralogenic) because of the low transfectability of ABSa15 cells and therefore luciferase activity close to background values. Fish bone-derived cells are known to be hard-to-transfect cells and although some reagents were reported to improved efficiency of transfection (Braga et al., 2006), rates achieved in HEK-293 cells are remarkable and resulted in results with better homogeneity. Castro and co-

workers (Castro et al., 2001) analyzed the basal activity of human *BHMT* promoter activity in human HepG2 cells and only constructs including the TATA box showed transcriptional activity. As for human and mouse *BHMT1* promoters (Neece et al., 2000; Park and Garrow, 1999), a canonical TATA box was identified upstream of the transcription initiation site of gilthead seabream *bhmt3* gene and the promoter construct – carrying the TATA box – exhibited basal activity in both cell lines. Functional analysis of *bhmt3* promoter region identified Ets1 as a positive regulator and both Mef2C and SP1 as negative regulators of *bhmt3* transcription. EST1 is central to osteoblast differentiation (Raouf and Seth, 2000) and was shown to be activated by Erk signaling pathway during chondrogenesis, to promote chondrogenic specification of neural crest cells (Sugiura and Ito, 2010) and to be a positive regulator of gilthead seabream S100-like calcium-binding protein (Fonseca et al., 2011; Rosa et al., 2014) and bone morphogenetic protein 2 (BMP2; (Marques et al., 2015)), two proteins with a critical role in skeletal development, osteogenesis and/or chondrogenesis (Urist, 1965; Crane and Cao, 2014; Fonseca et al., 2011). The positive effect of Ets1 on *bhmt3* transcription is another evidence of the role of Bhmt3 in mechanisms underlying cartilage/bone homeostasis. The inhibition of *bhmt3* transcription by SP1, a regulator of *COL2A1* and *COL10A1* collagen gene expression in differentiating chondrocytes in human (Ghayor et al., 2001; Magee et al., 2005) and of *SOX9* expression during chondrogenesis in human (Piera-Velazquez et al., 2007), could indicate a role in the transitional phase during chondral ossification when the cartilage precursor is replaced with a bony tissue. Although to a lesser extent, Mef2C, a factor that has been implicated in the regulation of bone and cartilage formation (Mackie et al., 2008; Dy et al., 2012) and in the regulation of *Col10a1* expression in mouse hypertrophic chondrocytes (Arnold et al., 2007), also appears to negatively regulate *bhmt3* transcription, further suggesting that expression of gilthead seabream gene may be repressed during chondrocyte differentiation. Despite being a key player of skeleton formation in mouse (Schorle et al., 1996; Zhang et al., 1996) possibly through its negative regulatory action on chondrocyte differentiation (Davies et al., 2002; Huang et al., 2004; Wenke and Bosserhoff, 2010), and despite the presence of a putative binding site in *bhmt3* promoter (also predicted in human and mouse promoters (Neece et al., 2000; Park and Garrow, 1999), AP2 failed to regulate *bhmt3* transcription in our system (predictions were based on mammalian matrices and predicted AP2 binding site may thus not be functional).The mechanisms behind the explosive diversity of *bhmt* genes in the genomes of neoteleosteans (possibly up to 6 genes in some species) are yet unknown and future studies should aim not only at determining the timing of the duplication events that originated this diversity and whether whole genome duplication event that affected teleost fish (Vandepoele

et al., 2004; Meyer and Van de Peer, 2005) could account for some of these genes, but also at understanding the variability in the number of genes that occurred in the species of this lineage, whether this number can be associated with a particular trait and what is the relationship between the different forms. Our current hypothesis on the evolution of vertebrate BHMT family is presented, in a simplified manner, in **Figure 4.11**.

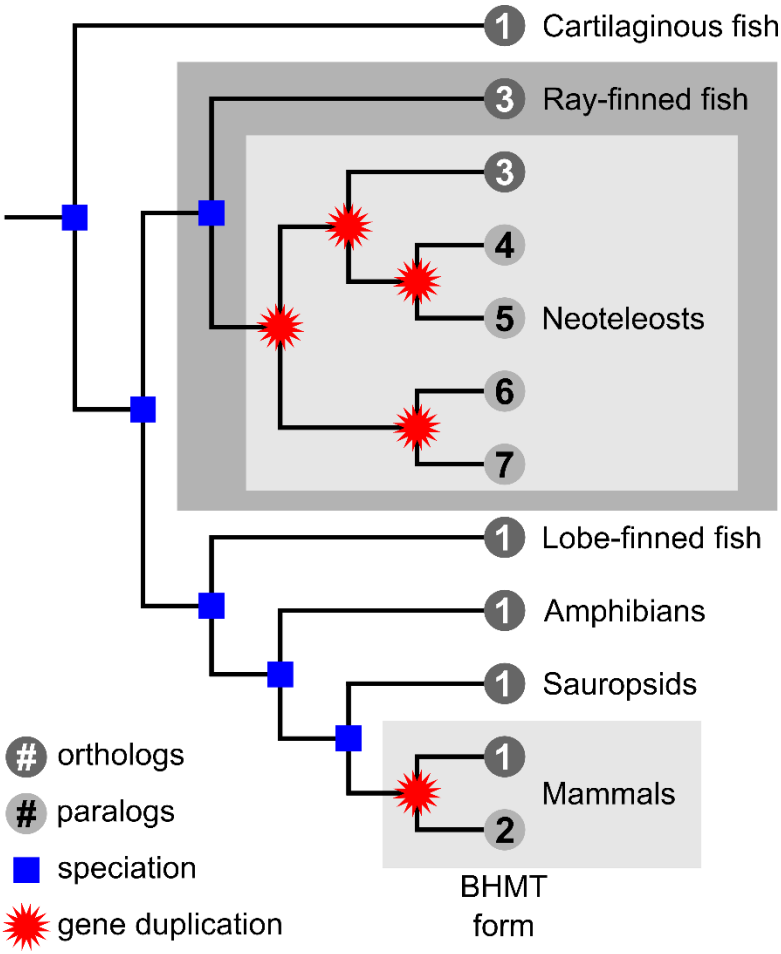


Figure 4.11. Simplified phylogeny of vertebrate *BHMT* genes

Although they are expressed in different tissues, mammalian *BHMT1* and fish *bhmt3* are most probably orthologous, while *BHMT2* and *bhmt4*, *bhmt5*, *bhmt6* and *bhmt7* would be their respective paralogs. It is important to emphasize that none of the *bhmt* genes in neoteleosteans is orthologous to mammalian *BHMT2*. During evolution, sequence of orthologous genes will diverge following speciation (Graur and Li, 1991). Rate of divergence of orthologous genes is usually not uniform and may evolve at a different pace in different taxonomic groups (Hall and Hallgrímsson, 2011). Similarly, sequence of paralogous genes will usually diverged at a faster

rate after gene duplication (Byrne and Wolfe, 2007). In that aspect, we believe that *BHMT* gene suffered a higher rate of divergence in the ray-finned fish than in the other vertebrates but not as a consequence of the duplication events that occurred in this lineage. Indeed, *bhmt3* gene from salmoniformes and cypriniformes (Teleostei species with a single *bhmt* gene and no gene duplication) is clustering with genes of the Neoteleostei in our phylogenetic analysis and not with genes from tetrapod and cartilaginous fish as it would be expected if rate of divergence was similar in ray-finned fish and in other vertebrate species. Similarly, the clustering of the *bhmt* gene from the spotted gar (Holostei species with a single gene and no gene duplication) together with genes of the Teleostei indicates that this increase in divergence rate may have happened early after the branching of the ray-finned fish from the tetrapods. This higher rate of divergence in orthologous genes and the explosive diversity of the family through the occurrence of several paralogous genes may indicate a greater need for Bhmt function in this lineage for reasons that remain to be determined but that could be related to the need to adapt to different substrates as methyl donor and/or to develop novel function. The lineage specific loss of paralogs in Neoteleostei is equally puzzling and the question remains: what environmental condition is driving the duplication and loss of *bhmt* genes in these fishes? The remarkable conservation of the genomic regions surrounding *bhmt6/7* in the Atlantic cod, the three-spined stickleback, the Japanese medaka and the Japanese fugu combined with the absence of *bhmt6/7* gene in the last 2 species suggest that lineage- or species-specific gene loss events are probably locus-specific. Although events are separated by million years of evolution, understanding gene duplication in mammalian species could give useful insights into the mechanisms behind gene duplications in neoteleostean species. Ganu and co-workers (Ganu et al., 2011; Ganu, 2011) suggested that the two mammalian *BHMT* genes have been retained during evolution due to the emergence of an alternative function yet to be identified but probably related to substrate specificity and adaptation to food source availability. Indeed, while betaine – present in wheat, spinach, shellfish and sugar beets – is the preferred substrate of BHMT1 (Sakamoto et al., 2002; Zeisel et al., 2003), S-methylmethionine – present in plants and lower organisms (e.g. *Escherichia coli*) – is used by BHMT2 (Chadwick et al., 2000). Duplication of *BHMT* gene has therefore offered mammals a selective advantage to produce methionine from several food sources and we propose that it may be also the case in Neoteleostei. Another hypothesis could be related to the need of keeping low levels of homocysteine in the plasma (accumulation of Hcy has been associated with a number of diseases in mammals (Look et al., 2000; Dudman et al., 1993; McGregor et al., 2001; Jacobs et al., 1998; Heil et al., 2000; Herrmann et al., 2005; Herrmann et al., 2008; Mizrahi et al.,

2002; Wijekoon et al., 2005) and therefore to have a higher capacity of Hcy transformation with 2 or more proteins. The central role of BHMTs in the regulation of tonicity in mammalian liver and kidney (Craig, 2004) and in the adaptation of fish to osmotic changes (Qian and Song, 2011) are also functions which will need to be explored in future studies to understand Bhmt diversity in neoteleosteans.

In conclusion, the presence of several *bhmt* genes was demonstrated in gilthead seabream and in neotelesotean fish (up to 6 genes). The explosive diversity of *bhmt* genes in this lineage was associated with multiple duplication events that lead to subfunctionalization in gilthead seabream, with expression of each gene restricted to few tissues while tissue expression is covered by a single gene in most of the non-neoteleostean species. We were also able to demonstrate that gilthead seabream *bhmt3* isoform expression was associated with calcified tissues and mechanisms of endochondral ossification, and is transcriptionally regulated by factors know to be involved in bone and cartilage development, supporting thus a new function for BHMT proteins during chondrogenesis and/or osteogenesis, which was never reported before.

4.1.6 Acknowledgments

JR, DT and PV acknowledge the financial support from the Portuguese Foundation for Science and Technology (FCT) through the grants SFRH/BD/47433/2008, SFRH/BPD/45034/2008 and SFRH/BPD/39189/2007, respectively. This work was partly funded by the FCT and the European Commission (ERDF-COMPETE) through PEst-C/MAR/LA0015/2011 project. The help of Natércia Conceição (functional promoter analysis) is gratefully acknowledged.

CHAPTER 5

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

CHAPTER 5. GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES

5.1 Overview

The importance to unveil mechanisms behind bone and cartilage formation regarding signaling pathways, regulatory factors and marker genes are of major importance due to increasingly prevalence of skeletal disorders. In this sense, identifying differentially expressed genes, determining spatiotemporal patterns and levels/sites of gene expression, but also studying the molecular and evolutionary characteristics of novel genes associated with chondrogenesis and osteogenesis will bring basic data that can provides valuable insights into underlying mechanisms of bone formation and mineralization and the basis for future studies on bone therapeutics. The work presented here aimed at the characterization of genes found to be differentially expressed during extracellular matrix mineralization of skeletal cell lines but never associated before with mechanisms involved in skeletogenesis or cartilage/bone homeostasis. The main conclusions and possible future directions of this work are presented in the following sections.

5.2 *S100-like*, a new marker for endochondral ossification in developing gilthead seabream

Calcium ion and S100 calcium-binding proteins are critical players in cell physiology. While the former is a universal, intracellular second messenger that functions as a signal for many cellular processes, the latter translate physiological changes in calcium levels into specific cellular responses (Donato, 2001). S100 proteins have long been associated with chondrocytes and chondrogenesis in mammals (Stefansson et al., 1982; Ushigome et al., 1984; Chano et al., 1995; Li et al., 2002), where they would be involved in cartilage repair and calcification (Mohr et al., 1985). Despite being important for skeletal development, their regulation by transcription factors normally required for chondrocyte differentiation and cartilage development has never been reported before. Although highly similar and with a structure typical of S100 calcium-binding proteins, gilthead seabream S100-like protein – and

other fish orthologs identified *in silico* – could not be associated with any of the twenty S100 protein subfamilies and may represent a novel isoform. This isoform is probably fish-specific since no ortholog has been identified in other vertebrates and may have arisen from a gene duplication event that occurred in the ray-finned fish (Actinopterygii) lineage after branching from tetrapods. Although this remains to be confirmed by functional assays, we proposed that fish S100-like protein is capable of binding calcium ions since it contains key structural features of S100 proteins responsible for calcium binding (Kraemer et al., 2008; Rintala-Dempsey et al., 2008) and crucial for protein function (Fritz and Heizmann, 2004). Expression of *S100-like* gene was found to increase during gilthead seabream development and to be restricted in juvenile and adult fish to cartilaginous structures undergoing peri/endochondral mineralization and to calcifying tissues, respectively. This data together with existing data on S100 proteins (Tuan and Scott, 1977) points toward a function of fish S100-like in calcium storage, possibly to fulfill the large demand for calcium in developing fish and in adult fish actively remodeling bone structures. Whether S100 is specific of teleost species exhibiting acellular bone and could be involved in alternative mechanisms of bone remodeling in the absence of osteocytes (Shahar and Dean, 2013) remains to be proven. Elevated expression in mineralizing tissues and up-regulation of gene expression in pre-chondrocyte and pre-osteoblast cells undergoing ECM mineralization could indicate a possible role during osteogenesis, as already reported for other calcium-binding proteins in mammals (Balmain et al., 2003; Berdal et al., 1996; Faucheux et al., 1998). However an active role in tissue mineralization would imply that the protein is secreted into the extracellular matrix environment. However, no signal peptide were predicted in gilthead seabream S100-like protein sequence, although this may not prevent S100-like secretion (Donato, 2001). In addition to functional data, it would be also important to determine in future studies the localization (i.e. intra or extracellular accumulation) of S100-like protein. To get further insights into the osteogenic role of S100-like protein, the activity of gene promoter was assessed and several DNA elements were shown to bind Ets1a transcription factor and increase gene transcription. Recent data reporting the blockage by Ets1 of chondrocyte differentiation in cardiac neural crest (Gao et al., 2010; Sugiura and Ito, 2010) further suggested the possible role of S100-like in chondrogenesis. The homodimerization of Ets1 and its central role in the transcriptional regulation of *S100-like* gene but also the possible interaction of Ets1 with Cbfa1 should be further demonstrated.

5.3 *sdr-like* a marker of osteoblast differentiation and gastroinstestinal tract

Members of SDR superfamily participate in the metabolism of various molecules – e.g. steroids, prostaglandins, glucocorticoids, lipids, amino acids, carbohydrates, retinoids, aliphatic alcohols and xenobiotics (reviewed in Oppermann et al., 2003). While several members of the SDR superfamily have been associated to pathogenic processes and degenerative defects, including osteoporosis (Tomlinson et al., 2004; Vihko et al., 2005; Mindnich et al., 2004), many remain uncharacterized and could represent potential candidates for monogenic and multifactorial diseases, including bone pathologies, and novel targets for drug development.

The analysis of gilthead seabream Sdr-like protein sequence revealed a high similarity with type C short chain dehydroganeses/reductases while the molecular phylogenetic analysis of closely related sequences (all unannotated) suggested that it probably belongs to a novel subfamily which function remains to be determined. Interestingly, no ortholog of Sdr-like was identified in mammals, an absence probably related to the selective loss of the gene in the genome of the mammalian ancestor after branching from other tetrapods. Expression data – up-regulation of *sdr-like* expression during *in vitro* mineralization, high transcript levels in calcified tissues and localization in pre-osteoblasts lining bone structures – and transcriptional data - regulation of gene promoter activity by bone-specific transcription factors retinoic acid (RA) and Runx2 – clearly suggested a role for gilthead seabream Sdr-like in mechanisms of osteoblast differentiation and mineralization. Although orthologs of Sdr-like exist in the genome of most vertebrate species, indicating its ancestrality, no functional data are available for any of these orthologs. We proposed, based on data available for other SDRs demonstrating the existence of feedback regulatory loops (Cerignoli et al., 2002; Kam et al., 2010)and on the presence of retinol binding domain signatures in Sdr-like, but also based on the expression of *sdr-like* in osteoblast and the central role of retinoic acid in bone mineralization, that Sdr-like could be involved in vitamin A / retinoic acid metabolism to control circulating RA levels to prevent the overstimulation of RAR and consequent effects on the differentiation of primary osteoblast, bone matrix synthesis in mature osteoblasts (reviewed in Imai et al., 2013). Future studies should aim at unveiling enzyme activity and substrate specificity to confirm this hypothesis. Expression data – high transcript levels in stomach and during gastrulation – and *in silico* data – prediction of binding sites for transcription factors associated with embryogenesis, differentiation, proliferation and cell growth in gene promoter – also evidenced

the role of Sdr-like in physiological processes not related to tissue mineralization, possible in gastrointestinal tract homeostasis and during embryonic development.

5.4 *bhmt3*, a fish specific marker for endochondral ossification

Homocysteine metabolism is nowadays the subject of a particular attention since hyper-homocysteinemia has been associated in human with several diseases, including bone disorders such as osteoporosis (Mizrahi et al., 2002; Herrmann et al., 2005; Wijekoon et al., 2005; Heil et al., 2000; Nelen et al., 2000; McGregor et al., 2001). BHMT proteins are responsible for Hcy methylation into methionine (Finkelstein et al., 1972) and getting insights into protein function in a bone context is of major importance to understand the relationship between Hcy metabolism and bone disorders. Gilthead seabream *bhmt3* gene was shown to be differentially expressed during *in vitro* mineralization of osteoblast and chondrocyte extracellular matrix and transcript was detected almost exclusively in calcified tissues and structures undergoing chondral ossification. Expression data collected for other gilthead seabream *bhmt* genes revealed high transcript levels in soft tissues. Tissue distribution of the different seabream paralogs was remarkably specific, i.e. *bhmt3* mostly in calcified tissues, *bhmt4* and *bhmt5* mostly in liver, *bhmt6* mostly in kidney and *bhmt7* mostly in intestine, a distribution covered in other vertebrates (at least for the soft tissues) by a single gene. Although this hypothesis is speculative we propose that the broad distribution of *bhmt* expression in soft tissues of non-neoteleostean species was split among the different paralogs in gilthead seabream after gene duplication events that occurred in the neoteleostean lineage. It is thus possible that duplication of *bhmts* in Neoteleost species was followed by a mechanism of subfunctionalization where sites of protein production/accumulation and substrate specificity would have been distributed among the paralogs. In that aspect, it would be interesting to determine the expression of the four *bhmt* paralogs throughout gilthead seabream development and their specific localization by ISH to get useful insights into their physiological roles and toward a better understanding of the distribution of Bhmt function during development. Whether BHMT1 and/or BHMT2 are also expressed in mammalian tissues undergoing peri/endochondral ossification should also be determined to get insights into a possible new role of BHMTs or into the specificity of fish Bhmt3. However, our data suggests that mammalian *BHMT1* and fish *bhmt3* are most probably

orthologous, while *BHMT2* and *bhmt4*, *bhmt5*, *bhmt6* and *bhmt7* would be their respective paralogs (note that mammalian *BHMT2* is not orthologous to any of the gilthead seabream genes). The central role of *bhmt3* in mechanisms underlying cartilage/bone homeostasis and/or ossification during juvenile and adult development was further demonstrated through the regulation of gene transcription by Ets1 (up-regulation) and Mef2C and SP1 (down-regulation), transcription factors known to regulate chondrocyte differentiation (Ghayor et al., 2001; Magee et al., 2005; Mackie et al., 2008; Dy et al., 2012). This is the first time, to the best of our knowledge, that BHMT function is associated with bone and cartilage formation. Underlying mechanisms remain however unknown and future studies should aim at collecting functional data toward the characterization of enzyme activity and substrate specificity.

5.5 Future perspectives

Unveiling the mechanisms underlying any biological process is an enormous challenge that requires a massive amount of data and all approaches even from remote fields of research, all systems even the most simple, all the tools even the most classical and all ideas even the most speculative, will provide valuable insights that put together will allow the scientific community to tackle this challenge. This work has initiated the idea that members of the SDR, S100 and BHMT protein families may be part of the complex mechanisms regulating skeletal development and bone formation and provided basic expression data in a fish background. This work merely represents a brick in a wall and much remains to be done.

There is a clear need for functional data for Sdr-like and Bhmt3 in order to confirm their enzyme activity – dehydrogenase/reductase and methyltransferase, respectively – and substrate specificity – vitamin A/retinoic acid and betaine/S-methylmethionine, respectively – but also for S100-like to validate calcium binding activity. Approaches based on gain- and loss-of-function have already provided unique data to explore the role of numerous protein in physiological or pathological processes and could be applied to the functional study of *sdr-like*, *S100-like* and *bhmt3*. However this will be difficult to realize in the current model, the gilthead seabream, but could be easily implemented in another fish model, the zebrafish where the manipulation of gene has proven to be somehow easier. Indeed various mutant lines and transgenic lines are already available for numerous zebrafish genes (see ZFIN, The Zebrafish Model Organism Database at http://zfin.org) or can be easily developed. While no zebrafish mutants are available for *sdr-like* and *S100-like*, a mutant (point mutation following ENU treatment) exist for *bhmt1*, the zebrafish ortholog of *bhmt3*, but no phenotype has been

described. Morpholinos silencing *bhmt1* expression during early zebrafish development are available and proved to be useful in determining new functions for Bhmt involving modulation of Shh signaling to control β-cell development (Yang et al., 2011b) and thus could be applied to access Bhmt functions in bone of developing embryo. Although from a different fish species, the availability of the promoter region of the 3 gilthead seabream genes will also certainly facilitate the development of transgenic lines where each promoter will drive the expression of a fluorescent protein (e.g. the green fluorescent protein GFP). Comparative analysis of fluorescence signal patterns in these new lines with those of established transgenic lines for bone and cartilage markers (e.g. runx2, osterix, osteocalcin, sox9, cathepsin K, collagens 1, 2 and 10, etc.) could provide useful insights on the type of cells expressing each gene (e.g. differentiation stage of osteoblast and chondrocytes) and into a role in particular mechanisms (e.g. endochondral versus intramembranous ossification). Silencing gene expression by morpholino in developing embryos could also give useful data on the role of these protein, in particular Sdr-like, during early development. Phenotypical analysis of zebrafish lines modified to silence or stimulate the different function of the 3 genes will certainly be fundamental to better understand their role in skeleton and bone homeostasis.

In a different aspect, validating and improving the data collected on promoter activity and on the regulation of gene expression by bone and cartilage transcription factors is also important. It would be relevant to localize the functional binding elements through site-directed mutagenesis (sdr-like and bhmt3), and further characterize them through electrophoretic mobility shift assay (EMSA) and/or chromatin immunoprecipitation (ChIP). The cooperative effect of Ets1 with Cbfa1 on *S100-like* gene transcription could also be assessed using the same methods. Again, a change from gilthead seabream to zebrafish system should be considered since many of the antibodies against the transcription factors studied here and needed for the ChIP assay have been only validated for zebrafish. The conservation of particular TF binding sites throughout vertebrate evolution (or ray-finned fish evolution for S100-like and bhmt3) could be evaluated *in silico* after collecting the promoter sequences of representative organisms from public sequence databases and analyze then for conserved regions/binding sites. The fact that Ets1 is a transcription factor stimulating the expression of the 3 genes is striking and could indicate that Ets1 is as a master regulator of bone marker genes. This hypothesis should be further investigated, first *in silico* by evaluating the presence of binding site in other bone-related genes, then *in vitro* through transient co-transfection of gene promoter constructs with expression vectors carrying zebrafish *ets1a* cDNA.

Finally, a possible relation between these genes and bone cellularity should be evaluated by collecting data in a teleost fish with acellular bone. Again, zebrafish represents a suitable model – it has osteocytic bones – and orthologs of the 3 seabream genes are already known.

5.6 Concluding Remarks

Although much remains to be done, novel and valuable data have been collected within the scope of this work toward a better understanding of S100, SDR and BHMT protein functions and regulation but also of the evolutionary relationship within respective gene families. This new data has been almost entirely collected using fish tools and it is worth to note that tools (pcDNA3/pCMX vectors driving gene expression through the human CMV promoter and HEK-293 cell line as an host for fish sequence-based vectors) developed for mammalian systems were functional in fish systems or when used with fish tools, supporting previous data showing a conservation of cellular machinery and regulatory mechanism between mammals and fish, further validating the use of fish as an alternative model to investigate molecular mechanisms of bone metabolism. Unfortunately, we were not able to find mammalian orthologs for S100-like and Sdr-like and on the basis of the existing data we cannot extrapolate and further access and confirm new functions for these proteins in mammals. Bhmt3, on the contrary, appears to be the orthologous to BHMT1 in mammals, and a possible role of mammalian proteins (i.e. BHMT1 and BHMT2) in cartilage/bone homeostasis should be further evaluated. Despite the fact of no orthologs were found for *S100-like* and *sdr-like*, they represent good marker genes to study osteoblast differentiation (*sdr-like*) and chondral ossification (*S100-like* and *bhmt3*) in fish systems.

REFERENCES

REFERENCES

Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* 38:W7–13.

Akiyama H (2008) Control of chondrogenesis by the transcription factor Sox9. *Mod. Rheumatol.* 18:213–219.

Apschner A, Schulte-Merker S, Witten PE (2011) Not all bones are created equal – using zebrafish and other teleost species in osteogenesis research. In *The Zebrafish: Disease Models and Chemical Screens*. pp. 239–256.

Arnold MA, Kim Y, Czubryt MP, Phan D, McAnally J, Qi X, et al. (2007) MEF2C transcription factor controls chondrocyte hypertrophy and bone development. *Dev. Cell* 12:377–89.

Babayeva ND, Baranovskaya OI, Tahirov TH (2012) Structural basis of Ets1 cooperative binding to widely separated sites on promoter DNA. *PLoS One* 7:e33698.

Balmain N, Moutahir F, Heizmann CW, Lieberherr M (2003) Immunolocalization of S100A2 calcium-binding protein in cartilage and bone cells. *Cell. Mol. Biol. (Noisy-le-grand).* 49:485–6.

Beck G, Zerler B, Moran E (2001) Gene array analysis of osteoblast differentiation. *Cell Growth Differ.* 12:61–83.

Behrens AN, Ren Y, Ferdous A, Garry DJ, Martin CM (2012) Nkx2-5 regulates tdgf1 (cripto) early during cardiac development. *J. Clin. Exp. Cardiolog.* 11:1–4.

Belloni E, Muenke M, Roessler E, Traverso G, Siegel-Bartelt J, Frumkin A, et al. (1996) Identification of Sonic hedgehog as a candidate gene responsible for holoprosencephaly. *Nat. Genet.* 14:353–6.

Bendtsen J, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340:783–795.

Benitez-Santana T, Saleh R, Betancor MB, Mesa A, Hernindez-Cruz CM, Izquierdo M (2013) Effects of combined phospholipids and selenium dietary contents on patterns of bone formation in the axial skeleton of Sparus aurata. *Commun. Agric. Appl. Biol. Sci.* 78:29–32.

Benjamin M (1989) Hyaline-cell cartilage (chondroid) in the heads of teleosts. *Anat. Embryol. (Berl).* 179:285–303.

Benjamin M (1990) The cranial cartilages of teleosts and their classification. *J. Anat.* 169:153–172.

Benjamin M, Evans E (1990) Fibrocartilage. *J. Anat.* 171:1–15.

Benjamin M, Ralphs J (2004) Biology of fibrocartilage cells. *Int. Rev. Cytol.* 233:1–45.

Benjamin M, Ralphs J, Eberewariye O (1992) Cartilage and related tissues in the trunk and fins of teleosts. *J. Anat.* 181:113–118.

Benjamin M, Ralphs JR, Eberewariye OS (1992) Cartilage and related tissues in the trunk and fins of teleosts. *J. Anat.* 181:113–118.

Berdal A, Hotton D, Saffar J, Thomasset M, Nanci A (1996) Calbindin-D9k and calbindin-D28k expression in rat mineralized tissues in vivo. *J. Bone Miner. Res.* 11:768–779.

GENERAL REFERENCES

Bhattacharya S, Bunick C, Chazin W (2004) Target selectivity in EF-hand calcium binding proteins. *Biochim. Biophys. Acta* 1742:69–79.

Bhattacharya S, Bunick CG, Chazin WJ (2004) Target selectivity in EF-hand calcium binding proteins. *Biochim. Biophys. Acta* 1742:69–79.

Blair HC, Zaidi M, Schlesinger PH (2002) Mechanisms balancing skeletal matrix synthesis and degradation. *Biochem. J.* 364:329–41.

Bobe J, Goetz F (2000) A S100 homologue mRNA isolated by differential display PCR is down-regulated in the brook trout (Salvelinus fontinalis) post-ovulatory ovary. *Gene* 257:187–194.

Boglino A, Wishkerman A, Darias MJ, Andree KB, de la Iglesia P, Estévez A, et al. (2013) High dietary arachidonic acid levels affect the process of eye migration and head shape in pseudoalbino Senegalese sole Solea senegalensis early juveniles. *J. Fish Biol.* 83:1302–20.

Bose N, Greenspan P, Momany C (2002) Expression of recombinant human betaine: homocysteine S-methyltransferase for x-ray crystallographic studies and further characterization of interaction with S-adenosylmethionine. *Protein Expr. Purif.* 25:73–80.

Braga D, Laizé V, Tiago D, Cancela M (2006) Enhanced DNA transfer into fish bone cells using polyethylenimine. *Mol. Biotechnol.* 34:51–54.

Bryne JC, Valen E, Tang M-HE, Marstrand T, Winther O, da Piedade I, et al. (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res.* 36:D102–6.

Byrne KP, Wolfe KH (2007) Consistent patterns of rate asymmetry and gene loss indicate widespread neofunctionalization of yeast genes after whole-genome duplication. *Genetics* 175:1341–50.

Caetano-Lopes J, Canhão H, Fonseca J (2007) Osteoblasts and bone formation. *Acta Reum. Port.* 1:103–110.

Cancela M, Williamson M, Price P (1995) Amino acid sequence of bone Gla protein from the african clawed toad Xenopus laevis and the fish Sparus aurata. *Int. J. Pept. Protein Res.* 46.

Cancela ML, Price PA (1992) Retinoic acid induces matrix Gla protein gene expression in human cells. *Endocrinology* 130:102–8.

Cao Z, Tanguay R, McKenzie D, Peterson R, Aiken J (2003) Identification of a putative calcium-binding protein as a dioxin-responsive gene in zebrafish and rainbow trout. *Aquat. Toxicol.* 63:271–282.

Cardeira J, Mendes AC, Pousão-Ferreira P, Cancela ML, Gavaia PJ (2015) Micro-anatomical characterization of vertebral curvatures in Senegalese sole Solea senegalensis. *J. Fish Biol.* 86:1796–810.

Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff a, et al. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 21:2933–42.

Castro C, Breksa AP, Salisbury EM, Garrow TA (2001) Betaine-homocysteine S-methyl-transferase (BHMT) transcription is inhibited by S-adenosylmethionine (AdoMet). In S Milstien, G Kapatos, RA Levine, and B Shane, editors *Chemistry and Biology of Pteridines and Folates*. Boston, MA: Springer US, pp. 549–556.

Cerignoli F, Guo X, Cardinali B, Rinaldi C, Casaletto J, Frati L, et al. (2002) retSDR1, a short-chain retinol dehydrogenase/reductase, is retinoic acid-inducible and frequently deleted in human neuroblastoma cell lines. *Cancer Res.* 62:1196–204.

Chadjichristos C, Ghayor C, Kypriotou M, Martin G, Renard E, Ala-Kokko L, et al. (2003) Sp1 and Sp3 transcription factors mediate interleukin-1 beta down-regulation of human type II collagen gene expression in articular chondrocytes. *J. Biol. Chem.* 278:39762–39772.

Chadwick LH, McCandless SE, Silverman GL, Schwartz S, Westaway D, Nadeau JH (2000) Betaine-homocysteine methyltransferase-2: cDNA cloning, gene sequence, physical mapping, and expression of the human and mouse genes. *Genomics* 70:66–73.

Chano T, Ishizawa M, Matsumoto K, Morimoto S, Hukuda S, Okabe H (1995) The identity of proliferating cells in bone tumors with cartilaginous components: Evaluation by double-immunohistochemical staining using proliferating cell nuclear antigen and S-100 protein. *Eur. J. Histochem.* 39:21–30.

Chapman K, Holmes M, Seckl J (2013) 11β-hydroxysteroid dehydrogenases: intracellular gate-keepers of tissue glucocorticoid action. *Physiol. Rev.* 93:1139–206.

Chico T, Ingham P, Crossman D (2008) Modeling cardiovascular disease in the zebrafish. *Trends Cardiovasc. Med.* 18:150–155.

Choi J, Pratap J, Javed A, Zaidi S, Xing L, Balint E, et al. (2001) Subnuclear targeting of Runx/ Cbfa/AML factors is essential for tissue-specific differentiation during embryonic development. *PNAS, Proc. Natl. Acad. Sci.* 98:8650– 8655.

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156–9.

Chun J-S, Oh H, Yang S, Park M (2008) Wnt signaling in cartilage development and degeneration. *BMB Rep.* 41:485–494.

Clarke B (2008) Normal bone anatomy and physiology. *Clin. J. Am. Soc. Nephrol.* 3:131–139.

Clément A, Wiweger M, von der Hardt S, Rusch M, Selleck S, Chien C-B, et al. (2008) Regulation of zebrafish skeletogenesis by ext2/dackel and papst1/ pinscher. *PLoS Genet.* 4, 7.

Cohen L, Dean M, Shipov A, Atkins A, Monsonego-Ornan E, Shahar R (2012) Comparison of structural, architectural and mechanical aspects of cellular and acellular bone in two teleost fish. *J. Exp. Biol.* 215:1983–93.

Cole A, Hall B (2004)(a) Cartilage is a metazoan tissue; integrating data from nonvertebrate sources. *Acta Zool.* 85:69–80.

Cole A, Hall B (2004)(b) The nature and significance of invertebrate cartilages revisited: distribution and histology of cartilage and cartilage-like tissues within the Metazoa. *Zool.* 107:261–273.

Conceição N, Laizé V, Simões B, Pombinho AR, Cancela ML (2008) Retinoic acid is a negative regulator of matrix Gla protein gene expression in teleost fish Sparus aurata. *Biochim. Biophys. Acta* 1779:28–39.

Cooper MS, Walker E a, Bland R, Fraser WD, Hewison M, Stewart PM (2000) Expression and functional consequences of 11beta-hydroxysteroid dehydrogenase activity in human bone. *Bone* 27:375–81.

Craig SA (2004) Betaine in human nutrition. *Am J Clin Nutr* 80:539–549.

GENERAL REFERENCES

Crane JL, Cao X (2014) Bone marrow mesenchymal stem cells and TGF-β signaling in bone remodeling. *J. Clin. Invest.* 124:466–72.

Crombrugghe B de, Lefebvre V, Nakashima K (2001) Regulatory mechanisms in the pathways of cartilage and bone formation. *Curr. Opin. Cell Biol.* 13:721–727.

Dallas S, Prideaux M, Bonewald L (2013) The osteocyte: an endocrine cell ... and more. *Endocr. Rev.* 34:658–690.

Davies SR, Sakano S, Zhu Y, Sandell LJ (2002) Distribution of the transcription factors Sox9, AP-2, and [Delta]EF1 in adult murine articular and meniscal cartilage and growth plate. *J. Histochem. Cytochem.* 50:1059–1065.

DeLaurier A, Nakamura Y, Braasch I, Khanna V, Kato H, Wakitani S (2012) Histone deacetylase-4 is required during early cranial neural crest development for generation of the zebrafish palatal skeleton. *BMC Dev. Biol.* 12, 16.

Doi M, Nagano A, Nakamura Y (2002) Genome-wide screening by cDNA microarray of genes associated with matrix mineralization by human mesenchymal stem cells in vitro. *Biochem. Biophys. Res. Commun.* 290:381–90.

Donato R (2003) Intracellular and extracellular roles of S100 proteins. *Microsc. Res. Tech.* 60:540–551.

Donato R (2001) S100: a multigenic family of calcium-modulated proteins of the EF-hand type with intracellular and extracellular functional roles. *Int. J. Biochem. Cell Biol.* 33:637–68.

Dong W, Hinton D, Kullman S (2012) TCDD disrupts hypural skeletogenesis during medaka embryonic development. *Toxicol. Sci.* 125:91–104.

Drissi H, Luc Q, Shakoori R, Chuva De Sousa Lopes S, Choi JY, Terry A, et al. (2000) Transcriptional autoregulation of the bone related CBFA1/RUNX2 gene. *J. Cell. Physiol.* 184:341–50.

Ducy P, Karsenty G (1995) Two distinct osteoblast-specific cis-acting elements control expression of a mouse osteocalcin gene. *Mol. Cell. Biol.* 15:1858–69.

Dudman NP, Wilcken DE, Wang J, Lynch JF, Macey D, Lundberg P (1993) Disordered methionine/homocysteine metabolism in premature vascular disease. Its occurrence, cofactor therapy, and enzymology. *Arterioscler. Thromb.* 13:1253–60.

Dy P, Wang W, Bhattaram P, Wang Q, Wang L, Ballock RT, et al. (2012) Sox9 directs hypertrophic maturation and blocks osteoblast differentiation of growth plate chondrocytes. *Dev. Cell* 22:597–609.

Eijken M, Hewison M, Cooper MS, de Jong FH, Chiba H, Stewart PM, et al. (2005) 11beta-Hydroxysteroid dehydrogenase expression and glucocorticoid synthesis are directed by a molecular switch during osteoblast differentiation. *Mol. Endocrinol.* 19:621–31.

Evans JC, Huddler DP, Jiracek J, Castro C, Millian NS, Garrow TA, et al. (2002) Betaine-homocysteine methyltransferase: Zinc in a distorted barrel. *Structure* 10:1159–1171.

Eyre D, Wu J (1983) Collage of fibrocartilage: a distinctive molecular phenotype in bovine meniscus. *FEBS Lett.* 158:265–270.

Fan H-Y, Liu Z, Johnson PF, Richards JS (2011) CCAAT/enhancer-binding proteins (C/EBP)-α and -β are essential for ovulation, luteinization, and the expression of key target genes. *Mol. Endocrinol.* 25:253–68.

Faucheux C, Bareille R, Amedee J (1998) Synthesis of calbindin-D28K during mineralization in human bone marrow stromal cells. *Biochem. J.* 333:817–823.

Faustino M, Power DM (1998) Development of osteological structures in the sea bream: vertebral column and caudal fin complex. *J. Fish Biol.* 52:11–22.

Faustino M, Power DM (1999) Development of the pectoral, pelvic, dorsal and anal fins in cultured sea bream. *J. Fish Biol.* 54:1094–1110.

Faustino M, Power DM (2001) Osteologic development of the viscerocranial skeleton in sea bream: alternative ossification strategies in teleost fish. *J. Fish Biol.* 58:537–572.

Feitsma H, Cuppen E (2008) Zebrafish as a cancer model. *Mol. cancer Res.* 6:685–694.

Fernández I, Tiago D, Laizé V, Cancela M, Gisbert E (2014) Retinoic acid differentially affects in vitro proliferation, differentiation and mineralization of two fish bone-derived cell lines: different gene expression of nuclear receptors and ECM proteins. *J. Steroid Biochem. Mol. Biol.* 140:34–43.

Finkelstein JD, Harris BJ, Kyle WE (1972) Methionine metabolism in mammals: Kinetic study of betaine-homocysteine methyltransferase. *Arch. Biochem. Biophys.* 153:320–324.

Finkelstein JD, Martin JJ (1984) Methionine metabolism in mammals. Distribution of homocysteine between competing pathways. *J. Biol. Chem.* 259:9508–9513.

Finkelstein JD, Martin JJ, Harris BJ (1986) Effect of Dietary Cystine on Methionine Metabolism in Rat Liver. *J. Nutr.* 116:958–990.

Fisher MC (2002) Perturbations in choline metabolism cause neural tube defects in mouse embryos in vitro. *FASEB J.* 16:619–621.

Fisher S, Jagadeeswaran P, Halpern ME (2003) Radiographic analysis of zebrafish skeletal defects. *Dev. Biol.* 264:64–76.

Fonseca VG, Laizé V, Valente MS, Cancela ML (2007) Identification of an osteopontin-like protein in fish associated with mineral formation. *FEBS J.* 274:4428–39.

Fonseca VG, Rosa J, Laizé V, Gavaia PJ, Cancela ML (2011) Identification of a new cartilage-specific S100-like protein up-regulated during endo/perichondral mineralization in gilthead seabream. *Gene Expr. Patterns* 11:448–55.

Foster J, Kumar S, Ganatra M, Kamal I, Ware J, Ingram J, et al. (2004) Construction of bacterial artificial chromosome libraries from the parasitic nematode Brugia malayi and physical mapping of the genome of its Wolbachia endosymbiont. *Int. J. Parasitol.* 34:733–746.

Foster PG (2004) Modeling compositional heterogeneity. *Syst. Biol.* 53:485–495.

Fritz G, Heizmann C (2004) 3D structures of the calcium and zinc binding S100 proteins. In W Bode, A Messerschmidt, and M Cygler, editors *Handbook of metalloproteins*. Chichester: John Wiley & Sons, Ltd, pp. 529–540.

Fukada K, Shibata S, Suzuki S, Ohya K, Kuroda T (1999) In situ hybridisation study of type I, II, X collagens and aggrecan mRNas in the developing condylar cartilage of fetal mouse mandible. *J. Anat.* 195:321–329.

Fukuoka H, Shibata S, Suda N, Yamashita Y, Komori T (2007) Bone morphogenetic protein rescues the lack of secondary cartilage in Runx2-deficient mice. *J. Anat.* 211:8–15.

GENERAL REFERENCES

Ganu R (2011) Developmental and evolutionary analysis of betaine homocysteine methyltransferase gene. PhD thesis from the University of Illinois at Urbana-Champaign.

Ganu RS, Garrow TA, Sodhi M, Rund LA, Lawrence B (2012) NIH Public Access. 473:133–138.

Ganu RS, Garrow TA, Sodhi M, Rund LA, Schook LB (2011) Molecular characterization and analysis of the porcine betaine homocysteine methyltransferase and betaine homocysteine methyltransferase -2 genes. *Gene* 473:133–138.

Gao Z, Kim GH, Mackinnon AC, Flagg AE, Bassett B, Earley JU, et al. (2010) Ets1 is required for proper migration and differentiation of the cardiac neural crest. *Development* 137:1543–51.

Garrow TA (1996) Purification, Kinetic Properties, and cDNA Cloning of Mammalian Betaine-Homocysteine Methyltransferase. *J. Biol. Chem.* 271:22831–22838.

Gathercole LL, Lavery GG, Morgan S a, Cooper MS, Sinclair AJ, Tomlinson JW, et al. (2013) 11β-Hydroxysteroid dehydrogenase 1: translational and therapeutic aspects. *Endocr. Rev.* 34:525–55.

Gaur T, Hussain S, Mudhasani R, Parulkar I, Colby JL, Frederick D, et al. (2010) Dicer inactivation in osteoprogenitor cells compromises fetal survival and bone formation, while excision in differentiated osteoblasts increases bone mass in the adult mouse. *Dev. Biol.* 340:10–21.

Gavaia PJ, Sarasquete C, Cancela ML (2000) Detection of mineralized structures in early stages of development of marine Teleostei using a modified alcian blue-alizarin red double staining technique for bone and cartilage. *Biotech. Histochem.* 75:79–84.

Gebhardt C, Németh J, Angel P, Hess J (2006) S100A8 and S100A9 in inflammation and cancer. *Biochem. Pharmacol.* 72:1622–1631.

Genten F, Terwinghe E, Danguy A (2009) Atlas of Fish Histology. In Science Publishers, p. 223.

Ghayor C, Chadjichristos C, Herrouin JF, Ala-Kokko L, Suske G, Pujol JP, et al. (2001) Sp3 represses the Sp1-mediated transactivation of the human COL2A1 gene in primary and de-differentiated chondrocytes. *J. Biol. Chem.* 276:36881–95.

Gil Martens L, Lock EJ, Fjelldal PG, Wargelius A, Araujo P, Torstensen BE, et al. (2010) Dietary fatty acids and inflammation in the vertebral column of Atlantic salmon, Salmo salar L., smolts: a possible link to spinal deformities. *J. Fish Dis.* 33:957–72.

Goldring MB, Tsuchimochi K, Ijiri K (2006) The control of chondrogenesis. *J. Cell. Biochem.* 97:33–44.

González B, Campillo N, Garrido F, Gasset M, Sanz-Aparicio J, Pajares MA (2003) Active-site-mutagenesis study of rat liver betaine-homocysteine S-methyltransferase. *Biochem. J.* 370:945–952.

González B, Pajares M a, Martínez-Ripoll M, Blundell TL, Sanz-Aparicio J (2004) Crystal structure of rat liver betaine homocysteine s-methyltransferase reveals new oligomerization features and conformational changes upon substrate binding. *J. Mol. Biol.* 338:771–82.

Graur D, Li W-H (1991) Fundamentals of Molecular Evolution.

Graves BJ, Petersen JM (1998) Specificity within the ets family of transcription factors. *Adv. Cancer Res.* 75:1–55.

Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, et al. (1996) The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development* 123:1–36.

Hall B (2005) Bone and cartilage: development and evolutionary skeletal biology.

Hall BK, Hallgrímsson B (2011) Strickberger's Evolution. Jones & Bartlett Publishers.

Hammond C, Moro E (2012) Using transgenic reporters to visualize bone and cartilage signaling during development in vivo. *Front. Endocrinol.* 3.

Heil SG, Lievers KJ, Boers GH, Verhoef P, den Heijer M, Trijbels FJ, et al. (2000) Betaine-homocysteine methyltransferase (BHMT): genomic sequencing and relevance to hyperhomocysteinemia and vascular disease in humans. *Mol. Genet. Metab.* 71:511–9.

Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel a E, Kel O V, et al. (1998) Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Res.* 26:362–7.

Heizmann C, Fritz G, Schafer B (2002) S100 proteins: stucture, functions and pathology. *Front. Biosci.* 7:d1356–1368.

Heizmann C, Schafer B, Fritz G (2003) The family of S100 cell signaling proteins. In R Bradshaw and E Dennis, editors *Handbook of cellular signalling*. USA: Elsevier Science, pp. 87–93.

Herrmann M, Umanskaya N, Wildemann B, Colaianni G, Widmann T, Zallone A, et al. (2008) Stimulation of osteoblast activity by homocysteine. *J. Cell. Mol. Med.* 12:1205–1210.

Herrmann M, Widmann T, Herrmann W (2005) Homocysteine--a newly recognised risk factor for osteoporosis. *Clin. Chem. Lab. Med.* 43:1111–7.

Hess J, Porte D, Munz C, Angel P (2001) AP-1 and Cbfa/runt physically interact and regulate parathyroid hormone-dependent MMP13 expression in osteoblasts through a new osteoblast-specific element 2/AP-1 composite element. *J. Biol. Chem.* 276:20029–20038.

Hill TP, Später D, Taketo MM, Birchmeier W, Hartmann C (2005) Canonical Wnt/beta-catenin signaling prevents osteoblasts from differentiating into chondrocytes. *Dev. Cell* 8:727–738.

Hofmann M, Drury S, Fu C, Qu W, Taguchi A, Lu Y, et al. (1999) RAGE mediates a novel proinflammatory axis: a central cell surface receptor for S100/calgranulin polypeptides. *Cell* 97:889–901.

Hsiao C, Ekker M, Tsai H (2003) Skin-specific expression of ictacalcin, a homolog of the S100 genes, during zebrafish embryogenesis. *Dev. Dyn.* 228:745–750.

Huang Z, Xu H, Sandell L (2004) Negative regulation of chondrocyte differentiation by transcription factor AP-2alpha. *J. Bone Miner. Res.* 19:245–55.

Ikura M (1996) Calcium binding and conformational response in EF-hand proteins. *Trends Biochem. Sci.* 21:14–17.

Imai Y, Youn M-Y, Inoue K, Takada I, Kouzmenko A, Kato S (2013) Nuclear receptors in bone physiology and diseases. *Physiol. Rev.* 93:481–523.

Infante JLZ, Cahu CL (2001) Ontogeny of the gastrointestinal tract of marine fish. :477–487.

Ingham P (2009) The power of the zebrafish for disease analysis. *Hum. Mol. Genet.* 18:R107–R112.

GENERAL REFERENCES

Jacobs RL, House JD, Brosnan ME, Brosnan JT (1998) Effects of streptozotocin-induced diabetes and of insulin treatment on homocysteine metabolism in the rat. *Diabetes* 47:1967–1970.

Jaillon O, Aury J, Brunet F, Petit J, Stange-Thomann N, Mauceli E, et al. (2004) Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* 31:946–957.

Javidan Y, Schilling T (2004) Development of cartilage and bone. *Methods Cell Biol.* 76:415–36.

Joernvall H, Persson B, Krook M, Atrian S, Gonzalez-Duarte R, Jeffery J, et al. (1995) Short-chain dehydrogenases/reductases (SDR). *Biochemistry* 34:6003–6013.

Jones FS, Kioussi C, Copertino DW, Kallunki P, Holst BD, Edelman GM (1997) Barx2, a new homeobox gene of the Bar class, is expressed in neural and craniofacial structures during development. *Proc. Natl. Acad. Sci. U. S. A.* 94:2632–7.

De Jong DS, Vaes BLT, Dechering KJ, Feijen A, Hendriks JMA, Wehrens R, et al. (2004) Identification of novel regulators associated with early-phase osteoblast differentiation. *J. Bone Miner. Res.* 19:947–58.

Jong R, van der Heijden J, Meijlink F (1993) DNA-binding specificity of the S8 homeodomain. *Nucleic Acids Res.* 21:4711–4720.

Julenius K, Molgaard A, Gupta R, Brunak S (2005) Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 15:153–164.

Kallberg Y, Oppermann U, Persson B (2010) Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. *FEBS J.* 277:2375–86.

Kam R, Chen Y, Chan S, Chan W, Dawid I, Zhao H (2010) Developmental expression of Xenopus short-chain dehydrogenase/reductase 3. *Int. J. Dev. Biol.* 54:1355–1360.

Kardong K V. (1998) Vertebrates: Comparative Anatomy, Function, Evolution.

Karg H, Burger E, Lyaruu D, Bronckers A, Woltgens J (1997) Spatiotemporal expression of the homeobox gene S8 during mouse tooth development. *Arch. Oral Biol.* 42:625–631.

Karreth F, Hoebertz A, Scheuch H, Eferl R, Wagner EF (2004) The AP1 transcription factor Fra2 is required for efficient cartilage development. *Development* 131:5717–5725.

Karsenty G (2001) Minireview: transcriptional control of osteoblast differentiation. *Endocrinology* 142:2731–2733.

Karsenty G (2003) The complexities of skeletal biology. *Nature* 423:316–318.

Karsenty G, Oury F (2014) Regulation of male fertility by the bone-derived hormone osteocalcin. *Mol. Cell. Endocrinol.* 382:521–6.

Karsenty G, Wagner EF (2002) Reaching a genetic and molecular understanding of skeletal development. *Dev. Cell* 2:389–406.

Keller B, Volkmann A, Wilckens T, Moeller G, Adamski J (2006) Bioinformatic identification and characterization of new members of short-chain dehydrogenase/reductase superfamily. *Mol. Cell. Endocrinol.* 248:56–60.

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203:253–310.

Knopf F, Hammond C, Chekuru A, Kurth T, Hans S, Weber C, et al. (2011) Bone regenerates via dedifferentiation of osteoblasts in the zebrafish fin. *Dev. Cell* 20:713–724.

Kobayashi T, Lu J, Cobb BS, Rodda SJ, McMahon AP, Schipani E, et al. (2008) Dicer-dependent pathways regulate chondrocyte proliferation and differentiation. *PNAS, Proc. Natl. Acad. Sci.* 105:1949–1954.

Kola I, Brookes S, Green a R, Garber R, Tymms M, Papas TS, et al. (1993) The Ets1 transcription factor is widely expressed during murine embryo development and is associated with mesodermal cells involved in morphogenetic processes such as organ formation. *Proc. Natl. Acad. Sci. U. S. A.* 90:7588–92.

Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, et al. (1997) Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* 89:755–764.

Kraemer A, Saraiva L, Korsching S (2008) Structural and functional diversification in the teleost S100 family of calcium-binding proteins. *BMC Evol. Biol.* 8:48.

Kranenbarg S, van Cleynenbreugel T, Schipper H, van Leeuwen J (2005) Adaptive bone formation in acellular vertebrae of sea bass (Dicentrarchus labrax L.). *J. Exp. Biol.* 208:3493–502.

Kriebitzsch C, Verlinden L, Eelen G, van Schoor NM, Swart K, Lips P, et al. (2011) 1,25-dihydroxyvitamin D3 influences cellular homocysteine levels in murine preosteoblastic MC3T3-E1 cells by direct regulation of cystathionine β-synthase. *J. Bone Miner. Res.* 26:2991–3000.

Kronenberg HM (2003) Developmental regulation of the growth plate. *Nature* 423:332–6.

Laizé V, Gavaia P, Viegas M, Caria J, Luis N (2014) Osteotoxicity of 3- methylcholanthrene in fish. *Aquat. Procedia* in press.

Laizé V, Gavaia PJ, Cancela ML (2015) Fish: a suitable system to model human bone disorders and discover drugs with osteogenic or osteotoxic activities. *Drug Discov. Today Dis. Model.* in press:doi:10.1016/j.ddmod.2014.08.001.

Laizé V, Pombinho A, Cancela M (2005) Characterization of Sparus aurata osteonectin cDNA and in silico analysis of protein conserved features: Evidence for more than one osteonectin in Salmonidae. *Biochimie* 87:411–420.

Lamber EP, Vanhille L, Textor LC, Kachalova GS, Sieweke MH, Wilmanns M (2008) Regulation of the transcription factor Ets-1 by DNA-mediated homo-dimerization. *EMBO J.* 27:2006–17.

Langheinrich U (2003) Zebrafish: a new model on the pharmaceutical catwalk. *Bioessays* 25:904–12.

Lartillot N, Rodrigue N, Stubbs D, Richer J (2013) PhyloBayes-MPI A Bayesian software for phylogenetic reconstruction using mixture models MPI version. :1–21.

Leclerc E, Fritz G, Vetter S, Heizmann C (2009) Binding of S100 proteins to RAGE: An update. *Biochim. Biophys. Acta - Mol. Cell Res.* 1793:993–1007.

Leśniak W (2011) Epigenetic regulation of S100 protein expression. *Clin. Epigenetics* 2:77–83.

GENERAL REFERENCES

Leśniak W, Swart GW, Bloemers HP, Kuźnicki J (2000) Regulation of cell specific expression of calcyclin (S100A6) in nerve cells and other tissues. *Acta Neurobiol. Exp. (Wars).* 60:569–75.

Li CL, Martinez V, He B, Lombet A, Perbal B (2002) A role for CCN3 (NOV) in calcium signalling. *Mol. Pathol.* 55:250–261.

Lieschke GJ, Currie PD (2007) Animal models of human disease: zebrafish swim into view. *Nat. Rev. Genet.* 8:353–67.

Lind T, Sundqvist A, Hu L, Pejler G, Andersson G, Jacobson A (2013) Vitamin A Is a Negative Regulator of Osteoblast Mineralization. *PLoS One* 8:e82388.

Look MP, Riezler R, Reichel C, Brensing KA, Rockstroh JK, Stabler SP, et al. (2000) Is the increase in serum cystathionine levels in patients with liver cirrhosis a consequence of impaired homocysteine transsulfuration at the level of gamma-cystathionase? *Scand. J. Gastroenterol.* 35:866–72.

Lu SC, Alvarez L, Huang ZZ, Chen L, An W, Corrales FJ, et al. (2001) Methionine adenosyltransferase 1A knockout mice are predisposed to liver injury and exhibit increased expression of genes involved in proliferation. *Proc. Natl. Acad. Sci. U. S. A.* 98:5560–5.

Mackie EJ, Ahmed YA, Tatarczuch L, Chen K-S, Mirams M (2008) Endochondral ossification: how cartilage is converted into bone in the developing skeleton. *Int. J. Biochem. Cell Biol.* 40:46–62.

Maden M, Sonneveld E, van der Saag PT, Gale E (1998) The distribution of endogenous retinoic acid in the chick embryo: implications for developmental mechanisms. *Development* 125:4133–44.

Magee C, Nurminskaya M, Faverman L, Galera P, Linsenmayer TF (2005) SP3/SP1 transcription activity regulates specific expression of collagen type X in hypertrophic chondrocytes. *J. Biol. Chem.* 280:25331–8.

Marenholz I, Heizmann CW, Fritz G (2004) S100 proteins in mouse and man: From evolution to function and pathology (including an update of the nomenclature). *Biochem. Biophys. Res. Commun.* 322:1111–1122.

Marenholz I, Lovering R, Heizmann C (2006) An update of the S100 nomenclature. *Biochim. Biophys. Acta*:1282–1283.

Marks SC, Odgren PR (2002) Principles of Bone Biology. Elsevier.

Maroulakou IG, Papas TS, Green JE (1994) Differential expression of ets-1 and ets-2 proto-oncogenes during murine embryogenesis. *Oncogene* 9:1551–1565.

Marques CL, Cancela ML, Laizé V (2015) Transcriptional regulation of gilthead seabream bone morphogenetic protein (BMP) 2 gene by bone- and cartilage-related transcription factors. *Gene* in press.

Marques CL, Rafael MS, Cancela ML, Laizé V (2007) Establishment of primary cell cultures from fish calcified tissues. *Cytotechnology* 55:9–13.

Matys V, Fricke E, Geffers R, Gössling E, Haubrock M, Hehl R, et al. (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* 31:374–8.

McGonnell IM, Fowkes RC (2006) Fishing for gene function - Endocrine modelling in the zebrafish. *J. Endocrinol.* 189:425–439.

McGregor DO, Dellow WJ, Lever M, George PM, Robson RA, Chambers ST (2001) Dimethylglycine accumulates in uremia and predicts elevated plasma homocysteine concentrations. *Kidney Int.* 59:2267–72.

McKeever MP, Weir DG, Molloy A, Scott JM (1991) Betaine-homocysteine methyltransferase: organ distribution in man, pig and rat and subcellular distribution in the rat. *Clin. Sci.* 81:551–6.

Mcmahon C, Gestri G, Wilson SW, Link BA (2009) Lmx1b is essential for survival of periocular mesenchymal cells and influences Fgf-mediated retinal patterning in zebrafish. *Dev. Biol.* 332:287–298.

Metz J, de Vrieze E, Lock E, Schulten I, Flik G (2012) Elasmoid scales of fishes as model in biomedical bone research. *J. Appl. Ichthyol.* 28:382–387.

Meunier FJ, Huysseune A (1992) The concept of bone tissue in osteichthyes. *Netherlands J. Zool.* 42:445–458.

Meyer A, Van de Peer Y (2005) From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *BioEssays* 27:937–945.

Micale V, Di Giancamillo a, Domeneghini C, Mylonas CC, Nomikos N, Papadakis IE, et al. (2008) Ontogeny of the digestive tract in sharpsnout sea bream Diplodus puntazzo (Cetti, 1777). *Histol. Histopathol.* 23:1077–91.

Michigami T (2014) Current understanding on the molecular basis of chondrogenesis. *Clin. Pediatr. Endocrinol.* 23:1–8.

Michigami T (2013) Regulatory mechanisms for the development of growth plate cartilage. *Cell. Mol. Life Sci.* 70:4213–4221.

Minagawa M, Kwan MEIYEE, Bettoun JD, Mansour FW, Dassa J, Hendy GN, et al. (2000) Dissection of Differentially Regulated (G ⁄. C) -Rich Promoters of the Human Parathyroid Hormone. 141:2410–2421.

Minagawa M, Yasuda T, Watanabe T, Minamitani K, Takahashi Y, Goltzman D, et al. (2002) Association between AAAG repeat polymorphism in the P3 promoter of the human parathyroid hormone (PTH)/PTH-related peptide receptor gene and adult height, urinary pyridinoline excretion, and promoter activity. *J. Clin. Endocrinol. Metab.* 87:1791–6.

Mindnich R, Möller G, Adamski J (2004) The role of 17 beta-hydroxysteroid dehydrogenases. *Mol. Cell. Endocrinol.* 218:7–20.

Minina E, Wenzel HM, Kreschel C, Karp S, Gaffield W, McMahon AP, et al. (2001) BMP and Ihh/PTHrP signaling interact to coordinate chondrocyte proliferation and differentiation. *Development* 128:4523–34.

Mizrahi EH, Jacobsen DW, Friedland RP (2002) Plasma homocysteine: a new risk factor for Alzheimer's disease? *Isr. Med. Assoc. J.* 4:187–90.

Mohr W, Kuhn C, Pelster B, Wessinghage D (1985) S-100 protein in normal, osteoarthrotic, and arthritic cartilage. *Rheumatol. Int.* 5:273–277.

Mustonen MVJ, Poutanen M. H, Kellokumpu S, Launoit Y, Isomaa V V., Vihko RK, et al. (1998) Mouse 17β-hydroxysteroid dehydrogenase type 2 mRNA is predominantly expressed in hepatocytes and in surface epithelial cells of the gastrointestinal and urinary tracts. *J. Mol. Endocrinol.* 20:67–74.

GENERAL REFERENCES

Nakamura I, Takahashi N, Jimi E, Udagawa N, Suda T (2012) Regulation of osteoclast function. *Mod. Rheumatol.* 22:167–177.

Nakashima K, Zhou X, Kunkel G, Zhang Z, Deng J, Behringer R, et al. (2002) The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. *Cell* 108:17–29.

Nakatani Y, Kawakami A, Kudo A (2007) Cellular and molecular processes of regeneration, with special emphasis on fish fins. *Dev. Growth Differ.* 49:145–154.

Nakatani Y, Takeda H, Kohara Y, Morishita S (2007) Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* 17:1254–1265.

Naumann A, Dennis J, Awadallah A, Carrino D, Mansour J, Kastenbauer E, et al. (2002) Immunochemical and mechanical characterization of cartilage subtypes in rabbit. *J. Histochem. Cytochem.* 50:1049–1058.

Neece DJ, Griffiths MA, Garrow TA (2000) Isolation and characterization of a mouse betaine-homocysteine S-methyltransferase gene and pseudogene. *Gene* 250:31–40.

Nelen WL, Blom HJ, Steegers EA, den Heijer M, Thomas CM, Eskes TK (2000) Homocysteine and folate levels as risk factors for recurrent early pregnancy loss. *Obstet. Gynecol.* 95:519–24.

Nichols J, Pan L, Moens C, Kimmel C (2013) Barx1 represses joints and promotes cartilage in the craniofacial skeleton. *Development* 140:2765–2775.

Nissen RM, Amsterdam A, Hopkins N (2006) A zebrafish screen for craniofacial mutants identifies wdr68 as a highly conserved gene required for endothelin-1 expression. *BMC Dev. Biol.* 6:28.

North TE, Zon LI (2003) Modeling human hematopoietic and cardiovascular diseases in zebrafish. *Dev. Dyn.* 228:568–83.

Notredame C, Higgins D, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302:205–217.

Oberlender S, Tuan R (2008) Expression and functional involvement of N-cadherin in embryonic limb chondrogenesis. *Development* 120:177–187.

Oikawa T, Yamada T (2003) Molecular biology of the Ets family of transcription factors. *Gene* 303:11–34.

Olsen B, Reginato A, Wang W (2002) Bone development. *Annu. Rev. Cell Dev. Biol.* 16:191–220.

Oppermann U, Filling C, Hult M, Shafqat N, Wu X, Lindh M, et al. (2003) Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem. Biol. Interact.* 143-144:247–53.

Otto F, Thornell A, Crompton T, Denzel A, Gilmour K, Rosewell I, et al. (1997) Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell* 89:765–771.

Pajares M a, Pérez-Sala D (2006) Betaine homocysteine S-methyltransferase: just a regulator of homocysteine metabolism? *Cell. Mol. Life Sci.* 63:2792–803.

Paré J-F, Malenfant D, Courtemanche C, Jacob-Wagner M, Roy S, Allard D, et al. (2004) The fetoprotein transcription factor (FTF) gene is essential to embryogenesis and cholesterol homeostasis and is regulated by a DR4 element. *J. Biol. Chem.* 279:21206–16.

Park EI, Garrow TA (1999) Interaction between dietary methionine and methyl donor intake on rat liver betaine-homocysteine methyltransferase gene expression and organization of the human gene. *J. Biol. Chem.* 274:7816–7824.

Parra-Torres A, Valdés-Flores M (2013) Molecular Aspects of Bone Remodeling. In *Topics in Osteoporosis*. pp. 1–28.

Peng H, Tan L, Osaki M, Zhan Y, Ijiri K, Tsuchimochi K, et al. (2008) ESE-1 is a potent repressor of type II collagen gene (COL2A1) transcription in human chondrocytes. *J. Cell. Physiol.* 215:562–573.

Person P, Mathews M (1967) Endoskeletal cartilage in a marine polychaete, Eudistylia polymorpha. *Biol. Bull.* 132:244–252.

Persson B, Kallberg Y, Bray JE, Bruford E, Dellaporta SL, Favia AD, et al. (2009) The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative. *Chem. Biol. Interact.* 178:94–8.

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:e45.

Piera-Velazquez S, Hawkins DF, Whitecavage MK, Colter DC, Stokes DG, Jimenez SA (2007) Regulation of the human SOX9 promoter by Sp1 and CREB. *Exp. Cell Res.* 313:1069–79.

Di Pietro S, Santome J (2002) Structural and biochemical characterization of calhepatin, an S100-like calcium-binding protein from the liver of lungfish (Lepidosiren paradoxa). *Eur. J. Biochem.* 269:3433–3441.

Pinto J, Conceição N, Gavaia P, Cancela M (2003) Matrix Gla protein gene expression and protein accumulation colocalize with cartilage distribution during development of the teleost fish Sparus aurata. *Bone* 32:201–210.

Pinto J, Ohresser M, Cancela M (2001) Cloning of the bone Gla protein gene from the teleost fish Sparus aurata. Evidence for overall conservation in gene organization and bone-specific expression from fish to man. *Gene* 270:77–91.

Pinto W, Figueira L, Dinis MT, Aragão C (2009) How does fish metamorphosis affect aromatic amino acid metabolism? *Amino Acids* 36:177–83.

Pirraco RP, Marques AP, Reis RL (2010) Cell interactions in bone tissue engineering. *J. Cell. Mol. Med.* 14:93–102.

Pombinho AR, Laizé V, Molha DM, Marques SMP, Cancela ML (2004) Development of two bone-derived cell lines from the marine teleost Sparus aurata; evidence for extracellular matrix mineralization and cell-type-specific expression of matrix Gla protein and osteocalcin. *Cell Tissue Res.* 315:393–406.

Provot S, Schipani E (2005) Molecular mechanisms of endochondral bone development. *Biochem. Biophys. Res. Commun.* 328:658– 665.

Qi H, Aguiar DJ, Williams SM, La Pean A, Pan W, Verfaillie CM (2003) Identification of genes responsible for osteoblast differentiation from human mesodermal progenitor cells. *Proc. Natl. Acad. Sci. U. S. A.* 100:3305–10.

Qian Y-X, Song J-J (2011) Effects of salinity and betaine on BHMT mRNA expression in Lateolabrax japonicus. *Zoological. Res. Exp.* 32:277–84.

Quach HNB, Tao S, Vrljicak P, Joshi A, Ruan H, Sukumaran R, et al. (2015) A multifunctional mutagenesis system for analysis of gene function in zebrafish. *G3 (Bethesda).* 5:1283–99.

GENERAL REFERENCES

Rafael MS, Laizé V, Cancela ML (2006) Identification of Sparus aurata bone morphogenetic protein 2: molecular cloning, gene expression and in silico analysis of protein conserved features in vertebrates. *Bone* 39:1373–81.

Rafael MS, Marques CL, Parameswaran V, Cancela ML, Laizé V (2010) Fish bone-derived cell lines: An alternative in vitro cell system to study bone biology. *J. Appl. Ichthyol.* 26:230–234.

Rameshwar R, Stegemann J (2013) Cell-based approaches to the engineering of vascularized bone tissue. *Cytotherapy* 15:1309–1322.

Rao PV, Garrow TA, John F, Garland D, Millian NS, Zigler JJS (1998) Betaine-homocysteine methyltransferase is a developmentally regulated enzyme crystallin in rhesus monkey lens. *J. Biol. Chem.* 273:30669–30674.

Raouf A, Seth A (2002) Discovery of osteoblast-associated genes using cDNA microarrays. *Bone* 30:463–71.

Raouf A, Seth A (2000) Ets transcription factors and targets in osteogenesis. *Oncogene* 19:6455 – 6463.

Rauch A, Seitz S, Baschant U, Schilling AF, Illing A, Stride B, et al. (2010) Glucocorticoids suppress bone formation by attenuating osteoblast differentiation via the monomeric glucocorticoid receptor. *Cell Metab.* 11:517–31.

Ravasi T, Hsu K, Goyette J, Schroder K, Yang Z, Rahimi F, et al. (2004) Probing the S100 protein family through genomic and functional analysis. *Genomics* 84:10–22.

Renn J, Buttner A, To T, Chan S, Winkler C (2013) A col10a1:nlGFP transgenic line displays putative osteoblast precursors at the medaka notochordal sheath prior to mineralization. *Dev. Biol.* 381:134–143.

Rety S, Osterloh D, Arie J, Tabaries S, Seeman J, Russo-Marie F, et al. (2000) Structural basis of the Ca(2+)-dependent association between S100C (S100A11) and its target, the N-terminal part of annexin I. *Structure* 8:175–184.

Ribes V, Le Roux I, Rhinn M, Schuhbaur B, Dollé P (2009) Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. *Development* 136:665–76.

Rintala-Dempsey A, Rezvanpour A, Shaw G (2008) S100-annexin complexes--structural insights. *FEBS J.* 275:4956–4966.

Rosa JT, Cancela ML, Laizé V (2014) Ets1 regulates the transcription of a cartilage-specific S100 protein in gilthead seabream. *J. Appl. Ichthyol.* 30:707–712.

Ross MH, Pawlina W (2011) Histology. A Text and Atlas.

Roy P, Witten P, Hall B, Lall S (2002) Effects of dietary phosphorus on bone growth and mineralisation of vertebrae in haddock (Melanogrammus aeglefinus L.). *Fish Physiol. Biochem.* 27:35–48.

Sakamoto A, Nishimura Y, Ono H, Sakura N (2002) Betaine and homocysteine concentrations in foods. *Pediatr. Int.* 44:409–13.

Sandell LL, Sanderson BW, Moiseyev G, Johnson T, Mushegian A, Young K, et al. (2007) RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development. *Genes Dev.* 21:1113–1124.

Sarropoulou E, Kotoulas G, Power D, Geisler R (2005) Gene expression profiling of gilthead sea bream during early development and detection of stress-related genes by the application of cDNA microarray technology. *Physiol. Genomics* 23:182–191.

Sarropoulou E, Power D, Magoulas A, Geisler R, Kotoulas G (2005) Comparative analysis and characterization of expressed sequence tags in gilthead sea bream (Sparus aurata) liver and embryos. *Aquaculture* 243:69–81.

Sato M, Morii E, Komori T, Kawahata H, Sugimoto M, Terai K, et al. (1998) Transcriptional regulation of osteopontin gene in vivo by PEBP2alphaA/CBFA1 and ETS1 in the skeletal tissues. *Oncogene* 17:1517–25.

Schneider T, Stephens R (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18:6097–6100.

Schorle H, Meier P, Buchert M, Jaenisch R, Mitchell PJ (1996) Transcription factor AP-2 essential for cranial closure and craniofacial development. *Nature* 381:235–8.

Scillitani A, Jang C, Wong BY-L, Hendy GN, Cole DEC (2006) A functional polymorphism in the PTHR1 promoter region is associated with adult height and BMD measured at the femoral neck in a large cohort of young caucasian women. *Hum. Genet.* 119:416–21.

Senger F, Priat C, Hitte C, Sarropoulou E, Franch R, Geisler R, et al. (2006) The first radiation hybrid map of a perch-like fish: The gilthead seabream (Sparus aurata L). *Genomics* 87:793–800.

Shahar R, Dean MN (2013) The enigmas of bone without osteocytes. *Bonekey Rep.* 2:343.

Shen Z, Peng Z, Sun Y, Poutanen M (2008) Overexpression of human hydroxysteroid (17β) dehydrogenase 2 induces disturbance in skeletal development in young male mice. *J. Bone Miner. Res.* 23.

Siegel N, Hoegg S, Salzburger W, Braasch I, Meyer A (2007) Comparative genomics of ParaHox clusters of teleost fishes: gene cluster breakup and the retention of gene sets following whole genome duplications. *BMC Genomics* 8:312.

Simes D, Williamson M, Ortiz-Delgado J, Viegas C, Price P, Cancela M (2003) Purification of matrix Gla protein from a marine teleost fish, Argyrosomus regius: Calcified cartilage and not bone as the pri- mary site of MGP accumulation in fish. *J. Bone Miner. Res.* 18:244–259.

Skiba WE, Taylor MP, Wells MS, Mangum JH, Awad WM (1982) Human hepatic methionine biosynthesis. Purification and characterization of betaine:homocysteine S-methyltransferase. *J. Biol. Chem.* 257:14944–8.

Slebos RJC, Oh DS, Umbach DM, Taylor JA (2002) Mutations in Tetranucleotide Repeats following DNA Damage Depend on Repeat Sequence and Carcinogenic Agent Mutations in Tetranucleotide Repeats following DNA Damage Depend on Repeat Sequence and Carcinogenic Agent. :6052–6060.

Spoorendonk KM, Hammond CL, Huitema LFA, Vanoevelen J, Schulte-Merker S (2010) Zebrafish as a unique model system in bone research: the power of genetics and in vivo imaging. *J. Appl. Ichthyol.* 26, 2:219–224.

Stefansson K, Wollmann RL, Moore BW, Arnason BGW (1982) S-100 protein in human chondrocytes. *Nature* 295:63–64.

Stein G, Lian J, Stein J, Wijnen A van, Montecino M (1996) Transcriptional control of osteoblast growth and differentiation. *Physiol. Rev.* 76:593–629.

GENERAL REFERENCES

Stemple DL (2004) The notochord. *Curr. Biol.* 14:R873–R874.

Stern CD (2004) Gastrulation: from cells to embryo. In *Cold Spring Harbor Laboratory Press*.

Sugiura K, Ito K (2010) Roles of Ets-1 and p70S6 kinase in chondrogenic and gliogenic specification of mouse mesencephalic neural crest cells. *Mech. Dev.* 127:169–82.

Sunden SLF, Renduchintala MS, Park EI, Miklasz SD, Garrow TA (1997) Betaine-homocysteine methyltransferase expression in porcine and human tissues and chromosomal localization of the human gene 1. *Arch. Biochem. Biophys* 345:171–174.

Suzuki N, Suzuki T, Kurokawa T (2000) Suppression of osteoclastic activities by calcitonin in the scales of goldfish (freshwater teleost) and nibbler fish (seawater teleost). *Peptides* 21:115–124.

Szegedi SS, Castro CC, Koutmos M, Garrow T a (2008) Betaine-homocysteine S-methyltransferase-2 is an S-methylmethionine-homocysteine methyltransferase. *J. Biol. Chem.* 283:8939–45.

Szegedi SS, Garrow T a (2004) Oligomerization is required for betaine-homocysteine S-methyltransferase function. *Arch. Biochem. Biophys.* 426:32–42.

Taichman RS (2005) Blood and bone: two tissues whose fates are intertwined to create the hematopoietic stem-cell niche. *Blood* 105:2631–2639.

Takagi Y, Yamada J (1992) Effects of calcium deprivation on the metabolism of acellular bone in tilapia, Oreochromis niloticus. *Comp. Biochem. Physiol.* 102:481–485.

Tamplin OJ, White RM, Jing L, Kaufman CK, Lacadie SA, Li P, et al. (2012) Small molecule screening in zebrafish: swimming in potential drug therapies. *Interdiscip. Rev. Dev. Biol.* 1:459–68.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10:512–26.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30:2725–9.

Tiago DM, Cancela ML, Aureliano M, Laizé V (2008) Vanadate proliferative and anti-mineralogenic effects are mediated by MAPK and PI-3K/Ras/Erk pathways in a fish chondrocyte cell line. *FEBS Lett.* 582:1381–5.

Tiago DM, Laizé V, Bargelloni L, Ferraresso S, Romualdi C, Cancela ML (2011) Global analysis of gene expression in mineralizing fish vertebra-derived cell lines: new insights into anti-mineralogenic effect of vanadate. *BMC Genomics* 12:310.

Tiago DM, Marques CL, Roberto VP, Cancela ML, Laizé V (2014) Mir-20a regulates in vitro mineralization and BMP signaling pathway by targeting BMP-2 transcript in fish. *Arch. Biochem. Biophys.* 543:23–30.

Tomlinson JW, Walker EA, Bujalska IJ, Draper N, Lavery GG, Cooper MS, et al. (2004) 11β-hydroxysteroid dehydrogenase type 1: a tissue-specific regulator of glucocorticoid response. *Endocr. Rev.* 25:831–66.

Trojanowska M (2000) Ets factors and regulation of the extracellular matrix. *Oncogene* 19:6464–71.

Tuan R, Scott W (1977) Calcium-binding protein of chorioallantoic membrane: Identification and developmental expression. *Proc. Natl. Acad. Sci.* 74:1946–1949.

Tuli R, Tuli S, Nandi S, Huang X, Manner P, Hozack W, et al. (2003) Transforming growth factor-beta-mediated chondrogenesis of human mesenchymal progenitor cells involves N-cadherin and mitogen-activated protein kinase and Wnt signaling cross-talk. *J. Biol. Chem.* 278:41227–41236.

Umesono K, Murakami KK, Thompson CC, Evans RM (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* 65:1255–66.

Urist MR (1965) Bone: formation by autoinduction. *Science* 150:893–9.

Ushigome S, Takakuwa T, Shinagawa T, Takagi M, Kishimoto H, Mori N (1984) Ultrastructure of cartilaginous tumors and S-100 protein in the tumors. With reference to the histogenesis of chondroblastoma, chondromyxoid fibroma and mesenchymal chondrosarcoma. *Acta Pathol. Jpn.* 34:1285–1300.

Uthus EO, Yokoi K, Davis CD (2002) Selenium deficiency in fisher-344 rats decreases plasma and tissue homocysteine concentrations and alters plasma homocysteine and cysteine redox status. *J. Nutr.* 132.

Vandepoele K, De Vos W, Taylor J, Meyer A, Van de Peer Y (2004) Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. *Proc. Natl. Acad. Sci. USA* 101:1638–1643.

Vanderpoele K, De Vos W, Taylor J, Meyer A, Van de Peer Y (2004) Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. *Proc. Natl. Acad. Sci.* 101:1638–1643.

Viegas M, Dias J, Cancela ML, Laizé V (2012) Polyunsaturated fatty acids regulate cell proliferation, extracellular matrix mineralization and gene expression in a gilthead seabream skeletal cell line. *J. Appl. Ichthyol.* 28:427–432.

Vihko P, Herrala A, Härkönen P, Isomaa V, Kaija H, Kurkela R, et al. (2005) Enzymes as modulators in malignant transformation. *J. Steroid Biochem. Mol. Biol.* 93:277–83.

Vijayakumar P, Laizé V, Cardeira J, Trindade M, Cancela ML (2013) Development of an in vitro cell system from zebrafish suitable to study bone cell differentiation and extracellular matrix mineralization. *Zebrafish* 10:500–9.

De Vrieze E, van Kessel M, Peters H, Spanings F, Flik G, Metz J (2014) Prednisolone induces osteoporosis-like phenotype in regenerating zebrafish scales. *Osteoporos. Int.* 25:567–578.

De Vrieze E, Sharif F, Metz J, Flik G, Richardson M (2011) Matrix metalloproteinases in osteoclasts of ontogenetic and regenerating zebrafish scales. *Bone* 48:704–712.

Wagner TU, Renn J, Riemensperger T, Volff J-N, Köster RW, Goerlich R, et al. (2003) The teleost fish medaka (Oryzias latipes) as genetic model to study gravity dependent bone homeostasis in vivo. *Adv. Space Res.* 32:1459–65.

Walker M, Kimmel C (2007) A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech. Histochem.* 82:23–28.

Wallace I, O'Sullivan O, Higgins D, Notredame C (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* 34:1692–1699.

Wang S, Kryvi H, Grotmol S, Wargelius A, Krossøy C, Epple M, et al. (2013) Mineralization of the vertebral bodies in Atlantic salmon (Salmo salar L.) is initiated segmentally in the form of hydroxyapatite crystal accretions in the notochord sheath. *J. Anat.* 223:159–70.

GENERAL REFERENCES

Wehner D, Jahn C, Weidinger G (2015) Use of the TetON system to study molecular mechanisms of zebrafish regeneration. *J. Vis. Exp.*

Weinstein RS, Jilka RL, Parfitt AM, Manolagas SC (1998) Inhibition of osteoblastogenesis and promotion of apoptosis of osteoblasts and osteocytes by glucocorticoids. Potential mechanisms of their deleterious effects on bone. *J. Clin. Invest.* 102:274–82.

Wenke A-K, Bosserhoff AK (2010) Roles of AP-2 transcription factors in the regulation of cartilage and skeletal development. *FEBS J.* 277:894–902.

Wenke A-K, Rothhammer T, Moser M, Bosserhoff AK (2006) Regulation of integrin alpha10 expression in chondrocytes by the transcription factors AP-2epsilon and Ets-1. *Biochem. Biophys. Res. Commun.* 345:495–501.

Wijekoon EP, Hall B, Ratnam S, Brosnan ME, Zeisel SH, Brosnan JT (2005) Homocysteine metabolism in ZDF (type 2) diabetic rats. *Diabetes* 54:3245–51.

Witten P, Huysseune A (2009) A comparative view on mechanisms and functions of skeletal remodelling in teleost fish, with special emphasis on osteoclasts and their function. *Biol. Rev. Camb. Philos. Soc.* 84:315–346.

Witten P, Huysseune A, Hall B (2010) A practical approach for the identification of the many cartilaginous tissues in teleost fish. *J. Appl. Ichthyol.* 26:257–262.

Wuelling M, Vortkamp A (2010) Transcriptional networks controlling chondrocyte proliferation and differentiation during endochondral ossification. *Pediatr. Nephrol.* 25:625–631.

Xua Z, Ouyanga L, del Castillo-Olivaresa A, Pandakc W, Gil G (2010) α1-Fetoprotein transcriptionfFactor (FTF)/liver receptor momolog-1 (LRH-1) is an essential lipogenic regulator. *Biochim. Biophys. Acta* 1801:473–479.

Yang S-L, Aw SS, Chang C, Korzh S, Korzh V, Peng J (2011) Depletion of Bhmt elevates sonic hedgehog transcript level and increases β-cell number in zebrafish. *Endocrinology* 152:4706–17.

Yoshida CA, Komori T (2005) Role of runx proteins in chondrogenesis. *Crit. Rev. Eukaryot. Gene Expr.* 15:243–254.

Zdobnov E, Apweiler R (2001) InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17:847–848.

Zeisel SH, Mar M-H, Howe JC, Holden JM (2003) Concentrations of choline-containing compounds and betaine in common foods. *J. Nutr.* 133:1302–7.

Zeng L, Murtaugh LC, Sato ME, Lassar AB (2005) Shh establishes an Nkx3.2/Sox9 autoregulatory loop that is maintained by BMP signals to induce somitic chondrogenesis. *Genes Dev* 16:1990–2005.

Zhang G, Cohn MJ (2006) Hagfish and lancelet fibrillar collagens reveal that type II collagen-based cartilage evolved in stem vertebrates. *Proc Natl Acad Sci U S A* 103:16829–16833.

Zhang G, Eames BF, Cohn MJ (2009) Chapter 2. Evolution of vertebrate cartilage development. *Curr. Top. Dev. Biol.* 86:15–42.

Zhang J, Hagopian-Donaldson S, Serbedzija G, Elsemore J, Plehn-Dujowich D, McMahon AP, et al. (1996) Neural tube, skeletal and body wall defects in mice lacking transcription factor AP-2. *Nature* 381:238–41.

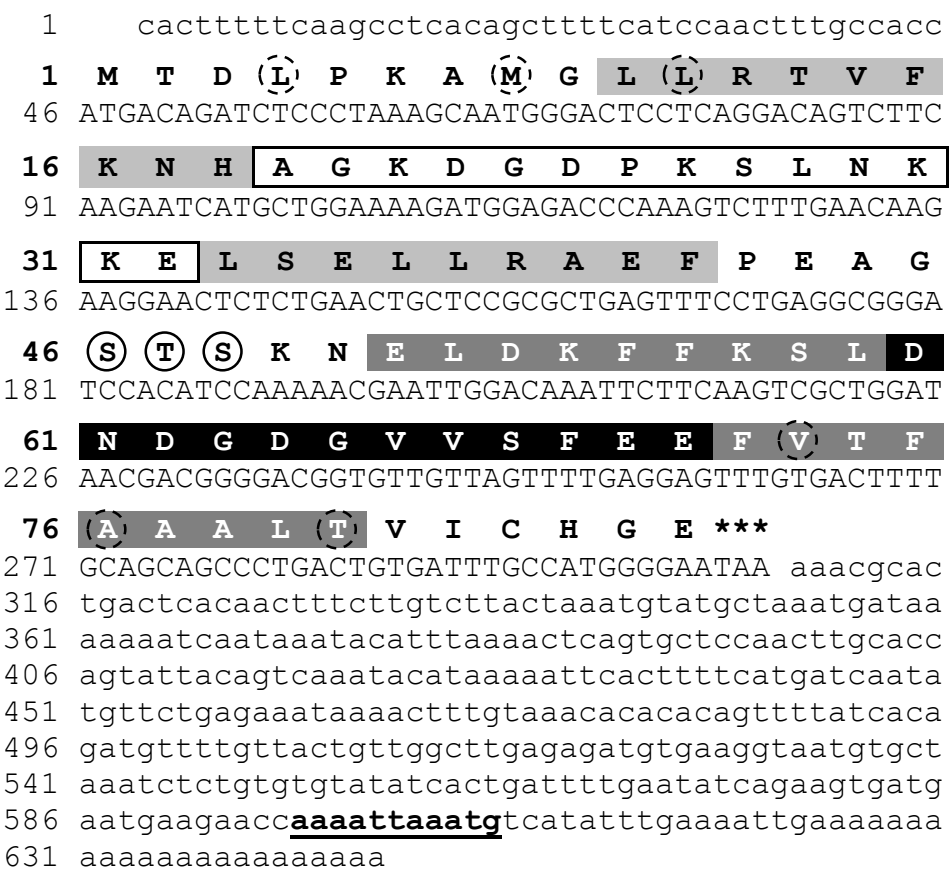
Zheng W, Li Z, Nguyen AT, Li C, Emelyanov A, Gong Z (2014) Xmrk, kras and myc transgenic zebrafish liver cancer models share molecular signatures with subsets of human hepatocellular carcinoma. *PLoS One* 9, 3.

APPENDIX

APPENDIX

Supplementary Data Chapter 2

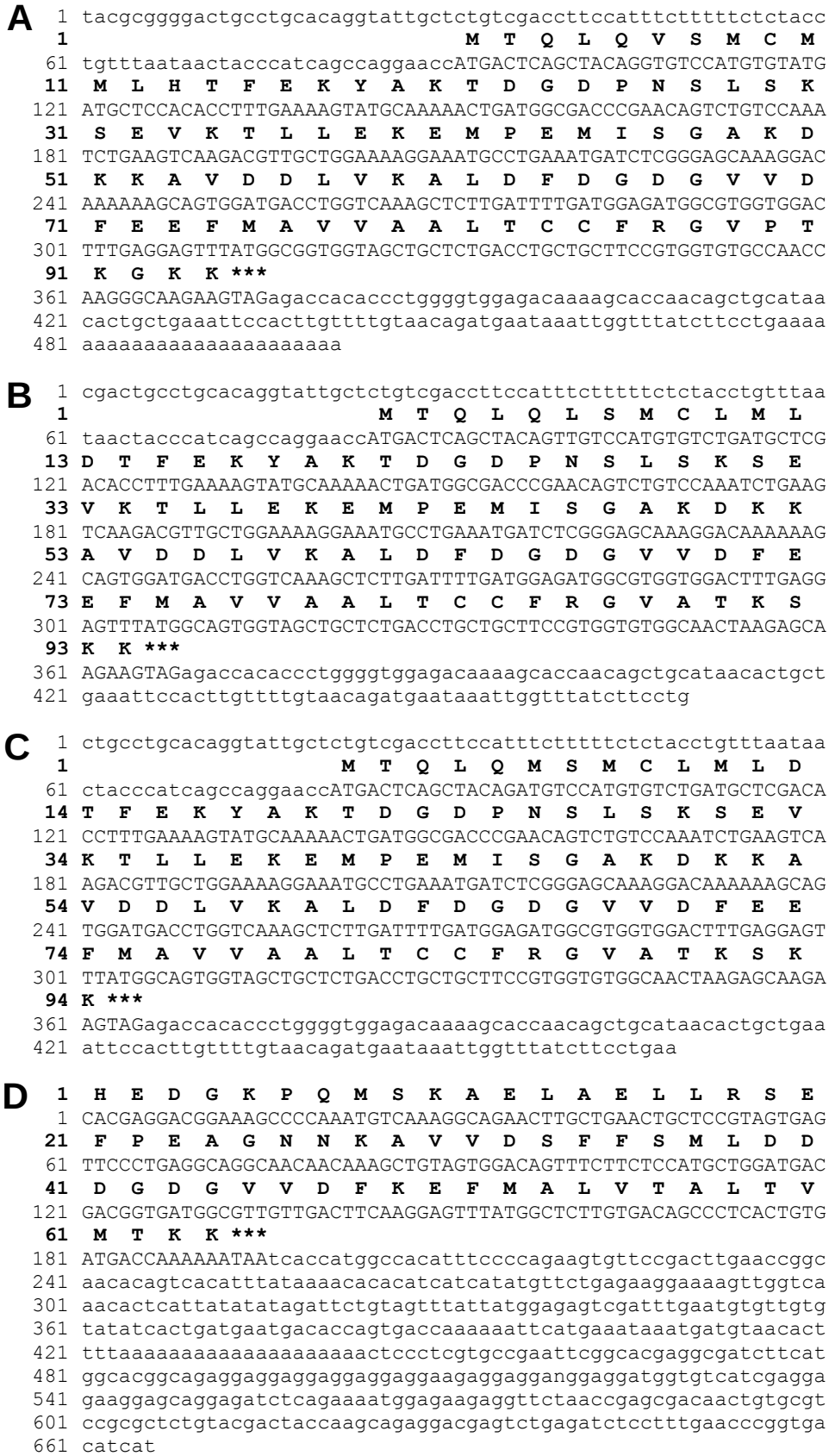
Supplementary Figure 2.1



Gilthead seabream *S100-like* cDNA and deduced amino acid sequences. Nucleotides and amino acids (*bold*) are numbered on the left. N-terminal non-canonical EF-hand is indicated by *light gray* (□-helixes) and *white* (calcium-binding loop) boxes. C-terminal canonical EF-hand is indicated by *dark gray* (□-helixes) and *black* (calcium-binding loop) boxes. Putative phosphorylated residues are *circled*, putative poly-adenylation site is *bold-underlined* and putative hydrophobic residues involved in dimer formation are *circled dashed*. *Asterisks* indicate stop codon. S100-like cDNA sequence can be retrieved from GenBank using accession number AY787209.

APPENDIX

Supplementary Figure 2.2



E

```
   1                                                          M
   1 gtttgtacaaaaagttggtatatgttctgtagcaagtctcacaaattttcgacaccATGA
   2 T  T  I  F  E  A  I  T  I  L  R  Q  V  F  D  K  Y  A  G  K
  61 CTACAATCTTTGAGGCAATCACAATTCTCAGGCAGGTTTTTGACAAGTATGCTGGGAAAG
  22 E  G  D  P  K  T  L  T  K  K  E  V  L  E  L  L  K  E  Q  L
 121 AAGGAGATCCAAAGACTCTAACAAAGAAGGAAGTCTTAGAACTGCTGAAGGAGCAGCTTG
  42 G  E  A  P  S  S  K  A  E  L  D  D  F  F  N  M  L  D  D  D
 181 GAGAAGCTCCTTCAAGCAAAGCAGAATTGGACGATTTTTTCAATATGCTCGATGATGATG
  62 G  D  G  V  V  D  F  N  E  Y  V  V  L  C  A  S  L  A  L  M
 241 GCGACGGCGTTGTCGATTTCAACGAGTATGTGGTCCTATGCGCAAGCTTGGCTTTGATGT
  82 L  A  P  ***
 301 TGGCCCCCTAAaggaaagaacagtctcttagttctggtttacttcacttaaaataagtgt
 361 ttgtttttttttaaaaaaggtaacaaccatgaactactgttgaatctcttttaaatgtg
 421 ctttttttgcaaccaccataatgtgtagtcttgtgcaagccatgctaaacttttctcttt
 481 ccccccccatgaaacctgaagttgtttgaactgtacttgcatgttatttgactgtccgatg
 541 gtggattctaacaaaatgcaaataaaatttgaatctctatcaaaaaaaaaa
```

F

```
   1                        M  S  G  V  L  T  T  V  A  L  L  K
   1 ggggaatcggcacgagaaacgccacATGTCTGGAGTATTGACCACAGTTGCGCTCCTTAA
  13   Q  T  F  D  K  Y  A  G  A  D  G  D  K  N  Q  M  S  K  K  E
  61 GCAGACCTTTGACAAGTATGCTGGAGCAGATGGAGACAAGAACCAAATGTCCAAGAAAGA
  33   L  A  T  M  L  K  A  E  L  P  G  C  C  D  K  Q  S  E  V  D
 121 ACTCGCTACCATGCTCAAAGCCGAGCTTCCTGGATGTTGTGACAAACAGTCAGAGGTGGA
  53   E  F  L  K  M  L  D  Q  D  G  D  G  S  L  S  F  E  E  F  V
 181 TGAATTTTTGAAGATGCTGGATCAGGATGGTGACGGTTCTTTAAGTTTTGAAGAGTTTGT
  73   T  F  V  A  T  F  A  V  L  V  G  S  A  *
 241 CACTTTTGTTGCAACGTTTGCTGTGCTCGTGGGCTCTGCTTAAaactttccagcctgact
 301 gctttgttttcttgtctcatctaatgtctcgtaaaaattctctgaaataaacttgctttg
 361 ccacactgtttgtggtttaaaataaagtttgagcaggttaaagtcacggtgaatgtggtg
 421 tacaatgttgactaaaacatggaaaataaagcctttctcacatgtcaaaaaaaaaaaaaa
 481 aaaa
```

S100-like **calcium-binding cDNA and deduced amino acid sequences**. **A**, *Astatotilapia burtoni* S100-L reconstructed from ESTs DY626581, DY628313, DY632220 and DY628572; **B**, *Lipochromis sp.* S100-L reconstructed from ESTs DB872930, DB864386, DB868592 and DB860133; **C**, *Paralabidochromis chilotes* S100-L reconstructed from BJ689580, BJ689818, BJ675903, BJ673443, BJ670934, BJ678694, BJ669607, BJ674669, BJ674665, BJ688240, BJ687903, BJ685119, BJ688257, BJ674683, BJ679005, BJ686725, BJ680573, BJ688472, BJ671093, BJ674879, BJ687962, BJ669882, BJ674512, BJ686579, BJ683112, BJ684770, BJ682379, BJ678378, BJ669315, BJ678718, BJ669627, BJ683523, BJ680397, BJ682586 and BJ682671; **D**, *Paralichthys olivaceus* S100-L reconstructed from CX284978; **E**, *Oryzias latipes* S100-L reconstructed from BJ886212, BJ879615, AU176585, BJ900256 and BJ893349; **F**, *Fundulus Heteroclitus* S100-L reconstructed from CN983748, CN972572, CN971173, CN954209 and CN957343

APPENDIX

Supplementary Table 2.1

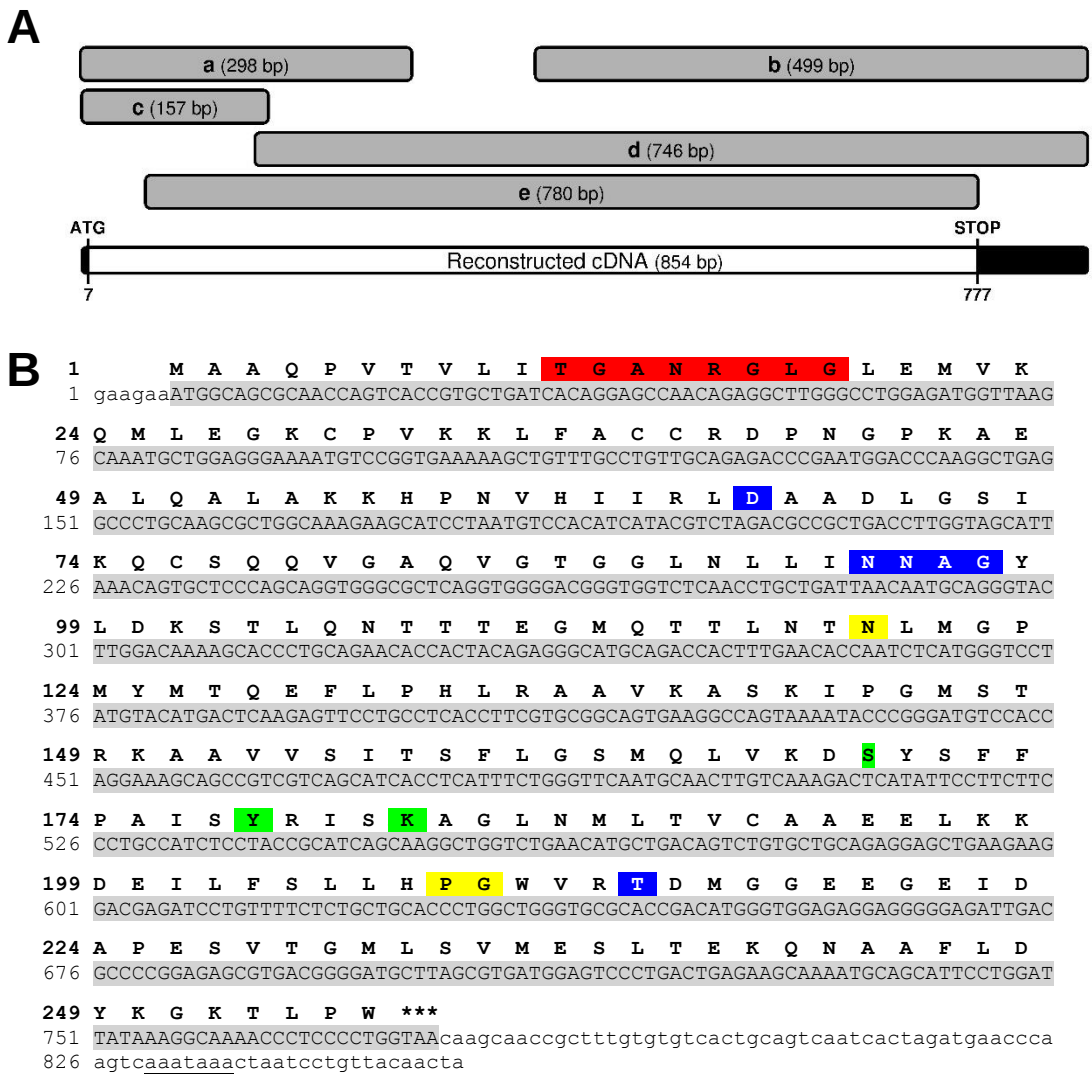
PCR primers used in this study

Name	Sequence (5'-3')*
S100_Rv1	GTCCTGAGGAGTCCCATTGCTT
S100_Rv2	CATGGTGGCAAAGTTGGATGA
S100_Rv3	CACGCAAGCTTCTGTCATGGTGGCAAAGTTGGATGAAAAG
S100_Fw1	CCGGAGCTCGAGCATGTAAGTGAAAGTAATCCCAGGGGCTG
S100_Fw2	CCGGAGCTCGAGGTTTCATGAGTCCTCAGATATG
S100_Fw3	CCGGAGCTCGAGGAATAGTAATTCCTTGCATTTC
S100_Fw4	CCGGAGCTCGAGCAGGATATGATGTTGGCACCAG
S100_Fw5	CCGGAGCTCGAGAAATCCTTTTAGATGATCAAT
S100_Fw6	CCGGAGCTCGAGGGCTTGTGATTTTCACAGTCAT
S100_Fw7	CCGGAGCTCGAGGGGGAGAAAAATCAGAAAGAAGG
S100_Ets1_1_Fw	GTTTCTACAAG*TC*AATCTTGACTCAACCCA
S100_Ets1_1_Rv	TGGGTTGAGTCAAGATT*GA*CTTGTAGAAAC
S100_Ets1_2_Fw	GATTAATCATAGACAG*TC*TATGATGTTGGCACCA
S100_Ets1_2_Rv	TGGTGCCAACATCATA*GA*CTGTCTATGATTAATC

* Underlined nucleotides indicate restriction site for *Xho*I in sense primers (Fw) and *Hin*dIII in antisense primers (Rv). Italicized nucleotides indicate mutations in putative ETS1 binding site core.

Supplementary Data Chapter 3

Supplementary Figure 3.1



Gilthead seabream *sdr-like* cDNA. (**A**) cDNA fragments obtained from the subtractive library (**a** and **b**) and amplified by PCR using primers AP1 and SDR-RV2 (**c**), AP1 and SDR-FW1 (**d**) and SDR-FW2 and SDRreal-RV (**e**). The structure of reconstructed cDNA is indicated below cDNA fragments. (**B**) cDNA and deduced amino acid sequences. Polyadenylation signal is underlined and coding sequence is highlighted in grey. Motifs conserved in SDRs are highlighted in red (cofactor binding site), blue (structure stabilizing motifs), green (active center) and yellow (catalysis enhancing sites). Sequence of gilthead seabream *sdr-like* cDNA can be retrieved from GenBank using accession No. EU557022).

APPENDIX

Supplementary Figure 3.2

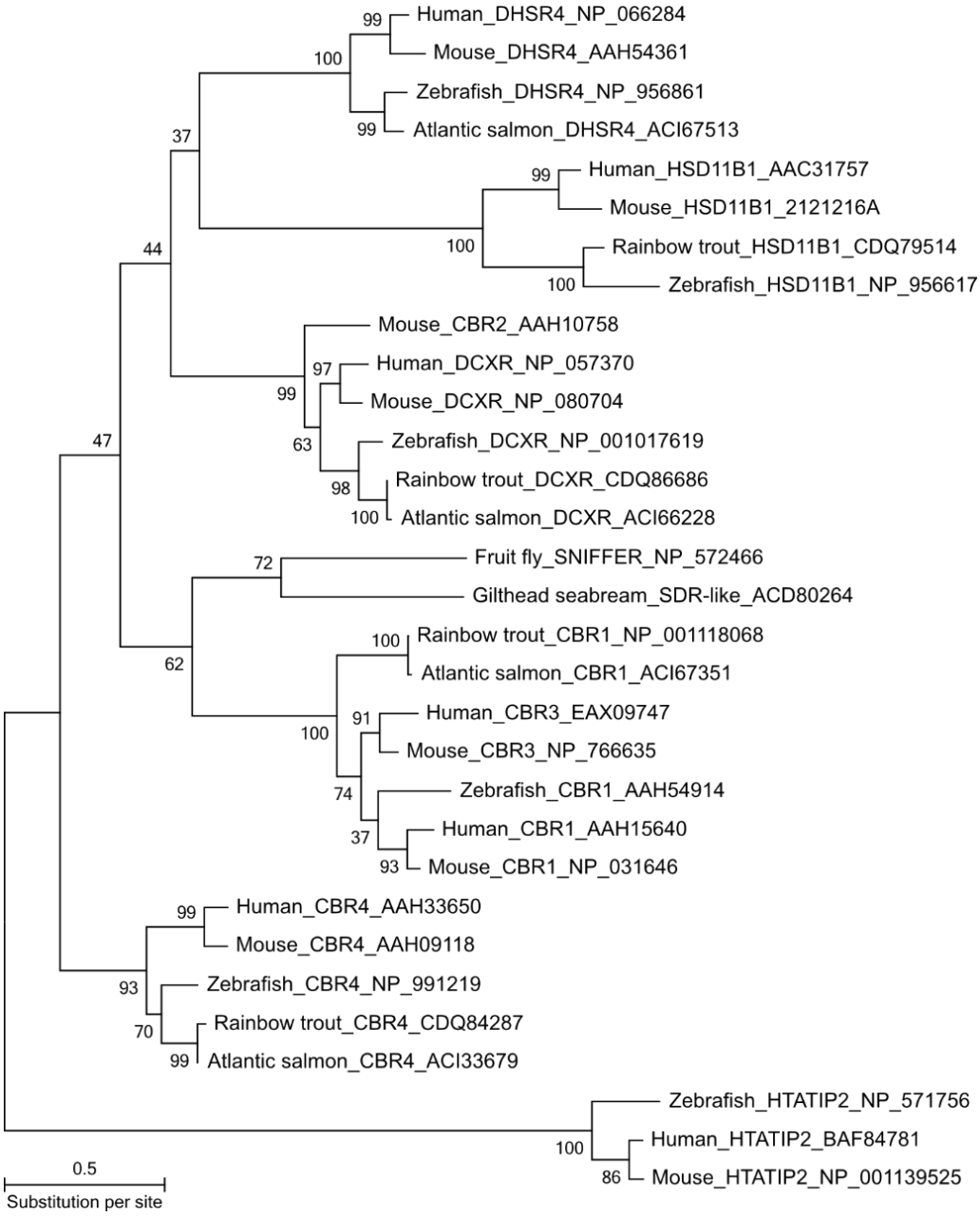
A

B

```
   1                                                    atcatctgctttaatctttaaatcatattgattaagtaatta
  43 tgttctgtgcacaagtctttaaagtcaccatgtggcgctgttccacacagccaccttggcggggcagatcttcag
 118 tggtttggggttgttgccattaatgttggctccagggcagaagaaggggaacatcttctcagtgaagaagtcggt
 193 gaaggtgtggatgggcgtcatgctgacagggttggagaatgccacctctccagattcgtagtccaactgcaccct
 268 gactgaattttgaggtagagaggaagcagatctttacacaagctgagacatgagaatgtgggtgtgctcgggcta
   1                                                                M  A  A  Q  P  V  T  V  L
 343 aaggccaacattcaacaggtgctttataaaagggttcagaggcagaagaaaATGGCAGCGCAACCAGTCACCGTGC
  10   I  T  G  A  N  R  G  L  G  L  E  M  V  K  Q  M  L  E  G  K  C  P  V  K  K
 418 TGATCACAGGAGCCAACAGAGGCTTGGGCCTGGAGATGGTTAAGCAAATGCTGGAGGGAAAATGTCCGGTGAAAA
  35   L  F  A  C  C  R  D  P  N  G  P  K  A  E
 493 AGCTGTTTGCCTGTTGCAGAGACCCGAATGGACCCAAGGCTGAGgtgagtctgtacaacctatgttccttcattg
 568 agctcagattgagctctagcaaaaatcagaaagtgttcttgcaaaccacagagttgcagtgattgagaggtgaga
  49                                                                                A  L  Q
 643 aaatgtgaagtgatcagctgtgtgcatcctctgaactacaaacatgtcttcactgtgtatttacagGCCCTGCAA
  52 A  L  A  K  K  H  P  N  V  H  I  I  R  L  D
 718 GCGCTGGCAAAGAAGCATCCTAATGTCCACATCATACGTCTAGgtatacactccttctttctgttatatactgac
 793 aaactctctgtcctgctcccaacagtcctgtgttgtgatatttatctgaatgtctgctactgactgactctcta
  67                          A  A  D  L  G  S  I  K  Q  C  S  Q  Q  V  G  A  Q  V
 868 ctgtctctccctctgtgtcagACGCCGCTGACCTTGGTAGCATTAAACAGTGCTCCCAGCAGGTAGGCGCTCAGG
  85   G  T  G  G  L  N  L  L  I  N  N  A  G  Y  L  D  K  S  T  L  Q  N  T  T  T
 943 TGGGGACGGGTGGTCTCAACCTGCTGATTAACAATGCAGGGTACTTGGACAAAAGCACCCTGCAGAACACCACTA
 110   E  G  M  Q  T  T  L  N  T  N  L  M  G  P  M  Y  M  T  Q
1018 CAGAGGGCATGCAGACCACTTTGAACACCAATCTCATGGGTCCTATGTACATGACTCAAgtaagtttgactcaag
 129                                                                  E  F  L  P  H  L
1093 cagtttgaacgtgttgtgcttcttcttgttgaccctcttgcttctggtcctcattagGAGTTCCTGCCTCACCTT
 135 R  A  A  V  K  A  S  K  I  P  G  M  S  T  R  K  A  A  V  V  S  I  T  S  F
1168 CGTGCGGCAGTGAAGGCCAGTAAAATACCCGGGATGTCCACCAGGAAAGCAGCCGTCGTCAGCATCACCTCATTT
 160 L  G  S  M  Q  L  V  K  D  S  Y  S  F  F  P  A  I  S  Y  R  I  S  K
1243 CTGGGTTCAATGCAACTTGTCAAAGACTCATATTCCTTCTTCCCTGCCATCTCCTACCGCATCAGCAAGgtgaga
1318 aacaacctgccaacactgtaacctgctgctcaaaggagtgacctcacattcattgagtgaaaatcgtcatgcatc
1393 acctcaactgacctgtcgctgttctatactttcctgtcaccctctttacctccttctcttcttttcccttctctt
 183       A  G  L  N  M  L  T  V  C  A  A  E  E  L  K  K  D  E  I  L  F  S  L  L
1468 ctcagGCTGGTCTGAACATGCTGACAGTCTGTGCTGCAGAGGAGCTGAAGAAGGACGAGATCCTGTTTTCTCTGC
 207   H  P  G  W  V  R  T  D  M  G  G  E  E
1543 TGCACCCTGGCTGGGTGCGCACCGACATGGGTGGAGAGGAGgttagttagcaactctcactgacctcacaggtcc
1618 agtcataaaaggatataatgaaactatcgtagaggacaggttcagctctgaatgatacaattgatttattgatta
1693 agaggtgcagtaattccatcatgttgtttgatctcactcaatcgcagtattccaaagaagttacagaaagcagga
1768 ttatcaagttgaacaggtcgctgatacaaattacatcatgaaatgttgctcattatttgcaagtcatcattacat
1843 aacagaatcattcaagtttgcttttgtgtgtccgagtcacaaaacaaataacactaacctctccgataattctct
1918 gtagtgattcaattcagtctggactatgacctaaaacatagaaatgtgcggttaaagctgactgtagataaactc
1993 tagtgcagcttttcacctacttatgaaggcagctagcagcaggactacaggaaacctgaggtttcttatttgcac
2068 agcaagctgcggttatctgcgtctgactgacgctctgctatttttaactgtaaaataaagtgttaagtaacgtga
2143 agttcgtataaatgtatctttaataatggatgcatgtttcattgttggtgggtgtgcatgcaaaagatgataact
2218 tgtgaaaatgcatcgactaccacagctgattgttgaggtgtgtgtaggtgagcccatcagtcagcccttctgatt
2293 aattagaagacgcctcttgtgttgcaattgtcagatattgtcttgatatataatcatatcagaaactatctaaac
2368 agtggatcagaattcttgttattacaggttttctgaagctttatgttaaaatatagcattattaaaattgcatta
2443 ttttgcacatatttccagagcaacaattcgataaagccaggttcaaatttcttgtttcattttgttgagatatta
2518 gactcaagggaattttggacatctctttttatcactgcttaaaaaagaaaacaaatcacttttttaggaatacaa
2593 ttttccagaaatcaggcacaatgaacaatatatgaacaaactcctcagtgaaaacctccggaacatagatgggaa
2668 aaaaacttgaaaagtttagtgcgggaaggtgatactaaagtggagatttctacaacagagcacaagaaaaaactg
2743 tagatttccaccaatttttaggggagaaaatgtaaatagggaaaaaaaaatttaactaatagacatcacagctaa
2818 gcttctttagtcgataactaaaataagacaacaattgtttgtttaatcaaatgttctcatcaaaatgaggctatc
2893 tagtaaaatattggcagatttgaataaatctccacaacaaagatatgaactaaaataaacaactagttgtagtta
2968 tggaagcaaaatatcccgaaatctaataattgaccagttcatgaaaatgttttcacttgtagagtgaagacgaca
3043 ttaaaaaacagttctgtgtcaaatacgaggggagacgaaagcagatttgaaccagttatattgctcaaagtgtgg
3118 atgttgctgcaaaacggtgttgcgagaacaaacattagccagaatctttgtgtgaggagtgacgaggattgacct
 220                                                        G  E  I  D  A  P  E  S  V  T
3193 gctgtgtgagccatcttgttaactttgtgtgtttccgaccaccagGGGGAGATTGACGCCCCGGAGAGCGTGACG
 230 G  M  L  S  V  M  E  S  L  T  E  K  Q  N  A  A  F  L  D  Y  K  G  K  T  L
3268 GGGATGCTTAGCGTGATGGAGTCCCTGACTGAGAAGCAAAATGCAGCATTCCTGGATTATAAAGGCAAAACCCTC
 255 P  W  ***
3343 CCCTGGTAAcaagcaaccgctttgtgtgtcactgcagtcaatcactagatgaacccaagtcaaataaactaatcc
3418 tgttacaacta
```

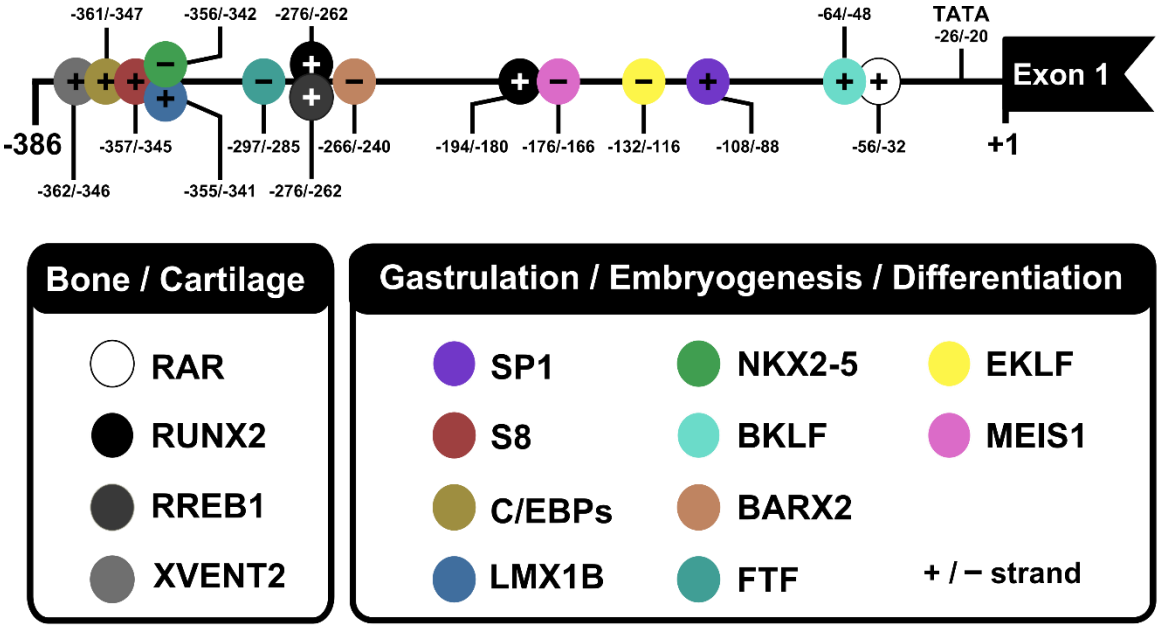
Gilthead seabream *sdr-like* gene. (**A**) PCR fragments obtained from GenomeWalker libraries. The structure of reconstructed gene is indicated below gDNA fragments. Black boxes indicate exons. (**B**) Gene sequence and structure. Canonical TATA box (tataaa) and polyadenylation signal (aaataaa) are underlined and coding sequence is highlighted in grey. Sequence of gilthead seabream SDR-like gene can be retrieved from GenBank using accession No. EU557023).

Supplementary Figure 3.3



Phylogenetic relationship between Sdr-like and members of closely related subfamilies (CBR1, CBR2, CBR3, CBR4, DHRS4, HSD11B1, DCXR, SNIFFER). Complete coding sequences were aligned using Translator X V1.1 (Abascal et al. 2010) with parameters set to default and molecular phylogeny (maximum likelihood tree; 1000 bootstrap replicates) was constructed from sequence alignment using MEGA version 6 (Tamura et al., 2013). Numbers represent branching probabilities. Tree was rooted with HTATIP2 as the outgroup.

Supplementary Figure 3.4



Schematic representation of gilthead seabream *sdr-like* gene promoter including transcription factors binding sites predicted using MatInspector library (matrix family library version 6.3 with all vertebrates as selected group, core similarity of 0,75 and matrix similarity optimized.

Supplementary Table 3.1

PCR primers used in this study.

	Name	Sequence (5'-3')*
cDNA/gene cloning	SDR-FW1	AGCTGTTTGCCTGTTGCAGAGACCCG
	SDR-FW2	ATGGCAGCGCAACCAGTCACCGT
	SDR-RV1	CCACCTGCTGGGAGCACTGTTTAATGC
	SDR-RV2	GCAGGGCCTCAGCCTTGGGTCCATTCG
	SDR-RV-P	CCCATCCACACCTTCACCGACTTCT
promoter constructs	SDR-RV-P2	CACGC**AAGCTT**TTCTTCTGCCTCTGAACCCTTTTA
	SDR-FW-P1	CCGGAG**CTCGAG**TGTTCTGTGCACAAGTCTTTAAAGTC
	SDR-FW-P2	CCGGAG**CTCGAG**ATGTTGGCTCCAGGGCAGAA
	SDR-FW-P3	CCGGAG**CTCGAG**GACAGGGTTGGAGAATGCC
qPCR	SDRreal-FW	GTGGAGAGGAGGGGGAGATTGACGC
	SDRreal-RV	TGTTACCAGGGGAGGGTTTTGCCTTTAT
	RPLreal-FW	AAGAGGAACACAACTCACTGCCCCAC
	RPLreal-RV	GCTTGCCTTTGCCCAGAACTTTGTAG

* Bold underlined sequence corresponds to cutting site for endonuclease *Xho*I in sense primer (FW) and *Hin*dIII in antisense primer (RV).

Supplementary Figure 3.5

Complete coding sequences (collected or reconstructed) for short-chain dehydrogenase/reductase (SDR) used to construct the phylogenetic tree presented in Figure 3.3. Aca, *Anolis carolinensis*; Ame, *Astyanax mexicanus*; Asu, *Ascaris suum*; Cmi, *Callorhinchus milii*; Dme, *Drosophila melanogaster*; Dre, *Danio rerio*; Fal, *Ficedula albicollis*; Gac, *Gasterosteus aculeatus*; Gga, *Gallus gallus*; Hbu, *Haplochromis burtoni*; Hsa, *Homo sapiens*; Lca, *Lates calcarifer*; Lcr, *Larimichthys crocea*; Ler, *Leucoraja erinacea*; Lme, *Latimeria menadoensis*; Loc, *Lepisosteus oculatus*; Man, *Misgurnus anguillicaudatus*; Mch, *Morone chrysops*; Mdo, *Monodelphis domestica*; Mmu, *Mus musculus*; Msa, *Morone saxatilis*; Ola, *Oryzias latipes*; Omy, *Oncorhynchus mykiss*; Oni, *Oreochromis niloticus*; Pma, *Petromyzon marinus*; San, *Sinocyclocheilus anophthalmus*; Sau, *Sparus aurata*; Sca, *Scyliorhinus canicula*; Ssa, *Salmo salar*; Tgu, *Taeniopygia guttata*; Tni, *Tetraodon nigroviridis*; Tor, *Thunnus orientalis*; Xla, *Xenopus laevis*; Xma, *Xiphophorus maculatus*.

>AcaSDR-like1 (green anole; Accession number XM_003229878)
ATGCCGGAGTCGGTGCTGGTGACGGGGGCCAACCGGGGCATCGGGCTGGAGCTGGTGAAGCAGCTGCTGGGCAAGGAGAGACC
CCCCAAGACCCTCTTTGCCACTTGCAGGGACCCCCAGGGACCAGGGGCCCAGGAGCTGAGGGCCTTGGCATCCAAACATCCCA
ACCTGGTGATTCTGAAGCTCGAAGCCACCAACCCAGCCAGCATCCAGTCTGCAGCCAAGGAAGCGAAGTCCCGCTTGGGGGAT
CAGGGCCTGAACCTGCTGATCAACAATGCCGGGGTGATGCCTGCCTGCACCTTGGAGTCGGCCACCGCCAAAGACATGCTGGA
CGTCTACAATATCAACCTAGTGGGGCCCATGCTGGTCACCAAGGCCTTCCTTCCTCTGCTGAAGCGGGCGGCCCAGGCCTCTC
CTCACAAGGCCCTCAGCTGTGGCAAGGCGGCCATCATCAACATGTCCACCATACTGTCTTCCATTGAGAAAACCCCGGAGATG
TACCACAGCTTCCCAGTCATCTCCTACCGCTGCAGCAAGGCAGCGCTGAACATGCTGACCAAGCTGCAGTCCATGGGCTACAA
GGAGGAGGGCATCCTCTGCACGGCAATCCACCCCGGGTGGGTCAAGACCGACTTGGGGACGGAGCAGGCGGACCTGACGGTGG
AGGAGAGCGTGGCCGGCATCCTGGAAGTGTTAAGCAGCCTCTCCGAGAAGCATCACGGGATCCTGGTCACTTGGGAAGGGAAA
ACCTTGCCCTGGTGA

>AcaSDR-like2 (green anole; Accession number ENSACAG00000028408)
ATGGCGCTGGGTTTCGGACCTCGGAGCGTCGTCCTCACGGGAGCCAATCGCGGGATCGGCTTGGAGCTGGTTCGGCAGCTTTT
GGGCGCCGCGGAGCCGCCGCAATGGATCTTTGCCTGCAGCCGGGAGCCCGACGGGGAGCGAGGAAAGGAGTTAAGGAAACTGG
CTTCCAGTCACCCAAATGTTATAATAGTGAAACTGGATGCTACAAATCCTGTGACCATTAAGGAAGCTGCCACTTGCATTGAG
GGGCACCTCAAAGGGGCAGGGCTGAATGTGCTCATAAACAATGCTGGCATTTTCACAGACAAGGAACTGGAAACCTTGGATCC
TGAAGACATGCTGAATGCCTACAAGACCAATGTGGTTGGGCCATTGTTAGTGAGCCAGGCTTTTCTGCCACTGTTAAAGAAGG
CAGCCAAGGAAAGCGCACAGGATGGACTCAGTTGCAGGAAAGCAGCCATCATCAACATTTCTTCTGTCTTAGGCTCCATTGAG
TCAGTCCCACTAAACTACAGCAAGCCTGCAGTCCCCTACCGCTGCAGCAAGGCTGCTCTAAATATGCTTACAAAGTGCCTCTC
TTTGAGCCTCAAAGAAGATAGGGTTCTGTGCACCGCAGTTCATCCTGGCTGGGTGAAGACAGACATGGGAACCCAAGAGGCCG
ACTTGAGTGTAGACGAGAGCGTGCGGGGGATTATCGGTGTGCTTTCCAAGCTTGGTGACACACACAATGGCGTTCTTGTGAAC
TGGAAAGGAAACAATGTTCCTTGGTGA

>AmeSDR (Mexican tetra; Accession number XM_007256232)
ATGAGTAGCGCGTTTAAGAGCTGCAGGAGTGTGCTGATAACTGGAGCTAACCGGGGGCTCGGTCTACAGATAGTAGAGAGCTT
AGTGAGTGGAAGCTTTTGTCCAGGGAAGATCATCGCCACGGCAAGGACTCCGGCCACCGCCAAGGAATTACAGAGCCTTGCAG
AAAAGCATTCCAATATACACATTATCCCCCTTGATGTCGTCAGTGAAGAGAGCATAGCGAAGGCTGCTGTGGAGGTGGATCAG
ATTGTGCAGGAGGATGGGCTCTGTTGCCTCATCAACAATGCAGGAATCAATGTTATCGCTAATTTTGAGACTGTAACGGCCGA
GAAGATGCTGGAGAACTTCCATACTAATTCTGTAGCTCCCCTCATGATCACCAAGGCTATGCTTCCAATGTTAAAGCGGGCAG
CAGCCAGAGAGACTGGCATGGGCATTCACAGAGCAGCAGTCATCAACATGACCTCTCTCCTGGGCTCAGTGGAGCTAAACTGG
GGTGAAGGTGCCAACAACTTCAAATGGTACCCATACAGAACATCAAAGAGTGCGCTGAACATGATCACTAGATGTATGGCGGT
GGACCTGGAGGCAGATGGAATCCTCTGCATGGCCATCCACCCTGGCTGGGTGCGCACTGATATGGGTGGGCAGCAGGCACCAC
TCAGCACAGAAGAAAGTATCTCCTCTGTCTTGGATGTGATTGGTGGATTAACAGAGAAGGATCACGGCTCATTTCTGCATTAC
ACCGGTGAACCTTTGCCCTGGTGA

>AsuSDR (round worm; Accession number AEUI02000020)
ATGACACCTGTGCCTCTCTCCGTGCTGATCACAGGGGCTAATCGAGGTATCGGACTTGGGCTTGTCAAACAGTGGGCATCACT
TGCTGGTGTTAAGCACATATTCGCATGCGCCAGAAAGCCGGATGCTGCCGAAGAGCTGAAGGACTTGTCAAGAAATGATGCAA
GAATTCACTGCGTTAAACTTGACGTAAATTCTGATAAATCGATTTTGGATGCGAAGAAACAAGTGGATTTGTTACTCGATGAG
GGTACCGGATTAAACGTGCTGATCAACAATGCAGGATGCATGATTGCTGAAGGTGGCACATTGGAGAACGCCGATAGAAGTGT
CTATTTACAGCACTTCGACACTAACGTGATCTCTGTTGCCAAGATGATAGAGGCATTCTTACCTCTGCTGAAAATGGCGTCGA
AAGAATGCATCTCGGACGAGTGGGGTGTTCATCGTGCAGCTATAATCAATATTAGTAGTGAATTGGGTTCCATTGAGAGCAAC
AATACTGGCTCAAAACTAATCAGAAGCATTCCATATAGATTATCCAAGGCTGCTCTAAATCAGCTTACAAAGACTCTTTCGGT
CGATCTTGCAGAGGATTCAATTTTGGTAGTGTCGGTTTGCCCAGGGTGGGTAAGGACGGATATGGGTGGACCGGAAGCTTCAC
TCACTGTGGAGCAATCCACGAAAACTCAAATTCAAACGGCGCTTTCACTTCGAAAAGAGCACTCGGGATTGCTGCTGTCGTAT
GATGGCACCAAACTTGCCTATTAA

APPENDIX

>CmiSDR-like (elephant shark; Accession number XM_007902227)
ATGGCTAACATCCTGCTTGGAGCTGCCCTGGTGACTGGAGCGAACCGTGGGATCGGACTGGAGCTGATCCGCCAGCTGGTGGA
CAGTGAGCGATCACCGCGGGTCCTGTTCGTTGGCTGCAGAGAGCCTGAGGGTCCCAGAGTCAGAGATCTCAAGAACATTGCAG
AAAAGCACCCCAACGTGATAGTGGTGAAGCTGGACGTGACCGACTCTCAGAGTATCGCAGAGTGTGTGGAGCAGGTGGAGAAG
GTGCTGGAGAAGGGTGGGCTGAACCTGTTGATTAATGATGCGGGGATAGCCACCTGCGACACACTGGAAACCCTGACTGCCGA
AATCATGGAGCAGACGTTCGCAACCAATATAGTGGCACCAATAATGATGGCCAAGGCTTTCGTGCCAACTCTGAAGCGAGCGG
CTGCGCTGAGCAACTTCAAGGGTCTGAGTTGCAGCAAAGCAGCTGTGATCAACATGTCCTCCATCTTGGGCTCCCTGGAGTTG
AACATAGACGTGAAGGGCCCCGATGATGCGTACTCCTACAAGACCAGCAAGGCTGCTGTGAACATGTTCACCCAGTGTCTGGC
TAAACAACTGAACCCTGACGGAATCCTTTATGCTGCGCTTCACCCCAGCTGGGTCCAAACGGATATGGGTGGGAATGAGGCAC
CGTTATCCGTAGAGGATAGTATCTGTGGCCTCATTCAAGTGATGACAAGTCTTTCGGATAAAAGCGAGCATTTCCTGGACTGG
AAAGGAGACCACTTGCCCTGGTGA

>DmeSNIFFER (fruit fly; Accession number NM_132238)
ATGAACTCCATCCTGATAACCGGCTGCAATCGAGGATTGGGTCTGGGCCTGGTCAAGGCGCTGCTCAATCTTCCCCAGCCGCC
GCAGCATCTATTTACCACCTGCCGGAATCGCGAGCAGGCAAAGGAGCTGGAGGATCTAGCCAAGAACCACTCGAACATACACA
TACTTGAGATTGATTTGAGAAATTTCGATGCCTATGACAAGCTAGTCGCCGACATCGAGGGCGTGACCAAGGACCAAGGCCTC
AATGTGCTCTTCAACAATGCCGGCATAGCGCCCAAATCGGCCAGGATAACGGCCGTTCGATCGCAGGAGCTGCTCGACACCTT
GCAGACCAACACGGTTGTGCCCATCATGCTGGCCAAGGCGTGTCTGCCGCTCCTTAAGAAGGCAGCCAAAGCGAACGAATCCC
AGCCGATGGGCGTGGGCCGTGCCGCCATTATTAACATGTCCTCGATCCTTGGCTCCATCCAGGGCAACACGGACGGCGGAATG
TACGCCTATCGCACCTCTAAGTCGGCCTTGAATGCGGCCACCAAGTCGTTGAGCGTGGATCTGTATCCGCAACGCATCATGTG
CGTCAGTCTGCATCCTGGCTGGGTGAAAACCGACATGGGTGGCTCCAGTGCCCCCTTGGACGTGCCCACCAGCACGGGACAAA
TTGTGCAGACCATCAGCAAGCTGGGCGAGAAACAGAACGGCGGTTTTGTCAACTACGACGGCACTCCGCTGGCCTGGTAA

>DreCBR1 (zebrafish; Accession number BC054914)
ATGAGTCAGTGCAAAGTAGCTCTGGTCACTGGTGCAAATAAAGGCATCGGCTTCGCGATCGTGCGTGCACTGTGTAAAGAGTA
CACCGGGGACGTGTATCTGTCATCCCGGGATGTCGGACGAGGAACCGCAGCGGTGGACAGCTTGAAGAAGGAGGGACTGCACC
CGCTTTTCCATCAGCTGGACATCAACGACCCAAACAGTGTCCGAACTGCACGCGACTTCTTCCAGGAGAAGTACGGGGGCCTG
GACGTGCTCATCAACAATGCTGGAATCGCCTTTAAAATGGCCGACACAACCCCCTTTGGGACCCAAGCTGATGTGACACTGAA
GACCAATTTCTTTGCCACCAGAGACATGTGTAATGTGTTTTTACCCATTATCAAACCAGGAGGTCGATTAGTAAATGTTTCCA
GCGGAATGGGCTCAATGGCACTAGGCCGCTGCAGTCCAGAACTCCAGGCCCGATTCCGAAGCGACGATATCACAGAGGAAGAG
TTAAACGGACTGATGGAGCGATTCGTCCGTGAGGCTCAGGAAGGCGTTCATTCGGAACGAGGATGGCCCTCAACTGCATATGG
CATCTCCAAAACTGGCCTCACCACCCTGACTCGAATTCAGGCCCGAAATCTGACCAAAGAAAGGCCTGGGGATGGGATCCTGT
GCAACGCCTGCTGTCCAGGATGGGTCAGGACTGACATGGCTGGACCAAATGCCACCAAATCACCAGATGAGGGCGCCATCACT
CCAGTGTATTTGGCGTTGCTTCCTGCAGGAGCAAAAGAGCCTCATGGCCAGTTTGTTTCAGAGATGAAAGTGCAGCCTTGGTA
A

>DreSDR (zebrafish; Accession number NM_001017792)
ATGAAGGGTAACTTCGCTAAGTGCGGATCGGTGCTCATCACTGGAGCGAGCAGGGGTCTCGGTCTGCAGATGGTCAAACAGCT
GCTCGCCACCCCAGAAAGACCCCAAAAGATCATAGCAACCGTCAGGAATCCAGCCGCTGCGGAGGAACTACAAAAACTCGCCA
AAGCTCATCCAGACGTTCATATCGTGACCCTAGATATCTCCAATGAAACGAGTGTGAACGCTGCATCTCAAGCAGTTGAAGCC
ATTGTTGGGGCCAATGGTCTGAACTGCTTAATTAACAATGCGGCGATTGGTATGAGCTCTGACCTGGACAGCGTTACTCGAGA
CGTCATGATGAAGACCTACGAGAGCAACACGGTTTCTCCTCTCTTTGTGACCAAGGCTCTCCTGCCGTTGTTGAGGAGAGCGG
CGGCTGAGGGGAGTGGTATGAGCATCCAGAGGGCCGCAGTGGTCAACGTTTCGTCTCTTCTGGGATCCGTGCAGCTCAACTGG
GGCGAGGGCGCGAGCTTTAAGAGCTACGCTTACAGGGCGTCCAAGAGTGCGCTGAATATGGTAACAAGGTGTTTGGCTGCTGA
TCTGGAAGCTGACGGCATCTTGTGCGTGGCTCTTCACCCCGGTTGGGTGCGCACTGATATGGGTGGACCAATGGCTCCTTTGA
GTCCAGAGGAGAGCATTTCATCCGTGCTGTCCGTCATTGCTGGATTAAAGGAGGAACATCACGGTGGATACGTGGACTACACT
GGGAAAAACCTACCCTGGTGA

>DreSDR-like (zebrafish; Accession number NM_001083559)
ATGGCAGCAGTTAAAGCATGCAAAATCCTGATCACAGGAGCCAACCGGGGCTTGGGCTTAGAAATGGTGAAGCAGCTTTCTGA
GAATTCTTGTCCAAAGCATATTTTTGCTACATGCCGTGACCCAGATGGGCCAAAATCTGCGGCACTAAGAGAGCTGGCGAAAA
AGCATCCCAATTTGATTACAATCATCCGTCTTGATGCTGATGATCCATGCAGTATCAAGGAGTCTGCCAAAAAAGTGGGTTCT
TTAGTGGGAGCAAATGGTTTGAACCTACTTGTAAATAACGCTGCAATCGTGGCAAATGGCACCATTCAGACCAGCAGTGTTGA
GGACCTGAAAAACACCTTCAACACCAATGTGATAGGACCCTTATTGATTATTAGGGAGTACAGACCATATCTACAAATAGCAG
CCAAAGCCAGTGGGACACCAGGCATGTCCAGTAAAAAAGCAGCGATCATCAATATCTCCACCGTGGCGGCTTCCATGACCAGA
ATGCCTCCCATATACAGCCACTTCCAAACCTTGCCATATGCTGTGTCTAAGGCAGGATTTAACATGCTGACTGTTTTGGCTGC
TGAGGAGGTAAAGACAGATGAGATCCTCTGCATGGCACTTCATCCAGGATGGGTGAAGACTGATCTGGGTGGACGAGATGCGA
CACTTGAACCAAATGAAAGTGTGGAGGGGATGCTGAAAGTCATCGGTGGTCTTACTGAGAAGCAGCATGGTGGATTTCTGGAC
TACACTGGAGCAACTGTGACCTGGTAA

>FalSDR-like (collared flycatcher; Accession number XM_005052439)
ATGGCGGCGGCGCGCACCGTGCTGCTGACCGGCTCCAACCGCGGCATCGGCCTGGAGCTGGTGAAGCAGCTGCTGGGCTCGCC
GCGACCCCCCGCCTGGATCTTCGCCACCTGCCGGGACCCCGAGGGGCCGCGGGCACAGGAGCTGAGAGATCTGGCATCCAAAC
ACCCAAACCTGGTTCTTGTAAAGCTGGATGTGGAAAACCCCTCGGCTATCACCGATGCGGCCAAGGTGGTGGAGGGGAAGCTG
GACGGAATGGGGCTGAACCTGCTGATAAACAACGCCGGCGTCTACACCCCCACGGCCTCGCTGGAGACGGTCGATGCTGAGGA
CATGGTCAGGTCCTACAAGACCAATGCAGTGGGGCCAATGCTGATGGCCCAGGCCTTCCTGCCTCTGCTGAAGAAGGCTGCCC
AGGACAGCAAAGAAAAGGGCCTGAGTTGCAGCAAGGCAGCCATCATCAACATCTCCACCATCTTGGGGTCCATCAAGAAAACA
CCTGATTCCTTCTTCCACCCTGTCATCTCCTACCGCTGCAGCAAGGCTGCCCTCAACATGCTGACCATGTGCCAGGCTCTGAC

CTACAAGGAAGCTGGGATCCTGTGTGTGGCACTGCACCCTGGCTGGGTGAAAACAGACATGGGCACCCAGGAGGCTGACCTGA
CAGTGGACACAAGTGTGCGGGGGCTGTTGTCTGTGCTGCCAATTCTTTCTGAGAAACACAGTGGGGCTCTGCTCAACTGGGAA
GGTAAAGCTATCCCCTGGTGA

>GacCBR1 (three-spined stickleback; Accession number DW611829)
ATGTCTATCAAAGTCGCCGTGGTAACGGGCAGCAACAAGGGCATCGGCCTGGCCATCGTCCGAGCGCTCTGCAAGCAGTACCA
AGGAGACGTCTACATCACCGCCAGAGACGTCGGCCGTGGTCAGGAAGCGGTGAAGTCTCTGGCCTCCGAAGGGCTGAAGACAC
AGTTCCACCAGCTGGACATCAACGACCAGGAGAGCATCACCACCGCTGCTGCTTACTTCAAAGAGAAGTACGGAGGGGTGGAC
GTCCTCATAAACAACGCTGGGATAGCGTTCAAAGCGGCGGACACGGCTCCGTTTTCGGTCCAGGCGGAGGTCACCCTGAAGAC
CAACTTCTTTGCCACCAGAGACATGTTGACTCACTTCCTGCCGCTCATCAAGGCCGGAGGGCGCGTGGTGAACGTCTCCAGCT
TCGTCGGCGCCCGCACTTTGAACAACTGCGACGCCGCCCTGCAGCAGCGCTTCCGCAGCGAGGACATCACCGAGGAGGAGCTG
TCGGGACTGATGGAGCAGTTCGTCGCCCAGGCCAAGAAGGGCGAGCACAAGAAGGGCGGCTGGCCCGACTCGGCGTACGGGAC
GTCCAAGACCGGACTCACGGTCCTGTCCATGATCCATGCTCGCCGTCTGTCCAAGGAGAGACCAAGCGACGGGATCCTGGTGA
ACGCCTGCTGCCCCGGGTGGGTGCGCACCGACATGGCCGGCGACAAAGCCCCCAAGTCTCCGGACGAGGGCGCCATCACCCCC
GTCTACCTGGCCCTGCTGCCGCCCGGAGCCACGGAGCCTCACGGGAAGTTCGTCTCCGATAAAGAAGTTCAGCCGTGG

>GacSDR-like (three-spined stickleback; Accession number BT026731)
ATGGCAACCAAAGCGGTCAGCGTGCTCATCACCGGAGCCAACAGAGGCCTCGGCCTGGAAATGGTAAAGCAACTGGTGGAGGC
GTCCCGTGACGTGAGGAAGCTGTTCGCCTGCTGCAGGGACCCAGATGGACCACGAGCCGAGCAGGCCTTGCAAACACTGGCAA
AGAAGCATCCCGACGTCATCGTTGTCGTCCGTCTGGGTACAAGCTGTCTGCCTTCTGTTGACGCAACTGACCTCTGTAGCATC
AAACAGTGTGCACAGCAGGTGGGCTCCGTGGTCGGGACGGGGGGCCTCAACCTGCTGATCAACAACGCGGCCGTGGCGGTTCG
AACCACCATGCAGGAAACCACTCCTGAGGAGCTGCAGCACATGTTCAGCACCAACGTCATGGGCCCCATGAACATCATCAAAG
AGTTCCTGCCTCACCTCCGTGCTGCGGCGAAGGCCAGCGGGACGCCCGGGATGTCCAGCTGCAAGGCAGCCATCGTCAACATC
TCCTCCATCCTCGGTTCAATCGAGGGCGCAAAGGAGTCTTACGCCACCTACCCTTGTTTCTCCTACCGCATCAGCAAAGCGGC
TTTGAACATGCTGACCCTGTGCGCCTCAGAGGAACTGAAGAAGGATGAGATCCTGCTGTCCGTGCTGCACCCTGGGTGGGTGC
GCACAGACATGGGCGGAGAAGAGGGGGAGATCGACGCTACGGAGAGTGTGCGGGGGCTGCTCGGGGTGATGGCCTCTCTCACG
GAGAAGCAGAACGGAGGCTTCCTGGACTGGAAGGGCAGGTCCATCCCCTGGTAG

>GgaSDR (chicken; Accession number XM_001233573)
ATGGCCCAGCCGCGCTGCCGCTCCGTGCTCATCACCGGCTGCAGCCGCGGCATCGGGCTGGGCTTGGTGAGGGGTCTCGCTGC
CTCCGATCCCCCCCCGGAGGTCGTCTTCGCCACGTGCCGCTACCCCGATAAGGCGCAGGAGCTCCAGCAGCTTAGCAAGCAAT
ACAGCAATATCAAGCTCCTCCAGCTGGATGTGGTCTGTGAAAACAGCATCAAAAAAGTAGTCAAAGAAGTGGAAGAAATCGTG
GGAGACAAAGGTCTGAACTGCCTCATTAACAACGCTGGCATCAATGTACTCGCCTCCTTGGAAGATGTCACAGCAGAGACAAT
GCTCACCATTTACGAGACCAACACAGTTGCCCAGCTGATGGTCACTAAGGCTTTTTTACCCCTTCTGAGGAAGGCAGCTCAGC
TGAGCACAGGAATGGGCTGCCATAGAGCTGCTATTATTAATATGTCCTCTCTTGCTGCCTCCATGCAACTCGTCCAGGCGAAT
GAGATGTTCCTCAAGGTCTACCCCTACAGGATAGCAAAGACTGCACTGAACATGATCACCAGATGTCTTGCTGCAGATTTGAA
ATCAGATGGAATTCTCTGTATTTCTCTGCATCCAGGCTGGGTTCAGACAGACATGGGCGGAAACATGGCTCCCTTGCAGGTGC
AAGAGGCCATCCCTGGTATTCTCTCTGTTCTGGACCGTCTTAGTGAAAAAGAAAATGGTTCCTTCCTGGACTGGCAAGGGGAA
ACTCTACCATGGTAG

>GgaSDR-like (chicken; Accession number NM_001277118)
ATGGCGGCTGTGCGGGCCCGCTGCGTGCTGCTCACCGGCTCCAACCGCGGCATCGGGCTGGAGTTGGTCAGGCAGCTGCTTGG
GGGGCCTCAGCCCCCCACCCACGTCTTTGCGACGTGCCGGGACCCCGAGGGGCCGCGGGGGAAGGAGCTGAGAGATCTGGCAT
CCAAACACCCAAACCTGGTGCTGGTGAAGCTGGATGTTGCCAATCCCTCAGCGATTGTCAGTGCTGCAAAGATTGTGGAGGGG
AAGCTGAACGGCGCCGGGCTGAACCTGCTGATAAACAACGCCGGCATCTTTACCCCCGTGTCACTGGAGACGGTGGACTCTGA
AGAGATGATACGGGCATACAAGACCAACGCGGTGGGACCGCTGCTGATGGCCCAGGCGTTCCTGCCCCTCTTGAAGAAGGCTG
CCCAGGACAGCAGGGAAGAAGGACTGAGTTGCAACAAGGCAGCCATCATCAACATCTCCACCCTCTTGGGGTCCATCGAGAGA
ACTCCTGAGTCCTATTTCAAGCCGGTCATCTCCTACCGCTGCAGCAAGGCTGCCCTCAACATGCTCACCAGGTGCCAGGCTCT
CACCTACGGCAAGGCTGGGATCCTCTGCGTGGCACTTCACCCTGGCTGGGTGAAAACTGACATGGGCACCCAGGAGGCCGACC
TGACAGTGGACACGAGTGTGCAGGGGCTGCTGTCGGTGCTGCCCATCCTCTCTGAGAAGCACAGCGGGATGTTGCTCAATTGG
AAAGGCAAAACAATTCCATGGTGA

>HbuSDR (Burton's mouthbreeder; Accession number GBBF01040636)
ATGAGTGCTGCTATGAATTTTAAAAAATGCGGCTCTGTGCTGATTACTGGAGCCAGCCGGGGCCTCGGGCTGCAGTTAGTGGA
CAGTCTGGCAAGTGGACAGTTTTCACCTGGAAAGATAATAGCCACGTGCAGAAACCCTGGTAACGCACAGGAACTTCAAGAAC
TGGCGGAGAAGCATCCCAACATCCACATAATCACTTTGGACGTAGTGAACCAGGAGAGCATAGAGAAGTCTGTAGAAGAGGTG
AGCAAGCTGGTGCAAGAAGAAGGTCTGAACTGCCTGATCAACAATGCGGGCATCAAAGTGGTGGCCGACTTTCATTCTGTTAC
CGCCGAGATGATGATAGAAAACTTCCACACCAACACTGTGGCTCCTCTAATGATCACTAAGGCTTATCTGCCTCTTCTAAAAA
AAGCTGCATCCAAAGGAGGGGCAGGAGGCTCAGCTACAATGGGAATCCAGAGGGCAGCGGTTATTAACATGACCTCACTGCTG
GGCTCTGTGGAGCTCGCTTGGGGTGAGAGGGCCAACAACTTCAAATGGTACCCCTACAGGACATCCAAGAGTGCACTAAACAT
GGTTAGTCGCTGCATGGCTGTAGACCTAGAGCCAGATGGGATTCTCTGTATGGCTATACACCCTGGCTGGGTACGCACTGACA
TGGGGGGGTCAGAGGCTCCACTGAGTCCCGAGGATAGTATTTCTTTCATGTTGTCTGTGATTGGAGGGCTAACTGAAAAGGAC
CATGGCTCATTTCTGAACTTTACAGGAGAGCAGATACCTTGGTGA

>HsaCBR1 (human; Accession number BC015640)
ATGTCGTCCGGCATCCATGTAGCGCTGGTGACTGGAGGCAACAAGGGCATCGGCTTGGCCATCGTGCGCGACCTGTGCCGGCT
GTTCTCGGGGGACGTGGTGCTCACGGCGCGGGACGTGACGCGGGGCCAGGCGGCCGTACAGCAGCTGCAGGCGGAGGGCCTGA
GCCCGCGCTTCCACCAGCTGGACATCGACGATCTGCAGAGCATCCGCGCCCTGCGCGACTTCCTGCGCAAGGAGTACGGGGGC

APPENDIX

CTGGACGTGCTGGTCAACAACGCGGGCATCGCCTTCAAGGTTGCTGATCCCACACCCTTTCATATTCAAGCTGAAGTGACGAT
GAAAACAAATTTCTTTGGTACCCGAGATGTGTGCACAGAATTACTCCCTCTAATAAAACCCCAAGGGAGAGTGGTGAACGTAT
CTAGCATCATGAGCGTCAGAGCCCTTAAAAGCTGCAGCCCAGAGCTGCAGCAGAAGTTCCGCAGTGAGACCATCACTGAGGAG
GAGCTGGTGGGGCTCATGAACAAGTTTGTGGAGGATACAAAGAAGGGAGTGCACCAGAAGGAGGGCTGGCCCAGCAGCGCATA
CGGGGTGACGAAGATTGGCGTCACCGTTCTGTCCAGGATCCACGCCAGGAAACTGAGTGAGCAGAGGAAAGGGGACAAGATCC
TCCTGAATGCCTGCTGCCCAGGGTGGGTGAGAACTGACATGGCGGGACCCAAGGCCACCAAGAGCCCAGAAGAAGGTGCAGAG
ACCCCTGTGTACTTGGCCCTTTTGCCCCCAGATGCTGAGGGTCCCCATGGACAATTTGTTTCAGAGAAGAGAGTTGAACAGTG
GTGA

>LcaSDR-like (Asian seabass; Accession number GAQL01020679)
ATGGCAGCCAGTCAACCAGTCAGCGTGCTAATCACTGGAGCGAACAGAGGCCTGGGCCTGGAGATGGTTAAGCAAATGGTGGA
GGGCCCCAATCCAGCGAAAAAACTGTTTGCCTGTTGCAGGGACCCAGACGGATCCAGGGCTGCGGCACTGCAAACACTGGCAA
AGAAGCATCCTAACATCATCTGCGTCGTCCGTCTGGACGCCACTGACCTTTGTAGCATAAAACAGGCTGCCCAGCAGGTGGGG
TCTCTGCTGGGGACGGGGGGCCTCAACCTGCTGATTAACAACGCAGGGTTCCTGGCCAAAGCCACCCTGCAGGACACCACTCC
CGAGGACATGCAGGGCTCCTTCAACACCAACGTCATGGGCCCTATGAATATCATTAAAGAGTTCCTGCCCCACCTCCGCACGG
CAGCGAAGGCCAGCGAAGTGCCAGGGATGTCCAGCAGCAAAGCAGCCGTCATCAGCATCTCCTCATTTCTGGGTTCATTTGAA
TCAGTGAAACACTCGTATTCCTTCATCCCTGCCGTCTCCTATCGCATCAGCAAGGCAGGTTTGAACATGCTGACGCTTTGTGC
TGCGGAGGAGCTGAAGAAGGACGGGATCCTGTTTTCAGTGCTGCACCCTGGCTGGGTGCGCACTGACATGGGTGGAGAGGAGG
GGGAGATTGATGCTCCAGAGAGTGTGCAGGGTATGCTCAGCGTGATGGCCTCTCTGACTGAGAAGCAGAATGGAGCCTTCCTT
GATTATAAGGGCAACACCATCGCCTGGTAA

>LcrSDR-like (Yellow croaker; Accession number JPYK01005734)
ATGGCAGCGAGACCAGTCAATGTGCTTATCACGGGAGCCAACAGGGGCTTGGGCCTGGAGATGGTTAAGCAAATACTGGAGGC
TCAACGTCCAGTGAAAAATCTATTCGCCTGTTGCAGGGACCCAGATGGACCAAGGGCTGAGGCCCTGCGAACGCTGGCAAAGA
AACATCCTAACATCTCCATCAGTCGTCTGGATGCTACTGACCTTTGTAGCATAAAAGAATGCTCCCAGAAGGTGGGCACTCTT
GTGGGGACGGAGGGTCTCAACCTGCTCATTAACAATGCTGGGTTCATGTCCAAAAGCACCCTGCAAGAGACCACTCCTGAGGA
CATGCAATACACCTTCAACACCAATGTCTTGGGCCCTATGAATGTCATAAAAGAGTTCCTGCCTCACCTTCGTGTTGGAGCAA
AGGCCAGTAAGATGCCAGGGATGTCCACCAGAAAAGCAGCCGTCGTCAACATCTCCTCAGTACTGGGTTCGATGCAAGCTGTA
AAAGAGTCATACAGCTACTTTTCTGCAGTCTCCTATCGCATCAGCAAAGCAGCCCTGAACATGCTGACATTGTGTGGTGCAGA
GGAGCTGAAGAAGGATGAGATCCTGTTTGCTTTGCTGCACCCTGGCTGGGTGCGCACCGACATGGGTGGAGAGGAGGGGGAGA
TTGATACTCCAGAGAGTGTGCAGGGGATGCTTCATGTGATGGCTTCCCTGACCGAGAAGCAGAATGGAGCCTTCCTAGATTAT
AAGGGGCAACCCATCCCCTGGTAA

>LerSDR-like (little skate; SkateBase accession number LS-transcriptB2-ctg93096)
ATGCCGATCGCGGTGCTGGTGACCGGAGCGAGTCGCGGCCTGGGCCTGGAGCTGATCCGCCAGCTGGGCCGCGGTGAGCGGCC
GCCCCGTTACCTGTTCGCTGCCTGCAGAGAGCCCGAGGGAACCGGAGCGACGGAACTTAAGAACTTTGCAAAATGTTTCTCAA
ATGTGAAAATAGTTAAGATGGATGTGGATGACCTTAACAGCATTCAAGAGACTGCCGAACAGATTCAGAAGGTCTTGAAAGAC
AATGGTCTAAATCTGTTGATAAATAATGCTGGGATCAACTTTGGTGGAGATCTGAATGGAATTACTCCCGAAATAATGATGAA
GACTTACAAGACCAATGTCATTGCACCAGTGATGGTTACCAAGGCTTTGTTGCCAAGTTTGATGCAAGCAGCTCAGCTGAGTG
GTGAACATGGCTTCAGCAGTGAGAAAGCAGCAGTTATCAACATGTCCTCTATCATGGCCTCCATGGAGATGTTCAACATTCAA
AATGGGAAAGCGTATTCCTACAGAGTCAGCAAGGCTGCACTGAACATGGTCACTAAATGCTTAGCCAGTGAGCTGAAACCCCA
TGGAATCCTTTGTGCATCAGTGCACCCAGGCTGGGTCAAAACAGACATGGGTGGAGAAAAGGCATCATTATTCAAAGAAGAAA
GTGTGCAGGGGATTCTCAAGGTACTTGCAAGGCTTTCAGAACAAGACAATGGATTGTTTTTAGACTGGAGAGGTCAACGTTTA
CCTTGGTGA

>LmeSDR (Indonesian coelacanth; Accession number GAPS01001932)
ATGGGCATTCTAAGATGCAAGTCGATATTGATTACTGGAGCAAGTCGAGGTATTGGGTTTGAACTGGCTCGACAGCTAGCATC
AGAAAGCCCGTCTGTAGAGAAGGTGATTGCAACCTGCCGGGAGCCAGAGAAGGCAAAGGAGCTTCAAGAATTAACTGGAAAAT
ATACTAAACTGAAAATTGTTCAGCTTGATATTGTCAGTAAAGAAAGCATAGAAAAAGCTGTAAAGGAGGTGGAAGCTTTGGTT
GGATCTAATGGACTGAACTGTCTCCTTAACAATGCTGGCGAGATGATTTTTGCGAACTTAGAAACTGTTACTGCTGAAGCAAT
GATGCGAATGTATGAAACAAATACAGTGGCTCACCTAATGGTTAGCAAGGCATTCCTGCCCCTCCTGAAAAAAGCAGCCCAGT
CAGGTTCTGGAATGGGTTGCCATAGAGCAGCAATCATGAACACCTCTTCAATTACAAGCTGTTTAAGCTTAGTCAATGATGCC
GAACAAAAAATATTTCTGCAAGTCTATCCTTACAGAATATCCAAGACAGCTCTAAACATGGTCACTAGATGCATTGCTGCTGA
TTTGAAGTCGGAGAACATTTTGTGCGTAGCTATGAATCCTGGCTGGGTGCAGACTAGCATGGGCGGCCCTCAGGCACCTTTAA
CCGCAAAAGACAGTGCCTCGGGGATTTTATCTGTTCTTGCCAACCTCGGTGAAAAGGATACTGGGAACTTTTTTGACTGGAAA
GGAGAGATCGTGCCCTGGTAG

>LocSDR (spotted gar; Accession number XM_006642467)
ATGGGCGCAGAGTTTCAGAGCTGCCAAACGGTGCTCATTACCGGAGCCAGTCGCGGCCTCGGCTTGCAGATGGTGAAACAGCT
GGTAACAGGGAACCACAGACCGCTGAAAATAATCGCAACTGCCCGGGAGCCGGCAGCAGCCAAGGAGCTAAAGGAACTGTCTG
AAAAATACAAAGGTCTACATGTGGTACCTCTTGATGTTGTCAGCCAAGGGAGCATAGAGAAAGCCTTGCAAGAGGTGGGGTCT
CTGGTTGGAGACGAAGGGCTGAACTGCTTGATTAACAATGCTGGGATTATGATCTCTGGAGACATTGAGACGGTGACGGCTGA
AGGCATGATGAAAAACTTCGAAACTAATACTGTGTCTCCTCTCATGATCACCAAGGCATTTTTGCCCCTGTTAAGAAAAGCTG
CAGCAAAGAGCACAGGAATGGGCATCCACAGAGCAGCCATCATCAACATCTCCTCCATGCTGGGATCAATTGAACTGAACTGC
CAGCCGGTGACAGGGTTGAAGTGTTGCCCTTACAGAACATCAAAGGCGGCTCTGAACATGGTGACTCGCTGTCTGGCTGTTGA
TTTAGAGGCAGACGGTATCCTGTGCACAGCCATCCACCCAGGCTGGGTCCGCACTGATATGGGGGGGGCAGAGGCTCCTTTGA
GTAAAGAAGACAGCATTTCTTCAGTTCTCAAAGTGATTGCCAGCTTATCCGAGAAGGAAAATGGTGGATTTTTGAGCTATGAT
GGAAAGCCCCTTCCTTGGTAA

>ManSDR-like (Pond loach; Accession number GAAD01004883)
ATGGCCACATTAAAGTCATGCAATGCTTTGGTCACAGGAGCCAACCGGGGATTAGGTTTGGAAATGGTTAAGCAACTTCTTGA
AGCCCACTGCTCTAAAATATTTGCTGGATGTCGGGATCCAGATGGGCCAAAGTCTGAGTCACTGAGAGAACTGGCCAATAAAC
ATCCGAATATTTTGACTATTGTGCGTCTAGATGTCGCTGATCCATGCAGTATCAAAGAGTCCGTCGAGAAAGTATCGTCTTTA
CTGGGGAATAGCGGCTTAAATCTGCTGGTGAATAATGCAGCCATACTAAAACAAAAAAGCATGCTGACCGCCACAGTGGAGGA
CATGCAGGACAACTTTAACACCAATGTCATAGGACCTATGTTAGTCATTAGGGAGTACTTGCCATACCTGCGTGCAGCAGCTA
AAGCCAGTGGTAAAGCAGGGTTGTCATGTGATAAAGCAGCTATAATCAACTTCTCCGATTCAGATACATCCTCCATGAGCATT
GTGCCAACAATAGAGTCACCATTTCCCCTCTTTCCTTACAGCATCAGCAAGGCAGCTCTTAACATGTTAACAGTGTGTTCTGC
TATGGAGTTGAAAGAAGATGAGATCCTCTGCGTTTCAATTCATCCTGGATGGGTGAAGACTGACATGGGGGGAGATAAGGCTC
CTCTTGAACCGAGAGAAAGTGTGGAGGGGATTCTGCGTGTCATTAGCGGTCTCAATGACAAACAGAATGGTTTATTCATTGAC
TACACTGGAAAAATAATGCCCTGG

>MchSDR (white bass; Accession number GAZY01115445)
ATGAGTGGTGCTGTGAATTTCAAGAACTGCGGCTCTGTGCTGGTAACGGGAGCCAGCCGAGGGCTCGGGCTGCAGATAGTCGA
CAGCCTGGCGAGTGGAGGTTTCTCACCCGGTAAGATTATAGCCACGACCAGAGAGCCGGCGAGAGCACAGAAACTTCAGGAAC
TGGCACAGAAACATCCCAACATCCACATAGTCACACTGGATGTAGTGAACCAGGCGAGTATAGAGGGGTCTGTAGAAGAGGTG
GGGCAGCTGCTACAACAACAGGGTCTGAACTGCCTGATCAACAATGCAGGGATCAATGTGGTGGCCGACTTCCATACGGTCAC
TGCAGAGATGATGATAGAAAACTTCCATACTAATGCTGTGGCTCCTCTAATGATCACCAAGGCCTTCCTGCCTCTGTTGAAGC
AAGCTGCGTCCAGAGGAGGAGCAGGTGGCACAGGGAGCATGGGCATCCAGAGGGCAGCAGTCATCAACGTGACCTCTCTGCTG
GGCTCTGTGGAGCTCAACTGGGGGGAACGGGCCAACAACTTCAAATGGTACCCCTACAGGACATCCAAGAGTGCCCTAAACAT
GGTGAGTCGCTCTATGGCTGTAGACCTGGAGCCTCATGGGATTCTCTGCATGGCTATACACCCTGGCTGGGTCCGCACTGACA
TGGGAGGGTCGCAGGCTCCACTGAGTCCAGAGGAGAGCGTTTCTTCCGTCTTGTCTGTGATTGGTGGACTGACTGAAAAGGAT
AATGGGTCATTTCTGAACTTTACAGGAGAGGTGTTGCCTTGGTGA

>MdoSDR (gray short-tailed opossum; Accession number XM_001364916)
ATGGCACTTGGGACCTGCCGGACTGTCCTCATCACTGGCTGCAGCCGCGGCATCGGCCTGGGTCTGGTGCAGGAGTTCTTGAC
CCGGAGCCCCACTCCGGACCTGGTCATCGCCACCTGCCGACATCCCGACCAAGCACAGGAACTCCAGCAGCTAAAGGAGCATC
ATCAACACCTGAGAATCCTACAGCTGGACATCCTGTGTGAAGACAGCATCAAGAAGGTGGTGCAGGAAGTGGAGGTACTGGTC
GGTGAACAGGGTCTTAACTGCCTCATCAACAATGCGGGAATTAACAGACTTGCCACCTTGGAGAGTGTCACAGCCAAGGATAT
GCTGACCATCTATGAGACCAACACGGTGGCCCAGCTCATGGTCAGCAAGGCATTTCTCCCCCTCCTGAGGAAGGCTGCCTGCC
AAGATGCTGTGATGGGCATCCATCGGGCTGCCATTATCAACGTGTCATCCCAAACTGCCTCCATGCAGCTCTTCTTCCAGAAT
AAGAAATCCCAGGAAGTGTATCCTTACCGGATATCCAAGACGGCCCTCAATATGATCACCTGCTGCCTTGCTGCTGACCTGAT
ATCAGATGGAATCCTCTGCATTTCTCTGCATCCGGGCTGGATCAAGACCGACATGGGGGGGAGCAAGGCTCCCCTACAATTAA
AGGAGGCCATCCCAGCGATTCTCAATGTCTTGGCCAATCTCAGAGAAAAGGATCATGGGGCCTTCCTAGACTGGAAAGGAGAA
GTGATCCCTTGGTAG

>MmuCBR1 (Mus musculus; Accession number NM_007620)
ATGTCTTCCAGCAGACCCGTGGCGCTGGTGACCGGTGCTAACAAAGGAATCGGATTCGCGATCACTCGTGACCTGTGTCGGAA
ATTCTCCGGGGACGTGGTGCTCGCGGCGCGGGACGAGGAGCGGGGCCAAACGGCAGTGCAAAAGCTGCAGGCGGAGGGCCTGA
GCCCACGCTTCCACCAGCTGGACATCGACAACCCGCAGAGCATTCGCGCACTGCGCGACTTTCTGCTCAAGGAATACGGAGGC
CTGGACGTGCTGGTCAACAACGCAGGCATCGCCTTCAAGGTCAATGACGACACCCCCTTCCACATTCAAGCAGAGGTGACAAT
GAAAACGAACTTTTTTGGTACCCGAGATGTCTGCAAGGAGCTGCTCCCTCTAATAAAACCCCAAGGCAGAGTGGTGAATGTGT
CCAGCATGGTGAGTCTCAGGGCCCTGAAAAACTGCAGGCTGGAGCTGCAGCAGAAGTTTCGAAGCGAGACCATCACAGAGGAG
GAGCTGGTGGGGCTCATGAACAAGTTTGTGGAAGATACAAAGAAAGGAGTCCATGCGGAAGAAGGTTGGCCTAATAGTGCATA
TGGGGTCACCAAGATTGGGGTGACAGTCCTGTCCAGAATCCTTGCCAGGAAACTTAATGAGCAGAGGAGAGGGGACAAGATCC
TTCTGAATGCCTGCTGCCCTGGGTGGGTCAGAACCGACATGGCAGGACCAAAAGCCACCAAAAGCCCAGAAGAAGGAGCAGAG
ACCCCTGTGTACTTGGCCCTTTTGCCTCCAGATGCAGAGGGGCCTCATGGGCAGTTTGTTCAAGATAAAAAAGTTGAACCATG
GTGA

>MsaSDR-like (striped bass; Accession number GBAA01172618)
ATGGCAGGGCAGGCAGTCAACGTGCTTATCACCGGAGCCAACAGGGGCTTGGGCCTGGAGATGGTTAAGCAGATGCTGGAGAC
CTACCGTCCAGTGAAAAAGCTGTTTGCTTGTTGCAGGGACCCAGATGGACCCAGGGCTGAGGCCTTGCGAACACTGGCAAAGA
AGCATCCTGATGTCATCTCCATCGTCCGCCTGGACGCCGCTGACCTTTGTAGTATAAAACAGTGCGCACAGCAGGTGGGCGCT
CTGGTGGGGACAGGCGGCCTCAACCTGCTGATTAACAATGCAGGGCTCCTGTGCAAAAGCGCCCTGCAGGAGACCAGTCCTGA
GGAGATGCAGAACAGCTTCACCACCAATGTCGTGGGCCCTTTGAACATCATTCAAGAGTTCCTGCCTCACCTTCGCACTGCAG
CAAAGGGCAGTAAGATGCCAGGAATGTCCTGCAGGAAAGCAGCAGTCGTCAGCATCTCCTCAGTTCTGGCTTCAATGCAATGT
GTGAAAGATTCATATTCCATTTTCTCTGCAATCCCCTACCGTATCAGCAAGGCAGGTCTGAACATGCTGACAGTGTGTGCTGC
AGAGGAGCTGAAGAAGGACGAGATCCTGTTTTCTGTGCTGCACCCGGGCTGGGTTCGCACCGACATGGGTGGAGAGGAGGCGG
AAATCGATGCTCCAGAGAGTGTGCAGGGGATGCTTCGTGTGATGACTTCCCTGACTGAGAAGCAGAATGCAGGCTTCCTGGAT
TATAAGGGCCATACCGTCCCCTGGTAG

>OlaSDR (Japanese medaka; Accession number XM_004086210)
ATGTCTGCGTGGGGGTCGGTGCTGATCACCGGCTGCAGCAGGGGCATCGGCCTGCAGCTCGTCAAGCAGCTGTCTGAGAGCAG
CAGCAGGCCGGCGAACATCATCGCCACGGCCCGGAACCCGGCCGGCTCCGCGGCTCTACAAGAACTCTCCAGAACATGCCCAG
GTGTTCACGCCGTCACCCTGGATGTTTGCAGTGAGCAAAGCATCTTCTCAGCTGTGGAAGAGGTTCAGTTTATTGTGGGAAAC
CAAGGACTCAACTGTCTGATCAACAATGCTGCTGTTGGATCTTCCAGCAGCATCAACAGTGTAACCCGTGAGGCCATGATGAG
GGCTTTTCAAGTCAACAGTGTTGCCCCTCTCTTTGTTACTAAGGCATTTCTGCCACTGTTGCAGACGGCTGCAGCTCGCTCCA

APPENDIX

CTGGGATGGGGGTCCATAGAGCGGCTGTCGTCAACATCTCCTCCATCCTCGGTTCAATCTCATTGAACTGTGGAGACAGCGCC
ACCTTCAAAAGCTATGCTTATGGCACCTCCAAGGCTGCTCTAAACATGGTGAGCCGCTGTCTGGCGGCTGATCTGGGACCTGT
TGGGATCCTGTGTGTGTCTCTACATCCTGGCTGGGTGAAGACAGACATGGGGGGGCCTCATGCCGACTTAACCGTAGAAGAAA
GTGTGTCTGAAATTCTATCTACTCTCGTCAGTTTCACGGAACAGGACCATGGCGGTTTCAAGGATTACCGTGGACAAACGTTA
CCATGGTAA

>OmyCBR1 (rainbow trout; Accession number NM_001124596)
ATGTCCAAAAAAGTGGCAGTAGTTACCGGTGCCAATAAAGGCATAGGACTTGCGATTGTGAGGGAGCTTTGTAAGGCAAAATT
TACCGGGGATGTTATTCTTACTGCTCGAAATGAGAAACTTGGAAATGAGGCAGTGAAGATGCTGAAGTCCGAAGGATTTGAAG
TTTCTTACCACCACCTTGATATCTGCGACCAGGGCAGCGCCAAGCAACTGAGTAACTTTCTGCAGAAGACGTATGGCGGATTG
GATGTGCTCATTAACAACGCTGGAATGGCTTTTAAGAATGATGCGACTGAGACTTTTGGGGAACAGGCTGAGGTGACCATGCG
CACCAACTTTTGGGGCACCCTGTGGGTGTGCCATGCTCTCCTACCCCTCCTCAGACCAAATGCCAGAGTGGTGAATGTCTCCA
GCTTTGTTAGCAAGAAAGCTCTTGATACATGCAGCCCTCAACTACAAGCCAAGTTCCGTGATACTGAGCTCTCTGAGGAGGAG
CTGTGCTTGCTGATGGGGCAGTTTGTTATTGCCGCTCAGCAGGGAAACCATCAGGCCCAGGGGTGGCCAAACACAGCCTATGG
CACAACAAAGATCGGAGTGACTGTGCTGTCCAGGATTCAGGCTCATTATCTGACTAAGACCCGGGCAGCTGATGGAATCCTGC
TCAACGCCTGCTGCCCTGGCTGGGTCCGCACTGACATGGCAGGCTCCAAAGCCCCCAAGAGTCCTGAAGAAGGAGCACAGACT
CCTACCTATCTGGCACTTCTTCCTGAAGGGGCCAAGGAGCCACATGGACAGTTGGTGTGGGACAAGACCGTTCAGGAATGGTA
G

>OmySDR (rainbow trout; Accession number CCAF010028677)
ATGAGTAGTGCTTTTAAAAATTGCCGAACGGTAATGGTGACAGGAGCCAGCCGAGGCCTCGGTCTGCGAGTTGTGGAAAGTCT
GGTGGCAGAGAGCTCTGATCCGGACAAGATTATTATTGCCACGGCCAGAAACCCGAGCGGGGCCCAGGAACTCCAGAAACTAG
CTGAAAGACATCCAAATATCCATATAATAACACTTGATGTTGTCAGTCAGGAGAACATAGAGAAAGCTGCACAAGACGTTGAC
CTGCTTGTTCAGGAGCTGGGCCTCAATTGTCTAATCAACAATGCTGGAATCAATGTTGTCGCCAATTTTGAGACAGTCACGGC
TGAGAAGATGTTGGAGAACTTCCACACCAACTCTGTGGCCCCCCTTATGATCACTAAGGCTCTCCTACCCCTGCTGAAGAGGG
CAGCTGCCAAAGGCACGGGCATGGGCATCCATAGAGCTGCGGTCATCAACATGACTTCCTTGTTAGGGTCAGTCGAGCTCAAC
TGGGGTGACCGGGCCAAGAACTTCAAGTGGTATCCTTACAGAACATCCAAGAGTGCTCTGAACATGGTGACCCGATGCATGGC
AGTGGATCTGGAGGCTGATGGCATCCTGTGCATGGGCCTCCACCCTGGCTGGGTACGCACTGACATGGGCGGACCAGAGGCAC
CTTTGAGTCCAGAGGAGAGCATATCTTCTGTGTTGTCAGTAATTGGTGGATTGACTGAAAAGGATCATGGGTCATTTCTACAC
TACACCGGGGAGCCACTACCCTGGTGA

>OniSDR (Nile tilapia; Accession number GBBU01025214)
ATGAGTGCTGCTATGAATTTTAAAAAATGCGGCTCTGTGCTGATTACTGGAGCCAGCCGGGGCCTCGGGCTGCAGTTAGTGGA
CAGTCTGGCAAGTGGACAGTTTTCACCTGGAAAGATAATAGCCACGTGCAGAAACCCTGGTAACGCACAGAAACTTCAGGAAC
TGGCAGAGAAGCATCCCAACATCCACATAATCACTTTGGATGTAGTGAACCAGGAGAGCATAGAGAAGTCTGTAGAAGAGGTG
AGCAAGCTGGTGCAAGAAGAAGGTCTGAACTGCCTGATCAACAATGCGGGCATCAAAGTGGTGGCCGACTTTCATTCTGTTAC
CGCCGAGATGATGATAGAAAACTTCCACACCAACACTGTGGCTCCTCTAATGATCACTAAGGCTTATCTGCCTCTTCTAAAGA
AAGCTGCGTCCCGAGGAGGGGCAGGAGGCTCAGCTACAATGGGAATCCAGAGGGCAGCGGTTATTAACGTGACCTCACTCCTG
GGCTCTGTGGAGCTCGCCTGGGGTGAGAGAGCCAACAACTTCAAATGGTACCCCTACAGGACATCCAAGAGTGCCCTAAACAT
GGTTAGTCGCTGCATGGCTGTAGACCTAGAGCCAGATGGGATTCTCTGTATGGCTATACACCCCGGCTGGGTACGCACTGACA
TGGGGGGGTCAGAGGCTCCACTGAGTCCCGAGGATAGTATTTCTTCCATGTTGTCTGTGATTGGAGGGCTAACTGAAAAGGAC
CATGGCTCATTTCTGAACTTTACAGGAGAGCAGATACCTTGGTGA

>PmaSDR-like (sea lamprey; Accession numbers AEFG01028136 and AEFG01028135)
ATGAACGGGCTCGCCAACCGCGAGCTGGTGCTCAACTCGGTGCTGGTGACGGGTGCGAACCGCGGCCTGGGACTCGGCATCGT
TCATCGCCTCCTCCAGCAGCAGCAGCAGGCGTCTCCACGCTTCATCTTCGCCACCTGCCGATGCACCGATGAGCAGAGCGCGA
AGGAATTGCGGAATTTGGCGGAACACCACGCAAGCCTGCTGGTGGTGGAAATGGACATCACAGACCGCGCGAGTGTGGAGCGG
GCGGTGAGGGTGGTGGATGAGGCTGTCGGGGAAGAAGGCCTCAACCTTCTGGTGAACAACGCCGGGACTCTACAAACGGGCTC
CCTCGGTTCACTGTCGGCCGACTCCATGCTGCAGTGCTATCATACCAACGCTGTGGCGCCTGCCATCCTTACGGAGGCTTTCC
TGCCACTGCTCAACCGGGCAGCAGATTTCAATGTGAAGCTGCCAATGGGCTGCCAAAAGGCAGCAGTGGTGAACGTGAGCTCT
GAGGCTGGGTCTATCTCAAGTGTGCCTGGCAACTTTCACATCTCCAAGGTCTACCCTTACCGTGCCAGCAAGGCCGCCCTGAA
CATGCTGATCAGGTGCCTTGCCGAGGAGCTCAAGGAACAGGGAATCCTGTGCGTGGCCTTACACCCAGGTTGGGTCAGGACCA
GAATGGGTGGATCGCGAGCACAGATCGGCATTGAAGAGAGCATCACCGGAATTCTAGGAGTCATGGCTTGCCTCACGGAGAAA
GATAATGGTTCGTTCCTTGACTGGCGGGGTCAGACTGTACCATGGTGA

>SanSDR (eyeless golden-line fish; Accession number GAHL01077601)
ATGAAGATCAACTTCTCGAAGTGTCACTCGGTGCTCATCACGGGGGCAAGCAGAGGTCTCGGACTGCAAATGGTCCGACAGCT
GCTCTATTCGCCGAACAGACCCCAAAAGATCATAGCAACCGCGCGGAATCCAGCCGCTGCACAGGAACTGCAAGAACTCGCCA
AATCGCATCCAAACGTTCATGTTTTGCCTCTAGATGTGACCAGCGACAGCAGTGTGGATGCAGCAGTGCAGTCAGTGGAATCC
ATTGTGGGCACTGATGGTTTGAACTGCTTAATTAATAATGCGGCGATTAACATTACGTCTGAACTGGACACGGTCACTCGAGA
TGCCATGATGAAAACCTACGAGAGTAACACAGTGGCTCCTCTTTTTATGACCAAGGCTTTCCTGCCGTTGTTAAGAAGGGCGG
CAGCTGAAGGCAGTGACATGGGAATCCAGAGAGCCGCGGTGATCAATGTGTCGTCCGTCCTGGGATCTGTGCAGCTCAACTGG
GGCGAGGGGGCGAGCTACAAGAGTTATGCATACAGGACATCTAAGAGCGCCTTGAATATGGTTACAAGGTGTTTGGCTATGGA
TCTGGAAGCTGAGAAGATCTTATGCGTCGCCCTTCACCCTGGTTGGGTGCGCACTGATATGGGCGGACCAACGGCACCTTTGA
GCCCAGAGGAGAGCATATCATCAGTGTTGTCTGTCATTTCTGAATTAACGGAGAAACATCATGGTGGATATGTGGACTACACT
GGGAAAAGCTTACCCTGGTGA

>SauCBR1 (gilthead seabream; Accession number JN811567)
ATGGCCGCCAAGGTTGTCATAGTAACGGGTGGCAACAAGGGCATCGGTCTGGCCATCGTCCGAGCGCTCTGCAAGCAGTTCCA
AGGAGACGTTTACCTGACAGCCAGAGACGTTGGTCGTGGTGAGGAAGCGGTGAAGTCTCTGTCTACGGAGGGACTGAAGCCGA
TGTTCCACCAGCTGGACATCAACGACGTGAACAGCATCACCACCGCTGCCGCCTTCTTTAAACAGAAGTATGGAGGAGTGGAC
GTCCTCGTCAATAATGCCGGGATAGCATTCAAAATGGCTGACACGGCGCCATTCGCAGTCCAGGCGGAGGTCACCCTCAAGAC
AAACTTCTTCGCCACCAGAGACATGCTGACTCACTTCCTGCCGCTCGTCAAAGCCGGAGGCCGTGTCGTGAACGTCTCCAGCT
TCGTCGGCTCCCGCACTCTGAACCAGTGCAGCCCGGCCCTGCAGCAGCGTTTCCGCAGCGAGGACATCACAGAGGAGGAGCTG
GTGGGGCTGATGGAGCAGTTCGTCGAGAAGGCCAAGAAGGGCGAACACAAGGACGCCGGCTGGCCTGAAACTGCGTATGGAGT
GTCCAAGACTGGACTGACGACTCTGACCATGATCCAAGCTCGACGTCTGTCCAAGGAGAGACCGAAGGACGGGATCTTGTGCA
ACGCCTGCTGTCCTGGATGGGTTCGCACCGACATGGCCGGTCCAAAAGCCCCCAAGTCACCAGACGAGGGCGCCATCACTCCG
GTGTACCTGGCCCTGCTGCCGGCCGGCGCCACAGATCCACACGGAAAGTTTGTCTCTGATAAAGAAGTTCAGCCGTGGTGA

>SauSDR-like (gilthead seabream; Accession number EU557022)
ATGGCAGCGCAACCAGTCACCGTGCTGATCACAGGAGCCAACAGAGGCTTGGGCCTGGAGATGGTTAAGCAAATGCTGGAGGG
AAAATGTCCGGTGAAAAAGCTGTTTGCCTGTTGCAGAGACCCGAATGGACCCAAGGCTGAGGCCCTGCAAGCGCTGGCAAAGA
AGCATCCTAATGTCCACATCATACGTCTAGACGCCGCTGACCTTGGTAGCATTAAACAGTGCTCCCAGCAGGTGGGCGCTCAG
GTGGGGACGGGTGGTCTCAACCTGCTGATTAACAATGCAGGGTACTTGGACAAAAGCACCCTGCAGAACACCACTACAGAGGG
CATGCAGACCACTTTGAACACCAATCTCATGGGTCCTATGTACATGACTCAAGAGTTCCTGCCTCACCTTCGTGCGGCAGTGA
AGGCCAGTAAAATACCCGGGATGTCCACCAGGAAAGCAGCCGTCGTCAGCATCACCTCATTTCTGGGTTCAATGCAACTTGTC
AAAGACTCATATTCCTTCTTCCCTGCCATCTCCTACCGCATCAGCAAGGCTGGTCTGAACATGCTGACAGTCTGTGCTGCAGA
GGAGCTGAAGAAGGACGAGATCCTGTTTTCTCTGCTGCACCCTGGCTGGGTGCGCACCGACATGGGTGGAGAGGAGGGGGAGA
TTGACGCCCCGGAGAGCGTGACGGGGATGCTTAGCGTGATGGAGTCCCTGACTGAGAAGCAAAATGCAGCATTCCTGGATTAT
AAAGGCAAAACCCTCCCCTGGTAA

>ScaSDR-like (small spotted catshark; SkateBase accession number SSC-transcript-ctg22409)
ATGCTGAGGTCGGTGCTGGTGACGGGCTCGAACCGCGGCCTGGGCCTGGAGCTGGTCCGCCAGTTGGCGCAGGCTGAGCCGCC
GCCCCGCAACCTGTTCGCCTGCTGCAGGCAGCCCGACAGCCCAGGCGCCCAGGATCTTGTGAACCTGGCAAAGTGTTACTCAA
ACATAAATGTACTGAAGTTGGATGTTGATGACTTTGGTAGCATTCAAGAGTGTGCTGATCAGATTGAGAAGATATTGGGAAAC
GATGGTCTAAATCTGCTAATAAATAATGCTGGAGTCAACTTTGGTGGAGGCCTCAACGAAGTTACTCCTGAAACAATGATGAA
GAGTTACAGAACCAATGCTGTTGGACCGATGATGGTAACCAAGGCTTTTTTGCCAGCTTTACAGCGAGCAGTTCAGCTGAATG
ATAAATTTAAAGCAATAGTTATCAACATATCCTCCATCTTGGCCTCCATGGAGCTGTTTGACAATAAAAGTGGGACAACATAT
GCATACCGGGTCAGCAAGGCTGCATTGAATATGGTAACCAAGTGCTTGGCGGATGAGTTGAAACTCCATGGAATCCTTTGTAC
TTCAGTGCACCCTGGCTGGGTTAAAACAGAGATGGGTGGCGAAAAGGCACCACTAACTAAGGAAGAAAGTGTTTGTGGAATTC
TCAAAGTACTATCGACGCTTTCAGACAAAGATAATGGATCATTTCTGGACTGGAGAGGTCAGCGTTTACCTTGGTGA

>SsaSDR (Atlantic salmon; Accession number BT058549)
ATGAGTAGCGCTTTTAAAAATTGCCGAACGGTAATGGTGACAGGAGCCAGCCGAGGCCTCGGTCTGCGAGTAGTGGAAAGTCT
GGTGGCAGGGAGCTCTGATCCGGATAAGATTATTATTGCCACGGCCAGAAACCCGAGCGGGGCCCAGGAACTGCAGAAACTAG
CTGAAAGACATCCAAACATCCATATAATAACACTTGATGTTGTCAGTCAGGAGAGCATAGAGAAAGCTGCACAAGACGTTGAC
CTGCTTGTTCAGGAGCTGGGCCTCAATTGTCTAATCAACAATGCTGGAATCAATGTTGTCGCCAATTTTGAGACAGTCACGGC
TGAGAAGATGTTGGAGAACTTCCACACCAACTCTGTGGCCCCCCTTATGATCACTAAGGCTCTCCTACCCCTGCTGAAGAGGG
CAGCTGCCAAAGGCACGGGGATGGGCATCCATAGAGCTGCGGTCATCAACATGACTTCCTTGTTAGGGTCAGTCGAGCTCAAC
TGGGGTGACCGGGCCAAGAACTTCAAGTGGTATCCTTACAGAACATCCAAGAGTGCTCTGAACATGGTGACCCGATGCATGGC
AGTGGATCTGGAGGCTGATGGCATCCTGTGCATGGGCCTCCATCCTGGCTGGGTACGCACTGACATGGGCGGACCAGAGGCAC
CTTTGAGTCCAGAGGAGAGCATATCTTCTGTGTTGTCAGTAATTGGTGGATTGACTGAAAAGGATCATGGGTCATTTCTACAC
TACACCGGGGAGCCACTACCCTGGTGA

>TguSDR-like (zebra finch; Accession number XM_002198285)
ATGGCGGCGGCGCGCACGGTGCTGCTGACCGGCTCCAACCGCGGCATCGGCCTGGAGCTGGTGAAGCAGCTGCTGGGCTCGCC
GCGACCCCCGGCCTGGATCTTCGCCACCTGCCGGGACCCCGAGGGGCCGCGGGCACAGGAGCTGAGAGATCTGGCATCCAAAC
ACCCAAACCTGGTTCTTGTAAAGCTGGATGTGGAAAACCCCTCGGCTATCACCGATGCGGCGAAGGTGGTGGAGGGGAAGCTG
AACGGAATGGGGCTGAACCTGCTGATAAACAACGCCGGCATCTACACCCCCACGGCCTCGCTGGAGACGGTCGACGCTGAGGA
CATGGTCAGGACGTACAAGACCAATGCAGTGGGGCCAATGCTGATGGCCCAGGCTTTCCTGCCTCTGCTGAAGAAGGCTGCAC
AGGACAGCAAAGAAAAGGGCCTGAGTTGCAGCAAGGCAGCCATCATCAACATGTCCACCATCTTAGGGTCCATCAAGAAAACA
GCTGATTCCTTCTTCCACCCTGTCATCTCCTACCGCTGCAGCAAGGCTGCCCTCAACATGCTGACCATGTGCCAGGCTCTGAC
CTACAAGGAAGCTGGGATCCTGTGCGTGGCACTGCACCCTGGCTGGGTGAAAACAGACATGGGCACCCAGGAGGCTGACCTGA
CAGTGGACACAAGTGTACGGGGGTTGTTGTCTGTGCTGCCAATCCTTTCTGAGAAACACAGTGGGACTCTGCTCAACTGGGAA
GGGAAAGCTATCCCCTGGTGA

>TniSDR-like (green spotted pufferfish; Accession number CAAE01014626)
ATGGCAGCGCAGAAGGTGAACGTGCTGATCACAGGAGCCAACAGGGGCCTGGGCCTGGAGATGGTCAGCCAAATGGTGAAGGG
CAGCGTTGCGGTGAGTAAGCTGATCGCCTGCTGCAGAGACCCGGATGGACCTGGGGCCGAGGCCCTGCAAGTACTGGCAAAGC
AGCATCCTGATGTCATCTCGGTTGTTCGTCTAGATACCTCCGACCTTTCTAGCATAAAAGAGTGCACCCAGCAGGCTGGCGCC
CTGGTGGGGAGCGAGGGGCTCAACCTGCTGGTTAACAACGCAGGGATTATGATCAATACCACCCTGCTGGACACCACCTGCGA
GGATATGCAGAAGATTTTCAACACCAATGTTTTGGGTCCCATGAACATGATCAAAGAGTTTCTGCCGTTTCTCCGCGCAGCAG
CAAAGGCCAGTAAGATATCGGGGATGTCCACCAGGAAAGCAGCCGTTATCTGCATTTCCTCACTTTTGGGTTCCGTTGAAGGC
ACCAGGCAAACCTACGAGTACTTCTCTGTGCTCCCCTACCGCGTCAGCAAGTGCGCCCTGAACATGCTAACTGTATGCGCCGC

APPENDIX

AGCAGAGCTCAGGAAGGAAGAGATCTTGTTTTCTTTGCTTCACCCCGGCTGGGTGCGAACCGACATGGGTGGGGAGAAGGCAT
CGATCGACGCCCCAGAGAGTGTGCAGAGGATGCTCGACGTGATGGCTTCAATGACGGAGGAGCAAAATGGAGCATTTCTGGAT
TATCATGGTCAAACTGTTGCTTGGTAG

>TorSDR-like (Pacific bluefin tuna; Accession number BADN01016024)
ATGGCAGCCCAACCAGTCAGCGTGCTTATCACAGGGGCCAACAGAGGCCTGGGCCTGGAGATGGTTAAGCAAATGATGGAGCC
CCCCTCTCAAGTGAGAAAACTGTTTGCCTGTTGCAGGAATCCACATGGACCGAAGACCGAGGCCTTGCAAACTCTGGCAAAGA
AGCACCCCAAAATCATATCCGTCATCCGCCTGGACGTCACTGACCTTTGTAGCATAAAACAGGCTGCCCCACTGGTTGGCTCT
CTGGTGGGGACGGGGGGTCTCAACCTGCTGATTAACAATGCAGGGATCCTTGCCAAAGGCAACCTGCAGGACACCAGTCCCGA
GGACATGCAGAGTGCCTTCAACACCAATGTCATGGGCCCGATGAACGTTATTAAAGAGTTCCTGCCTCACCTACGTGCAGCAG
TGAAGGCCAGTGGTATGCCGGGGATGTCCTGCAACAAAGCAGCCGTTGTCAACATCTCCTCAGTTATGGCGTCAATGGAATCT
TTAAAACAGTCATATGCCTCCTTACCTGCCATCTCCTACCGCATTAGCAAGGCAGCTTTGAACATGCTGACTCTGTGTGCTTC
AGAGGAGCTGAAGGAGGATGGGATTCTGTTTTCAGTGCTTCATCCTGGCTGGGTGCGCACTGACATGGGTGGAGAAGAGGCGG
ATATTGCTGCTCCGGAGAGTGTGCAGGGGATACTCAGGGTGATGGCCTCCTTGACTGAAAAGCACAATGGAGCCTTCCTGGAT
TATAAGGGCAAGACTGTCCCCTGGTAG

>XlaSDR-like (African clawed frog; Accession number NM_001096655)
ATGTCTGAGCTGAACCTCCGCACTGTCTTGGTGACTGGGTCGAACAGGGGGATTGGATTTGAATTTGTGCAACAAATTATTAA
CAGCCAAAATTCTCCTCATAAGATTTTTGCCACATGCCGGGACCCAGGTGCTCAGCAAAGTCAGGAGTTGAGGAAGCTGTCTG
AAAAGCATCCAAATGTTGTAGTAATTCAGCTTGATACTACCAACCCTGCAAGTGTTAATGCGTCTGTGAAAGAAGTAGAGAAG
CATCTAAATGGAGAAGGCCTGGACCTGCTGATCAATAATGCTGGGATTCTGACTCAAAACTCTCTAGAGACCCAAACCTCAGA
AGACATGATGAATGTCTACAACGTGAATGTGGTTGGCCCTATGCTAATGACTCAGGCATATCATCATTTGCTAAAACGATCTG
GAGTAGAAAGCTCAGGAAAATCAGCTATTGTTCATATCTCAGCTCTGCTGGGCTCCCTAGAAGAGCTTCCCAATTTGTTTTCA
GCATTGCCAGTTATCTCCTATCGCTGCAGCAAGGCTGCTCTGAACATGCTCTCAAGGTGTCACATGGAAGGTTACAGGCAGGA
TGGAATTATTTCGATTGCCATTCATCCAGGCTGGGTGCAAACTGACATGGGTGGTGAAAAGGCACCATTAACAAAACAAACCA
GTGTTGCTGGCATGATGAAGATAATCTTTTCCCTGAATGAACAACATAATGGCACCTTTGTGGATTGGGAAGGCAAAACAATA
CCTTGGTGA

>XmaSDR (Southern platyfish; Accession number XM_005810114)
ATGAGTCAGGGTGTGTGGAGGTCCGTGTTGATGACTGGTTCCAGCAGAGGGATAGGACTGCAGCTGGTTACACAGCTGGCTAA
CAGCAGCAACAGACCCGCTACCATCATAGCGACAGCCCGCAACCCTGCTGCTTCCACGGCTCTGCAGGAGCTGTCAAAAACAC
ACTCAGGTCTTCATATTGTTACATTTGATGTGAGCAGTGAGCAGAGCATAAGTTCAGCCTCTGAACAGGTTCAGTCTATTGTG
GGAAATTCAGGACTCAACTGCTTGATTAATAATGCTGCCATTGGATTCTCCACTAACATCAGCACTGTAACCCCAGAGGCAAT
GATGACGATGTTCCAGGTCAACACTGTTGCACCACTCTTTGTAACAAAGGCCTTCATGCCTCTGCTGGAGACTGCTGCTGCCA
GGTCAGGTGGGATGGGGATCCACAGAGCAGCTGTCATCAACATGTCCTCTGTCCTTGGTTCCATCACACTGAACTGTGGAGAT
GCCGCCAAATTCAGGAGCTATGCCTATTGTACCTCTAAGGCAGCCCTAAACATGGTGACAAGGTGTCTAGCTGCTGATTTGGA
GTCCAGTGGAATTCTCTGCATGTCTCTTCACCCTGGCTGGGTGAGGACGGACATGGGAGGACCTCATGCTGACTTAACAGTAG
AGCAGAGTGTTTCTGGAATCCTGTCTGTCCTCTTCAGCCTTAGCGAGCGGGACCATGGAGAATTCCTTGACTACCATGGACAG
AAGCTACAGTGGTGA

Supplementary Data Chapter 4

Supplementary Table 4.1

PCR primers used in this study.

Name	Sequence (5'-3')*
SauBHMT3_Rv2_promoter	TTCTCCAGGGCGAAGACAAAGC
SauBHMT3_Rv1_promoter	TCCCGCATCCAGACGCTCC
SauBHMT3_promo_Rv_pGL4_HindIII	CACGC**AAGCTT**ATCCTGCTCCCCTTCTTCTGTCCAGC
SauBHMT3_promo_Fw_pGL4_XhoI	CCGGAG**CTCGAG**TATCCTCATCCCTCTTCAGTTTGGTCAGC
SauBHMT3_promo_Fw2_pGL4_XhoI	CCGGAG**CTCGAG**GTCAAATGTATCAGAGAGAG
SauBHMT3_promo_Fw3_pGL4_XhoI	CCGGAG**CTCGAG**GATCAATTACGGTAAGTACAATT
SauBHMT3_promo_Fw4_pGL4_XhoI	CCGGAG**CTCGAG**GTGCGGGAGCAGGCGG
BHMT3 qPCR Fw	CCACTGAGCGCCTGCTGACCACGTTA
BHMT3 qPCR Rv	CATTTACCATCAACGGGTGCCTGTTTG
BHMT4 qPCR Fw	CGTGCTGGAGATGCAGAAGAAGCC
BHMT4 qPCR Rv	GACTCTTTTCTCCCTCGTCCCTCTGTTGTAG
BHMT5 qPCR Fw	CTGAGTCAGCTCATCGCCTCCAGTAGAG
BHMT5 qPCR Rv	CCACAGAACAGTGCCAGTCTCAAGTGAAGG
BHMT6 qPCR Fw	GTTGAGCCATGTCACCGGCATCAC
BHMT6 qPCR Rv	GACTTCGCATAGCAGGGAGTCATGGAAAC
BHMT7 qPCR Fw	GGAGATCAGTGGTAACGTCCCCAATATCAC
BHMT7 qPCR Rv	CTGTGGCTCGTCACATAACACGGAGTCTTAG
RPLreal-FW	AAGAGGAACACAACTCACTGCCCCAC
RPLreal-RV	GCTTGCCTTTGCCCAGAACTTTGTAG

* Bold underlined sequence indicates cutting sites for endonucleases *Xho*I and *Hind*III in forward (FW) and reverse (RV) primers, respectively.

APPENDIX

Supplementary Figure 4.1

Complete coding sequences (collected or reconstructed) for betaine homocysteine S-methyltransferase (BHMT) used to construct the phylogenetic tree presented in Figure 4.10. Cmi, *Callorhinchus milii*; Ler, *Leucoraja erinacea*; Sca, *Scyliorhinus canicula*; Loc, *Lepisosteus oculatus*; Afi, *Anoplopoma fimbria*; Dla, *Dicentrarchus labrax*; Dre, *Danio rerio*; Fhe, *Fundulus heteroclitus*; Gac, *Gasterosteus aculeatus*; Gmo, *Gadus morhua*; Hbu, *Haplochromis burtoni*; Ola, *Oryzias latipes*; Omy, *Oncorhynchus mykiss*; Pch, *Paralabidochromis chilotes*; Pfl, *Perca flavescens*; Plf, *Platichthys flesus*; Pma, *Pagrus major*; Pol, *Paralichthys olivaceus*; Ppr, *Pimephales promelas*; Sau, *Sparus aurata*; Sma, *Scophthalmus maximus*; Spa, *Stegastes partitus*; Sra, *Sebastes rastrelliger*; Ssa, *Salmo salar*; Tni, *Tetraodon nigroviridis*; Tru, *Takifugu rubripes*; Lch, *Latimeria chalumnae*; Cpy, *Cynops pyrrhogaster*; Xla, *Xenopus laevis*; Xtr, *Xenopus tropicalis*; Aca, *Anolis carolinensis*; Asi, *Alligator sinensis*; Cbe, *Chrysemys picta bellii*, Cmy, *Chelonia mydas*; Gga, *Gallus gallus*; Psi, *Pelodiscus sinensis*; Hsa, *Homo sapiens*; Mdo, *Monodelphis domestica*; Mmu, *Mus musculus*; Oan, *Ornithorhynchus anatinus*.

>CmiBHMT1 (elephant shark; Accession number NM_001292237)
ATGCCTCCCGTGGGAGCCACCGGCAAGAGCGCCGCCAAGAGGGGAATTTTGGAGCGTTTGGAAGCAGGAGAGATCGTTATTGG
TGATGGCGGGTTTGTTTTTGCTCTGGAGAAGAGGGGATATGTCAAAGCAGGTCCCTGGACTCCAGAGGCAACAGTAGAGTATC
CCGAGGCTGTAAGACAACTGCACCGTGAATTCCTCAGAGCAGGCTCTACAGTTATGCAGACTTTCACCTTCTATGCTAGTGAT
GACAAGTTGAAGAATAGAGGAAACATTGCAGGTCAAAAGTTCTCAGGCCAAAGTGTGAATGAAGCTGCATGTGATATTGCTCG
GGAGGTGGCCAACGAAGGGGATGCATTGGTTGCTGGAGGAGTCTCTCAAACTCCAGCCTACCTCAGCTGTAAGAGTCAGTCTG
AAGTGAAGGAAATCTTTAGGAAACAGCTGGAAGTCTTTGTCAAGAAGGATGTGGACTTCATGATTGCTGAGTACTACGAGCAT
GTGGAAGAAGCAGTGTGGGCAGTTGAAGTTTTAAAGGAATCTGGGAAACCAGTTGCCGCCACCCTATGTATTGGTCCAGAGGG
TGATCTGAAAGGAGTTTCACCTGGGGATTGTGCAATCAGGCTGACCAAGGCTGGTGCTCAGATTGTTGGAGTGAACTGCCATT
TTGACTATGCCACTAGTCTCAAAACTATGAAGCTTATGAAGGAAGGTTTAGATGCTGCAAAACTCAAAGCACATCTGATGGTG
CAGCCTGTGGCTTTCCATACACCTGATTGTAACAGCCAGGGATTTATTGATCTTCCAGAGTTTCCTTTTGCTTTGGAGCCCCG
AATTTCTACCAGATGGGACATGCATAAATTTGTAAGGGAAGCATACAATCTGGGAATCAGGTACATTGGCGGCTGCTGTGGAT
TTGAGCCCTATCACATCAGGGCTGTAACAGAGGAGCTCTCAGCTGAGAGGGGCTTCCTTCCCCAAGCTTCAGAGAAACACGAT
AGCTGGGGGACTGGCCTCTCCATGCACACAAAACCGTGGGTGAGAGCAAGAGCCCGGAGGGACTATTGGGAGAACCTGGTGCC
TGCAAGTGGCTTGCCTTATTCACCCTCCTTATCCAAACCAGATAGCTGGAAAGTGACCAGGGGAGATAAAGATCTGATGCAGC
AAACAGATGCCACAACTGAGCAACAACAGAAAGAATTGTTTGAAAAGCAGAAAGCTAAGGCTCTACTGTCAGCATAG

>LerBHMT1 (little skate; Accession numbers FF600473, DR713938, CV067236, GH161124, FL592306, CO050181 and GD273388)
ATGCAGCCAGTGAGGAAGGACAAAAAGAACGCGCCGCACAGAGGACTTCTGGAACGTTTAGCGTCCGGGGAGATCATTATTGG
AGATGGGGGATTTGTTTTTGCACTTGAGAAGAGAGGATATGTCAAAGCTGGTCCTTGGACTCCAGAAGCAACAGTGGAACATC
CAGAGGCAGTGAGGCAGCTGCATCGTGAATTCCTCCGAGCAGGTTCCACAGTCATGCAGACATTCACTTTTTATGCAAGTGAT
GACAAGTTGATGAACAGAGGGAATTATGTGCAACAAAAGATTTCAGGCAACAAGGTGAATGAAGCTGCGTGTGACATTGCTAG
GGAGGTGGCCAATGAAGGCGATGCCCTAGTTGCTGGTGGAGTCAGCCAAACACCTTCCTACCTGAGCAACAAAAGTGAGACTT
TAGTTAAGGCCATCTTCAGAAAGCAGCTGGATGTTTTTGTCAAGAAGAATGTCGACTTCCTTATTGCTGAGTACTTTGAACAT
GTAGAGGAGGCGGAATGGGCAGTTCATGTTATGAAGGAGTCTGGCAAAGCTGTCGCTGCCACATTGTGCATTGGCCCAGAAGG
TGATATGGCCGGGGTATCAGCTGGGGACTGTGCTGTCCGACTTGTAAAAGCAGGTGCTGACATTGTTGGAGTGAACTGCCATT
TTGACCCAGCCACTTGCCTTAAAACAGTGAAACTTATGAAGGAAGGATTGGAAGCTGCAAAACTAAAGGCACACTTGATGGTG
CAGCCTCTTGCCTTTCATACACCTGACTGCAGCAAGCAGGGATTCATTGATCTACCAGAGTTTCCATTTGCTCTGGAACCACG
AATTGCCACCAGATGGGACATGCATAAATATGCAAGGAATGCGTATGATCTGGGCATCAGATACATTGGTGGCTGCTGTGGCT
TTGAGCCTTATCACATTCGGGCCATTGCAGAAGAATTGGCAACTGAGAGGGGCTTCCTGCCAGCAGGTTCAGAGAAACATGAT
GCCTGGGGCAGTTGTCTTGCCATGCACACAAAGCCATGGGTGCGAGCAAGAGCTCAGCGGGCATATTGGGAGAAAATGTTGCC
TGCCACTGGCCTACCATATTCACCTTCCTTCTCAAAGCCAGACAGCTGGGGTGTGACCAAAGGAGATGAAGTGCTCATTCAAC
AAGTTGAAGTGACATCTAATGATCAACTGAAAGAGATGTTTGAGAAGCAGAAGAGATTTGCTGGAATCGTAGCATAA

>ScaBHMT1 (small-spotted catshark; Accession number SkateBase-ctg15221
ATGCCGCCCGTGGGGAACGACGGAAGGACCGCTGCGAAAAGAGGACTTCTGGAGCGTTTGGCTTCTGGGGAAGTCATTATTGG
AGATGGAGGATTTGTTTTTGCTCTTGAGAAGAGAGGCTATGTCAAAGCTGGACCATGGACACCAGAAGCAACAGTAGAGTATC
CAGAGGCTGTGAGACAGCTGCACCGTGAGTTCCTCAGAGCAGGCTCCACAGTCCTGCAGACTTTCACTTTTTATGCAAGTGAT
GACAAGTTGCAGAACAGAGGAAATTATGTGGAGCAAAAGATTTCAGGCCAGAAAGTGAATGAAGCTGCTTGTGATATTGCGAG
GGAGGTGGCTAACGAAGGAGATGCCCTCGTTGCTGGAGGAGTCAGCCAAACCCCTTCCTACCTCAGCTGCAAAAGTGAGACTG
ACGTAAAGGCCATCTTCAGGAAACAACTAGACGTCTTTATCAAAAAGAACGTAGACTTTATGCTTGCTGAGTACTTTGAACAC
GTGGAGGAAGCAGAATGGGCAGTTCATGTTTTGAAGGAGTCTGGCAAACCAGTTGGTGCCACCTTATGCATTGGTCCAGAGGG
TGATTTGGCTGGAGTGTCAGCAGGGGACTGTGCAGTTAGACTCGTAAAGGCTGGTGCTGACCTCATTGGAGTGAATTGCCATT
TTGACCCAACCACTTCCCTCAAAACTATGAAGTTTATGAAAGAAGGCTTAGAAGCTGCACATCTAAAAGCACACTTAATTGTG

CAGCCTCTTGCTTATCACACACCTGATTGCAACAAGCAGGGATTCATTGACCTACCAGAATTTCCATTTGCCCTGGAGCCACG
AATTCTTACCCGATGGGAAATGCATAAATATGCAAGGGCTGCATATGATTTGGGAATCCGATACATTGGTGGCTGTTGTGGAT
TTGAACCCTATCATATTAGGGCTATAGCGGAAGAACTAGCAGCTGAAAGGGGCTTCCTTCCAGTGGCTTCAGAGAAGCACGAT
GCCTGGGGAAGTGGTCTTGCCATGCACACAAAGCCTTGGGTGCGAGCAAGAGCTCAGCAGGCATATTGGGAGAAGCTGGTGCC
TGCGTCTGGCCGGCCCTACTCAGCCTCCCTTTCAAAGCCGGACAGCTGGGGCGTGACCAGAGGAGACAAGGAGCTAATCCAAC
AAAAAGAAATGACAACAGAAAACCAATTGAAAGAAATGTTTGAGAAGCAGAAGATTGTTTCTGGAATGTAG

>LocBHMT3 (spotted gar; Accession number XM_006626723)
ATGGCACCAGCGGGGTCCAAGAAAGGTATTCTAGAGCGCTTGGATGCTGGGGAGATTGTTATTGGTGACGGGGGGTTTGTGTT
TGCCCTGGAGAAAAGAGGCTATGTGAAGGCAGGGCCATGGACTCCAGAAGCTGCGGTGGAGCACCCTGAGGCAGTTCGACAGC
TGCACAGGGAATTTCTGAGAGCGGGAGCAAATGTCATGCAGACGTTCACATTCTATGCCAGTGACGATAAACTGGAGAACCGG
GGCAACGACTTGAGAATATCTGGAGAAAAAATCAATGAAGCAGCCTGTGACTTAGCCAGGGAAGTCGCCAATGAAGGAGATGC
CTTGGTTGCCGGAGGAGTGTCCCAGACCCCTTCCTACCTGAGCTGTAAGAGTGAGAATGAAGTGAAGGCAATCTTCAAGAAAC
AACTGGAGGTCTTCATTAAAAAGAATGTGGACTTTCTGATTGCTGAGTACTTTGAGCACGTGGAGGAAGCTGAGTGGGCAGTT
CAGGTTCTGAACACCAGTGGAAAGCCGGTTGCAGCCACTCTGTGTATTGGGCCGGAAGGAGACCTGAATGGAGTCAGCTGTGG
TGAATGTGCTGTCAGATTGGTCAAGGCTGGTGCCAAGATTGTCGGTGTGAATTGTCACTTTGATCCCATGACGTGCCTGAAGA
CTGTGAAGCTCATGAAGGAAGGCCTGGAGGCAGCAAAGCTGAAAGCTCATCTCATGATCCAGCCCCTTGCTTACCACACCCCT
GACTGTAACCATCAGGGATTCATTGACCTCCCAGAATTCCCCTTTGCCTTGGAGCCAAGAATCCTAACCAGGTGGGACATGCA
CAAATACGCCAGAGAGGCCTACAAGATGGGCATCCGCTACATTGGTGGCTGCTGTGGCTTCGAGCCGTACCACACCCGTGCCC
TGGCAGAAGAGCTGGCACCCGAGAGAGGTTTCCTCCCTGACGCCTCTCAGAAGCATGGCAGCTGGGGTAATGGTCTGGACATG
CACACCAAGCCCTGGGTCAGGGCCAGGGCTCGTCGTGATTACTGGGAGAAGTTGCTGCCGGCATCAGGGCGCCCCCTGTGTCC
TTCCATGTCTACACCAGATGGATGGGGAGTGACCAAGGGACATGCTGACCTGATGCAGCAGAGAGAGGCGACCACTGAGGGGC
AGCTGAAGGAGCTGTTTGAAAAGCAAAAGAAGGCCAAGGTCACAGCTTGA

>AfiBHMT6/7 (sablefish; Accession number BT082667)
ATGGAGAACAAGAGGAGGGGTATCTTGGAGAGACTTAATGCTGGGGAGGTGGTTGTGGGTGATGGAGGTTATGTGATGCAGCT
AGAGCGACGTGGCTACGTGAAGGCTGGACACTGGACACCTGAGGCAGCTGTTGAACATCCTGAAGCAGTGCGGCAGCTGCACA
GGGAGTTCCTGAGAGCGGGAGCCAATGTGCTTCAGACATTCACCTTCTACTGCAGTGAGGATAAACTGGAGATCAGTGGCAAT
GTCACCAACATTACTGGGGCTCAGATCAATGAGGCAGCTTGTGACCTGGCCAGGGAGGTAGCCAATGAGGGGGATGCATTGGT
CGCTGGCGGTGTGTCTCAGACTCCCTGTTATGTGACGAGCCACAGTGAGACTGAAGTCAAGGCCATCTTTAAGAAGCAGATGG
ACGACTTCCTCAAGAAGGACATTGATTTCTTTATAGTAGAGTATTTTGAGCACGTGGAAGAGGCAGTGTGGGCAGTGGAGGTG
CTGAAGACCAGCGGTAAACCAGTGGGTGCAACACTGTGCATCTCCCCTCAAGGAGACATGCACGGAGTCCCACCTGGAGAGTG
TGCTGTCAGACTGGTCAGAGCTGGAGCTGACATTGTTGGAATAAATTGCCACTTGGACCCTCTAACATGCATTCGTACAGTGA
AGTTGATGAAAGCAGGATTAGAGAATGCTGGTCTCAAAGCTCATCTCATGATCCAGCCACTGGGTTTTCACACGCCTGAGTGC
AACCACACTGGATATCTCTCGCTACCTGAGTTCCCCTTCGCATTGGAGACTAGGTCAATGACCCGCTGGGACATCCATCAATA
TGCCAGAGAGGCTTACAATGCTGGAATTCGCTACATTGGTGGCTGCTGTGGATTTGAGCCCTACCATATCAGAGCTATAGCAG
AAGAGGTGGCTGCAGAGAGAGGATTCCTCCCACCAGCTTCAGAGAAGCACGGATTGTGGGGAGCTGCCCTTGAGATGCACACT
AAACCCTGGGTCAGAGCCAGGTCTCGTCGAGACTACTGGGAAAACCTTTTGCCTGCTTCTGGACGTCCCAAATGCCCTTCCAT
GGCCATACCAGCAGATGAATAA

>DlaBHMT4 (European seabass; Accession numbers FK941601, FM021854, FK941836, FK942038 and FK941354)
ATGGCACCTGGCAAGAAGGGTATCCTTGAGCGTCTGAATGCTGGGGAGGTGGTGATTGGCGATGGAGGCTTTGTGTTTGCCCT
GGAGAAGAGGGGCTATGTGAAGGCCGGGCCATGGACTCCTGAGGCTACTGTCACTCACCCTGAGGCTGTGCGGCAGCTGCACA
GGGAGTTCCTGAGGGCAGGAGCTAATGTCATGCAGACATTCACATTCTACGCCAGTGATGACAAACTGGAGAACAGGGGTCAA
TCATTGAGACTCTCTGGGGCACAAATCAACGAGGCAGGCTGCGACCTGGCCAGAGAAGTAGCCAGTGAAGGAGATGCTCTGGT
AGCTGGCGGAGTGTGTCAGACTCCTTCCTACCTGAGCTGCAAGAGCGAGACAGAAGTGAAGGCCGTCTTCAAGAAACAGATGG
AAGTGTTTGTGAAGAAGGGCGTGGATTTCATGATTGCTGAGTACTTTGAGCACGTTGAGGAGGCGGAATGGGCCGTGCAGGTG
CTAAAGACCAGCGGAAAGCCCGTGGCTGCTTCTCTGTGCATCGGACCAGAGGGAGACATGCATGGCGTCTCACCTGGAGAGTG
TGCTGTGAGGCTAGTGAAGGCCGGTGCCCAGATCATTGGAGTCAACTGCCACTTTGACCCCATGACCTGTGTGAAGGCTGTTA
AGTTGATGAAGGAGGGAGTGGAGAAGGCCGGACTGAAGGCTCACTACATGGTGCAGCCCCTGGCATTCCACACCCCCGACTGC
AACTGTCAGGGATTCATTGACTTGCCAGAATTCCCCTTCGGCCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACCAGTA
CGCTAGAGAGGCTTACAAGGCTGGCATCCGATTCATTGGCGGCTGCTGTGGTTTTGAGCCTTACCACATCAGGGCTGTGGCAG
AGGAGCTGGCCCCTGAGAGAGGGATAACGCCCCCTGGCTCAGAGAAACATGGCATGTGGGGTGCTGGTCTTGAGATGCACACC
AAACCCTGGGTCAGAGCCAGGGCCCGTCGTGACTACTGGGAGCATCTCTTGCCTGCATCCGGTCGTCCCAAGTGCCCGTCCCT
GTCCACACCAGAGGGCTGGGGTGTGACAAAGGGCCACGCCGACTTGCTGCAACACAAAGAGGCCACCAGCATCCAGGAAATGA
AGCACGTGATGGAGATGCAGAAGAAGGCCAAGTCCTCTGCTTGA

>DreBHMT3 (zebrafish; Accession number NM_001012480)
ATGGCACCAGTCGGATCCAAACGGGGAGTTCTGGAGCGCCTGAACGCGGGTGAGGTGGTGATTGGAGACGGAGGGTTCGTGTT
CGCTCTGGAGAAGCGCGGGTATGTGAAGGCCGGACCCTGGACACCAGAGGCCGCTGCGGAGCACCCGGAGGCCGTGCGGCAGC
TGCACCGGGAGTTCTTGCGCGCTGGCTCAAATGTCATGCAGACGTTCACTTTCTACGCCAGTGACGACAAGCTGGAGAACAGA
GGCAACAAACTGAGCTTCACTGGCCAGCAGATCAATGAGGCGGCCTGTGACCTGGCCAGAGAAGTGGCCAATGAGGGCGATGC
GCTGGTGGCGGGCGGAGTCTCGCAGACGCCTTCATACCTGAGCTGCAAGAGTGAAGAGGAGGTGAAGAAGACCTTCAAGAAAC
AGCTGGATGTCTTCATCAAGAAGAATGTAGACCTGCTGATCGCTGAGTACTTTGAGCATGTGGAGGAGGCTGAATGGGCCGTC
CAGGTTCTGAAGGCCACAGGAAAGCCTGTAGCAGCGACTCTGTGTATTGGACCTGACGGAGACATGCCTGGAGTTACACCTGG
AGAGTGTGCCGTCAGACTGGTGAAAGCAGGAGCTGATATTGTGGGTGTGAACTGTCACTTTGACCCCCTGACCTGTGTGAAAA
CTGTGGTGATGATGAAGGCTGCAGTGGAGAAAGCGGGTCTGAAAGCGCATTATATGACGCAGCCGCTGGCGTACCACACACCC

APPENDIX

GACTGCAGCTGCCAGGGCTTCATTGACCTGCCCGAGTTCCCATTCGCTCTGGAGCCGCGGATCCTGACGCGCTGGGAGATGCA
GCAGTACGCCAGAGAGGCCTATAAAGCTGGAATCCGCTACATCGGTGGCTGCTGCGGGTTCGAGCCGTACCACATCCGCGCTG
TAGCTGAAGAGCTGTCGGCCGAGAGAGGATTCCTGCCCGAAGCCTCACAAAAACACGGCCTGTGGGGCAGCGGCCTGGAGATG
CACACCAAACCCTGGGTCAGGGCCAGGGCTCGCCGTGACTACTGGGAGAAGCTGAAGCCGGCGTCTGGTCGTCCACTCTGCCC
CTCCATGTCCACTCCAGACGGTTGGGGGGTCACCAGGGGTCACGCGGCCCTCATGCAGCAGAAGGAGGCCACCACGGCGGAGC
AGCTGCGGCCACTGTTCCAGCAGGCGGACGCTAAACACTGA

>FheBHMT3 (mummichog; Accession number XM_012853912)
ATGGCACCAGCTGGAGCAAAGAAAGGCATTTTGGAGCGCCTGGATGCAGGAGAGGTGGTCATCGGAGACGGAGGGTTTGTGAT
CGCCCTGGAGAAAAGGGGTTATGTGAAAGCGGGACCATGGACTCCAGAGGCTGCAGCCGAGCACCCTGAAGCAGTGCGGCAGC
TGCACAGGGAGTTCCTGAGGGCTGGCTCTAATGTCATGCAGACCTTCACTTTCTATGCCAGCGATGACAAGCTGGAGAACAGA
GGGCACACTCAGCGTTTTACTGGGGAGCAAATAAACGAGGCGGCTTGTGATCTGGCCAGGGAGGTGGCCAACGAGGGCGATGC
CCTGGTGGCGGGAGGAGTCTCTCAGACGCCCACATACCTCAGCTGTAAGAGCGAGGAGGAGGTCAAGGCCATTTTCAAGAAGC
AAACTGACGTGTTTGTCAAGAAAAACGTGGACTTCTTGATTGCAGAGTACTTTGAGCACGTGGAGGAGGCCGAGTGGGCGGTC
CAGGTTCTGAAGAGCACCGGGAAGCCCGTGGCTGCGACTCTGTGCATCGGACCTGAGGGAGACCTGAACGGAGTCAGTCCTGG
GGAATGCGCCGTCAGGCTTGTGAAAGCGGGGGCCCAGATTGTTGGCATTAACTGTCACTTTGACCCGGAGACCTGCGTGAAGA
CTGTGAAGCTGATGAAGGAAGGAGTGGAGAAGGCCGGTCTGAAGGCCCACTACATGTCCCAGCCGCTGGGCTTCCACACCCCC
GACTGCAACCGACAGGGATTCATCGATCTTCCCGAGTTCCCGTTTGCTCTGGAGCCCAGGATCCTGACCAGGTGGGACATGCA
TAAATATGCCCGGGAAGCATACAACGCCGGCATTCGCTTCATCGGCGGCTGCTGCGGATTCGAGCCTTACCACGTCCGCGCTC
TGGCTGAGGAGCTGTCGCCTGAGCGGGGTTTCCTGCCCGCTGGCTCCGAGAAGCACGACAACTGGGGCAGCGGCTTGGCGCTG
CACACGAAGCCCTGGGTCAGAGCTAGGGCCCGTCGAGACTTCTGGGAGACACTAAAGCCCGCCTCCGGGCGTCCATTCTGTCC
ATCCATGTCCAAGCCCGATGACTGGGGCATCACCAAAGGACACGCCGACCTGATGCAACAAAAAGAGCCCACCAGCAAGGAGC
AGCTCAACGCTCTTTTTGAGAAGGCCACCCACAATGCCTGA

>FheBHMT5 (mummichog; Accession number XM_012870075)
ATGGCACCAACAAAGAAGGGCATCTTGGAGCGTCTCAACGCTGGAGAGGTCGTCATTGGCGATGGAGGGTTTGTGTTTGCCCT
CGAGAAGAGAGGATACGTGAAGGCTGGACCCTGGACTCCCGAGGCTGCCGCAGAGTACCCAGACGCTGTGCGTCAGCTGCACA
GGGAGTTCCTGCGAGCTGGAGCCAACGTCATGCAGACCTTCACCTTCTATGCCAGTGATGACAAGTTGGAGAACAGAGGCAAC
AAGGTCACCTTCACTGGAGCTCAGGTCAACGCGGCCGCCTGCGATCTGGCCCGGGAGGTTGCCAATGAGGGTGACGCTCTGGT
GGCGGGAGGCGTGTCCCAGACTCCGTCCTACCTGAGCTGCAAGAGTGAGACGGAGGTGAAAGCCATCTTCAAGAAGCAGCTGG
ACGTGTTCACCAAGAAGAATGTCGACTTCCTGATTGCTGAGTACTTTGAGCATGTTGAAGAGGCCGTGTGGGCCGTGGAAGTC
CTGAAGGAGACGGGGAAGCCAGTGGCTGCCAGTCTTTGCATCGGACCTGAGGGAGACATGCACGGCGTCTCACCCGGAGAGTG
TGCCGTCAGGCTCGTCAAAGCTGGTGCCCAGATCGTGGGAGTCAACTGCCACTTTGACCCCCTGACCTGCGTGAAGGCCGTCA
AGAAGATGAAGGAGGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTGCAGCCGCTGGCTTTCCACACTCCCGACTGC
AACTGCCAGGGATTCATCGATCTGCCAGAATTCCCCTTCGGTCTAGAGCCCAGGATCCTGACCCGTTGGGACATGCACCAGTA
CGCCAGGGAGGCATACAACGCCGGCATCCGCTTCATCGGGGGCTGCTGTGGGTTTGAGCCTTACCACATCAGGGCCGTGGCCG
AGGAGCTGGCTCCAGAGAGGGGCTTCCTCCCTGCTGCGTCAGAGAAACACGGAATGTGGGGTGCTGGTCTGGAGATGCACACC
AAGCCCTGGGTCAGAGCCAGAGCCCGCCGTGATTACTGGGAGAACCTGAAGCCAGCCTCTGGCCGTCCCTTGTGTGCGTCCAT
GTCGACCCCAGACAAATGGGGCGTTACCAGGGGCCACGCTGAACTCATGCAGCAGAAGGAGGCCACCTCCAAGGACCAACTCA
AGCAGCTGTTTGACAGGTCAAAGACCCACTGA

>GacBHMT3 (three-spined stickleback; Accession number ENSGACT00000020863)
ATGGCACCAGGCGCAGGGAAGGGCATTCTGGCGCGTCTGGACGCGGGGGAGGTCGTCATCGGGGACGGAGGCTTCGTCTTCGC
CCTGGAGAAGAGGGGTTACGTCAAGGCAGGGCCGTGGACCCCCGAGGCCGCAGCCGAGCACCCAGAAGCAGTGCGGCAGCTGC
ACAGAGAGTTCCTGAGAGCGGGCTCCGATGTCATGCAGACGTTCACTTTCTACGCCAGCGACGACAAACTGGAGAACAGGGGC
CACACTCAGCGCTTCACTGGGCAGCAGATCAACGAGGCTGCTTGTGACCTGGCCAGGGAGGTGGCCGATGAAGGTGGTGCTCT
GGTGGCGGGAGGAGTCTCTCAGACCCCCGCTTACCTCAGCTGCAAGAGCGAGGACGACGTGAAGGCCATCTTCAAGAAACAGA
TCGACGTCTTTGTCCAGAAGAACGTGGACTTCTTGATTGCAGAGTACTTTGAGCACGTGGAGGAGGCTGAGTGGGCCGTGCAG
GTTTTGAAGAGCACCGGGAAGCCCGTGGCCGCAACCCTGTGCATCGGCCCGGAGGGGGACCTGAACGGAGTCCGCCCCGGGGA
CTGTGCGGTCAGACTCGTCAAAGCCGGAGCTCACGTTGTGGGCGTCAACTGCCACTTTGACCCGGAGACCTGCGTGAAGACAG
TGAAGATGATGAAGGAGGCGGTGGAGAAAGCCGGGCTGAAGGCTCACTACATGAGCCAGCCGCTGGCGTACCACACTCCCGAC
TGCAACTGCCAGGGTTTCATCGATCTGCCAGAGTTCCCTTTCGGTCTGGAGCCGAGGGTCCTGACCAGGTGGGACATGCAGAA
GTACGCCCGCGAGGCCTACAAAGCCGGCATCCGCTACATCGGGGGCTGCTGTGGATTCGAGGCCTATCACGTCCGCGCCCTGG
CTGAGGAGCTGGCGCCGGAGAGGGGCGGCCTGCCCGCCGGCTCCGAGAAGCACGGCACCTGGGGCGCCGGCCTGGAGATGCAC
ACGAAGCCTTGGGTCAGAGCCAGGGCCCGTCGTGACTACTGGGAGAACCTGAAGCCCGCGTCCGGCCGTCCCCTCTGCCCCTC
CATGGCCACGCCCGACGCCTGGGGGGTCACCAAAGGTCACGCTGACCTGATGCAGCAGAAGGAGGCCACCTCTCAGGACCAGC
TGAAGGCTCTCTTTGACAAGGCCAACAAAAACCAGTGA

>GacBHMT4 (three-spined stickleback; Accession number ENSGACT00000008034)
ATGGCACCAGTCAAGAAGGGTATCATTGAGCGCCTGAATTCTGGGGAGGTGGTGATCGGTGATGGGGGTTTCGTCTTCGCTCT
GGAGAAGAGAGGCTACGTGAAGGCCGGGCCGTGGACCCCAGAGGCTACCGTCACACACCCCGAGGCCGTGCGGCAGCTGCACA
GGGAGTTCCTGAGGTCAGGATCCAATGTCATGCAGACCTTCACCTTCTACGCCAGCGACGACAAACTGGAGAACAGGGGGCAG
AACTTGAGAATCACTGGGGCACAGATCAACGAGGCAGCCTGTGACCTGGCGAGGGAAGTAGCCAGTGAAGGAGACGCGATGGT
GGCGGGTGGCGTGTGTCAGACTCCGTCCTACCTGAGCTGCAAGAGTGAGACAGATGTGAAGGCCATCTTCAAGAAACAGCTGG
ACGTGTTCATGAAGAAGAACGTGGACTTCCTGATAGCCGAGTACTTCGAGCATGTGGAGGAGGCCGTGTGGGCCGTAGAGGTG
CTGAAGACCAGCGGGAAGCCCGTGGCCGCCTCCATGTGCATCGGACCGGAGGGGGACATGCACGGCGTCTCGCCCGGCGACTG
TGCTGTCAGGCTGGTGAAGGCCGGTGCCCAGATTGTGGGAGTCAACTGCCACTTTGACCCCATGACCTGTGTGAAAGCCGTGA
AGATGATGAAGGAGGGGGTGGAGAAGGCCAGGCTGAAGGCCCACTACATGGTGCAGCCCCTGGCATACCACACCCCCGACTGC

AACTGTCAGGGATTCATCGACCTGCCAGAATTCCCATTCGCCCTTGAGCCCAGAATCCTGACCCGCTGGGACATGCACGAGTA
CGCCAGAGAGTCCTACAAGGCTGGCATCCGCTTCATCGGCGGCTGCTGCGGATTCGAGCCGTACCACATCAGGGCCATAGCAG
AGGAGCTGGCCCCCGAGAGAGGGATCGTGGCCCCTGCCTCAGAGAAACATGGAATGTGGGGTGCCGGTCTGGAGATGCACACC
AAGCCCTGGGTCCGAGCCAGGGCCCGCCGGGACTACTGGGAGAATATAGCTCCTGCATCCGGGCGGCCCAAGTGCCCATCCAT
GTCCACGCCGGAGGGCTGGGGCGTGACCAAGGGCCACGCTGACCTGCTGCAGCACAAAGAGGCCACCAGCAGTCAGGAAATGA
AGCACGTGCTGGAGATGCAGAAGAAGTCCAAGTCCTCCACATGA

>GacBHMT6/7 (three-spined stickleback; Accession number ENSGACT00000013149)
ATTGTACCATGTAATATTTGTAAAGGTATCTTGGAGCGACTAAATGCTGGGGAGGTGGTTGTAGGTGATGGAGGTTACGTGAT
GCAGCTGGAGCGCCGTGGCTATGTGAAGGCTGGACACTGGACACCTGAGGCGGCTGTAGAACATCCTGAAGCAGTGAGGCAGC
TGCACAGGGAGTTCCTGAGAGCAGGAGCCAATGTGATTCAGACGTTCACCTTCTACTGCAGCGAGGATAAACTGGAGATCAGC
GGCAATGTCCCCAACATTACCGGGGCTCAGATCAACGAGGCAGCTTGTGACCTGGCCAGGGAGGTCGCCAATGAGGGGGAGGC
CTTGGTCGCCGGCTGCGTGTCCAAGACTCCCTGTTATGTGGATACTCACAATGAGACTGAAGTCAAGGCCATCTTTAAGAAAC
AGATGGACGACTTCCTCAAGAAGGAGATTGATTTCTTTATTGCAGAGTTCTTTGAGCATGTGGAAGAGGCGGCGTGGGCCGTG
GAGGTGCTGAAGACCGGCGGTAAACCAGTGGGTGCAACACTGTGCATCTCCCCTGAGGGAGACATGCACGGAGTCCCGCCTGG
AGAGTGTGCTGTCAGGCTGGTCAGAGCCGGTGTAGGAGCTGACATTGTTGGAATAAACTGCCATTTGGATCCTCTGACGTGTG
TTCGCACGGTGAAGTTGATGAAAGCAGGATTAGAGAAAGCTGGTCTCAAAGCTCATCTCATGATCCAGCCTCTGGGCTTCCAC
ACACCTGAGTGCAACCACAGCGGATATCTCACGCTACCCGAGTTCCCGTTCGCATTGGAGACCAGAGCGATGACCCGCTGGGA
CGTCCATCAATACGCCCGAGAGGCTTACAACGCCGGCATTCGCTACATCGGTGGCTGCTGTGGATTCGAGGCCTACCACATCA
GAGCCATAGCAGAGGAGCTGGCTGCAGAGAGAGGATTCCTCCCACCGGCGTCAGAGAAGCACGGCCTGTGGGGAGCTGCTCTG
GAGATGCACACTAAACCGTGGGTCAGAGCCAGGTCTCGTCGAGAGTACTGGGAAAAGCTTTTACCTGCTTCTGGACGCCCCAA
ATGCCCCTCCATGGCTACACCAGCTGATGAATAA

>GmoBHMT3 (Atlantic cod; Accession number ENSGMOT00000009251)
AACATTCTGGAGCGTCTGGATGCAGGCGAGGTGGTGCTGGGGGACGGAGGCTATGTGTTTGAACTGGAGAAGAGAGGCTACGT
GAAGGCGGGGCCGTGGACCCCGGAGGCGGCCTCAGAACACCCTGAAGCAGTGCGGCAGCTGCATCGGGAGTTCCTCAGAGCAG
GAGCAGACATCCTCCAAACGTTTACCTTCTACGCCAGTGATGACAAACTGGAGAACAGGGGCAATGTGCAGAGTTTCACTGGA
AAGCAGATCAATGAGGCGGCGTGTGATCTGGCCAGGGAGGTGGCGGACGAGGGCGACGCCCTGGTGGCTGGTGGCGTCTCGCA
GACTCCATCTTACCTCAGCTGCAAGAGTGAGGCTGAAGTAAAGGCCATCTTCATGAAGCAGATCGAGGTCTATGTCCAGAGGA
AGGTGGACTTTCTCATTGCAGAGTACTTTGAGCATGTGGAGGAGGCTGAGTGGGCCGTCCAGGTCCTGAAGACCACAGGGAAG
CCCGTGGCCGTGACGCTGTGTATCGGACCGGAAGGTGACCTGAATGGGGTCAGCCCTGGAGACTGTGCAGTCAGACTGGTGCA
AGCAGGAGCTGACATAGTGGGGGTCAACTGCCACTTTGACCCGGTGACCGCTGTGAAGACGGTGAAGATGATGATGGAGGGGG
TGGGGAAGGCTGGGCTGAAGGCTCACTACATGGTCCAGCCCCTGGCCTACCACACCCCCGACTGCAATTGCCAGGGCTTCATC
GACCTCCCTGAGTTTCCCTTCGCCCTGGAGCCCAGGGTGCTGACCCGGTGGGACATGCAGAAGTACGCCCGGGACGCCTACGA
CTCTGGCATCCGCTACATTGGAGGCTGCTGTGGCTTCCAGCCCTACCACACGCGGGCCCTGGCAGAGGAGTTGGCGGCAGAGC
GGGGGTTCTTACCGGCTGCCTCAGAAAAGCACGGCTCCTGGGGCAGTGGCCTGGAGATGCACACCAAGCCCTGGGTCCGAGCC
AGGGCTCGGCGAGACTACTGGGAGAAGCTCCCGCCGGCGTCCGGCCGCCCCCACTGCGCCTCCATGTCCACTCCCGCCGGATG
GGGCGTCACCAAGGGCGACGCCCAGCTCATGCAGAAGAAGGAGTCCACCACTCAGGACGAGCTGAAGGTCCTGTTTGGGAAGT
CGGTGACCAATCAGTGA

>GmoBHMT4 (Atlantic cod; Accession number ENSGMOT00000007895)
ATGGCACCAGGCAAGAAGGGGATCCTTGAGCGTCTGAACGCTGGAGAGGTGGTGATTGGCGATGGAGGCTTTGTCTTCGCTTT
GGAGAAGAGGGGCTACGTCAAGGCTGGCCCCTGGACTCCTGAGGCTGCCCACACACACCCTGAGTCTGTCCGCCAACTGCACA
GGGAGTTCCTGAGGGCAGGATCAAACGTCATGCAGACTTTCACCTTCTACGCCAGTGATGACAAACTGGAGAACAGGGGTCAG
GCCCTGAAGATCACTGGTGCACAGGTGAACGAGGCTGCCTGCGACCTTGCCAGAGAGGTGGCTAATGAGGGAGATGCCCTGGT
GGCGGGCGGAGTCTCCCAGACCCCCTCCTACCTTAGTTGCAAGAGCGAGACTGAAGTCAAGGCGATCTTCAAGAAACAGCTGG
ATGTCTTCATGAAGAAGAATGTGGACTTCATGATTGCCGAGTACTTCGAGCATGTGGAGGAGGCTGAGTGGGCCGTCCAGGTG
CTGAAGGCCAGCGGCAAGCCGGTGGCTGCCAACATGTGCATCGGACCCGAGGGAGACATGCACGGCGTCTCCACAGGAGAGTG
TGCCGTGAGGCTGGTCAAGGCCGGGGCCCAGATTGTGGGAGTCAACTGCCACTTTGACCCCATGACCTGTGTGAAGGCTGTGA
AGATGATGAAGGATGCCGTGGAGAAGGCTGGTCTTAAGGCCCACTATATGGTGCAGCCTTTGGCCTTCCACACCCCTGACTGC
AACTGCCAAGGATTCATCGATCTGCCAGAGTTCCCCTTCGGACTGGAGCCCAGGATCCTGACCCGTTGGGACATGCATGCATA
CGCCAGGGAGGCCTACAACGTGGGCATCCGCTACATCGGGGGCTGCTGCGGCTTCGAGCCCTACCACATCAGGGCCGTGGCGG
AGGAGCTGGCCACCGAGAGGGGCATCCTACCCGCTGCCTCAGAGAAACACGGCATGTGGGGCGCCGGCCTGGAGATGCACACC
AAACCGTGGGTCAGAGCCAGGGCCCGTCGGGATTACTGGGAGCAGCTCAAGCCTGCGTCAGGACGTCCCCTGTGCCCGTCCAT
GTCCACGCCAGAGGGATGGGGCGTGACCAAGGGGCATGCAGACCTCCTGCAGCACAAGGAGGCCACCAGCAGCCAGGAGATGA
AGCACGTGGTGGAGCTGCAGAAGAAGGCCAAATCCACGTGA

>GmoBHMT5 (Atlantic cod; Accession number ENSGMOT00000007909)
ATGGCACCCTCAGGAAAAAAGGGTATCTTGGAGCGCCTTGCCGCTGGCGAGGTCGTCATTGGAGATGGTGGTTTTGTGTTCGC
CCTGGAGAAGAGGGGCTACGTCAAGGCAGGACCATGGACCCCTGAAGCCGCTGCAGAGCACCCAGAAGCAGTGCGACAGCTGC
ACAGGGAGTTCCTGAGGGCTGGGGCCAACGTTATGCAGACATTCACCTTCTACGCCAGTGATGACAAACTGGAGAACAGGGGC
AACCAGCTTGGCTTCACAGGAGCCCAAATCAACGAGGCGGCGTGTGACCTGGCTCGCGAGGTGGCCAATGAGGGTGACGGTCT
GGTGGCCGGGGGTGTCTGCCAGACTCCCTCTTATCTGAGTTGCAAGAGTGAAACGGAGGTGAAGGCCATCTTCAAGAAGCAGC
TGGACGTGTTTGTTAAGAAAGATGTGGACTTCCTGATCGCTGAGTACTTTGAGCATGTTGAAGAGGCGGAGTGGGCTGTACAG
ACGCTGAAGGAGACAGGGAAGCCGGTGGCCGCCTCTATGTGCATCGGACCGGACGGGGACATGCACGGGGTCACCCCCGGTGA
CTGTGCCGTCAGGCTGGTGAAGGCTGGTGCCGAGATTGTGGGAATCAACTGCCACTTTGACCCCATGACCTGTGTGAAGACTG
TGAAGATGATGAAGGATGCCGTGGAGAAGGCTGGTCTTAAGGCCCACTATATGGTGCAGCCTTTGGCCTTCCACACCCCTGAC
TGCAGCTGCCAGGGATTCATCGATCTGCCAGAGTTCCCCTTCGGACTGGAGCCCAGGATCCTGACCCGTTGGGACATGCACCA

APPENDIX

GTACGCCAGGGAGGCCTACAACGTGGGCATCCGCTTCATCGGAGGCTGCTGCGGCTTCGAGCCCTACCACATCAGGGCCGTGG
CGGAGGAGCTGGCCACCGAGAGGGGATGCCTGCCCGCCGCCTCAGAGAAACACGGCCTCTGGGGCGCCGGCCTGGAGATGCAC
ACCAAACCATGGGTCAGAGCCAGGGCTCGTCGTGACTACTGGGAGGGCCTGAAGCCAGCGTCCGGACGGCCCCTGTGCCCCTC
TATGGCCAGCCCTGACGGCTGGGGCGTCACCAAGGGCCACGCCGACCTCATGCAGCAGAAGGAGGCCACTACCCAGGAGCAGC
TCAAGCCCCTGTTCGAGAAGGCCAAGACCCATTAA

>GmoBHMT6/7 (Atlantic cod; Accession number ENSGMOT00000011836)
GGAATTCTGGAGCGCCTGAATGCAGGTGAGGTGGTGATAGGCGACGGGGGCTATGTGATGCAATTGGAGAGGCGAGGCTATGT
GAAGGCTCCTCACTTCACCCCTGAAGCTGCCGTTGAGCATCCAGAAGCGGTGCGGCAGCTCCACAGGGAGTTCCAGAGGGCTG
GCGCCGATGTTCTTCAGACGCTAACCTTCTATTCCAGTGACGAAAAACTGAAGATCAGCGGCAACGTCACCGACATCACCGGA
ACACAGATCAATGAGGCAGCCTGCGACCTGGCCAGGGAGATCGCCAACGAGTCCGATGCATTGGTGGCCGCGGGGGTGTCCCA
GACGCCCGTCTACCCGGTGAATCGCAACCAGGCTGAGGTCAAAGCCATCATCCAGAAGCAGCTAGACGTCTTCAAAAGGAAAG
ACATCGACTTCTACATGGTGGAGTTCTTCAAGCACGTGGAAGAGGCCGTGTGGGCGGTGGAGGTGATGAAGGCCGGGGGAAAA
CCGGTGGGGGCGACCCTCTGCATCTCCCCCCAGGGAGACAAGGACGGGACCCCACCCGGAGAGTGTGCGGTCAGGCTGGTCAG
AGCCGGTACAGGGGCCGACATGGTTGGGATCAACTGCCACCTGGACCCCCTGACGTGTGTCCGTACGGTGAAGCTCATGAAGG
AGGGACTGGAGAAGGCCGGACTCAAAGCCCACCTGATGGTCCAGCCGCTGGGCTTCCACACTCCCGAGTGCAACCACAGCGGC
TACACCACCCTGGCCGAGTATCCCTTCGCCATGGAGGCCCGAGCTATGACGCGGTGGGACATGCAGAAGTATGCCAGGGAGGC
GTACAACGCTGGAATCCGCTACATCGGCGGCTGCTGTGGCTTCGAGCCCTATCACATCAGGGCCATCGCCGAGGAGATGTCCC
CCGAGAGAGGCGTCCTACCGGCGGCGTCCGACAAACATGGGCTCTGGGGCGCCGGCCTGGCAACTTACGGCACATCACTGAGA
GGTAAGGCTTGCCGTGAATATTGGGAAAATCTTCTGCCGGCTTCAGGGCGACCGAATTGTCCCTCCATGTCCAGACCA

>HbuBHMT5 (Burton's Mouthbrooder; Accession number XM_005919107)
ATGGCACCAGCAAAAAAGGGCATCCTTGAACGTCTCAATGCTGGTGAGGTGGTGATTGGTGATGGAGGATTTGTGTTTGCTCT
GGAGAAGAGGGGGTACGTAAAGGCCGGACCATGGACGCCTGAAGCTTCTGTACAGTACCCTGATGCAGTGCGACAGCTGCACA
GGGAGTTCCTGAGAGCCGGGTCCAACGTTATGCAGACCTTTACATTCTACGCCAGTGATGACAAACTTGAGAACAGAGGCAAC
AACCTCAAACTGACTGGAGCTCAGATCAATGAGGCTGCCTGTGATTTAGCCCGTGAGGTTGCCAATGAGGGTGATGCTCTTGT
TGCTGGAGGAGTGTCTCAAACTCCATCTTACCTGAGCTGCAAGAGCAAGGCAGATGTGAAGGCCATCTTCAGAAAACAGTTGG
ATGTGTTCATCAAGAAAAATGTGGACTTCCTGATTGCTGAGTACTTTGAGCATGTTGAAGAGGCCGTGTGGGCTGTGGAGGTG
ATGAAGGAGACAGGGAAGCCTGTGGCTACTTCTCTGTGTATTGGACCAAAGGGGGACATGCACGGTGTCTCACCCGGAGAGTG
TGCAGTCAGGCTGGTCAAAGCCGGTGCCCAGATTGTGGGAGTCAACTGTCACTTTGACCCCATGACCTGTGTGGAGACCGTGA
AGCAGATGAAGGAGGGAGTGGAGAAGGCTGGTCTGAAGGCTCACTACATGGTTCAGCCGCTGGCATACCACACTCCTGACTGC
AACTGCCAGGGATTTATTGACCTGCCAGAGTTCCCCTTCGGCCTGGAGCCCAGGATTCTGACTCGCTGGGACATGCACAAGTA
TGCCAGAGAAGCTTACAATGCTGGCATTCGCTTCATCGGTGGCTGCTGTGGCTTTGAGGCTTATCACATTAGGGCTATAGCAG
AGGAACTGGCACCTGAGAGAGGCTTCCTTCCAAGTGCATCAGAGAAACATGGCTCCTGGGGGTGCTGGTCTGGAAATGCACACC
AAGCCTTGGGTTAGAGCCAGAGCCCGACGTGATTACTGGGAGAACCTGAAGCCAACCTCCGGTCGTCCCCAGTGTCCCTCCAT
GTCTGTCCCTGATTCTTGGGGTGTTACCAAGGGCCATGCTGATCTGATGCAGCAGAAAGAGGCAACCTCTAAAGAAGAACTCG
AACAGCTGTTCGACAAATCAAAAACCCACTGA

>OlaBHMT3 (Japanese medaka; Accession number ENSORLT00000003020)
ATGGCACCATCCGGAACCAAGAAGAACATCCTGGAGCGCCTGGAGGCCGGAGAGGTCATCATCGGGGACGGAGGCTTCGTGGT
CGCTCTGGAGAAGAGGGGTTATGTGAAGGCGGGACCCTGGACCCCCGAGGCCGCCGCCGAGCACCCAGAGGCGGTGCGCCAGC
TGCACAGGGAGTTCCTGCGCGCGGGCTCCACGGTCATGCAGACCTTCACCTTCTACGCCAGCGACGACAAGCTGGAGAACCGA
GGCAACACGCAGCGCTTCACAGGGCAGCAGATCAACGAGGCCGCCTGCGACCTGGCCCGGGAGGTGGCCAACGAGGGGGACGC
TCTGGTGGCGGGGGGGGTCTCCCAGACCCCGTCCTACATGAGCTGCGAGAGCGAGCAGGCTGTGAAGGCCATCTTCAAGAAGC
AGATCGACGTCTTCATCAGGAAGGACGTGGACTTTCTGATCGCAGAGTACTTTGAGCACGTGGAGGAAGCAGAGTGGGCGGTT
CAAATCCTGAAGACCAGCGGGAAGCCCGTGGCTGCGACTCTGTGCATCGGACCGGAAGGGGACCTGAACGGCGTCAGCCCCGG
AGACTGCGCCGTCAGGCTGGTGAAGGCCGGCGCAGACATCGTGGGGGTCAACTGCCACTTTGACCCGGAGACCTGCGTGAAAA
CCGTGAAGCTGATGAAGGAGGGCGTGGAGAGGGCGGGGCTAAAAGCCCACTACATGAGCCAGCCGCTGGCCCTGCACACCCCC
GACTGCAATCGGCAAGGATTCATCGACCTGCCGGAGTTCCCGTTCGCTTTGGAGCCCAGAGTGTTGACCCGATGGGACATGCA
CAGGTACGCGCGGGGCGCCTACAACGCCGGCATCCGTTACATCGGCGGCTGCTGCGGGTTCGAGCCGTACCACATCCGGGCCG
TGGCCGAGGAGCTGGCGCCCGAGCGCGGGTTCCTGCCGCCCGCCTCCGAGCACGGCAGCTGGGGCGCCGGCCTGGAGATGCAC
ACCAAGCCGTGGGTCAGGGCCAGGGCTCGCCGTGAATACTGGGAGACGCTGAAGCCCGCATCTGGACGTCCATTCTGCGCCGC
CATGTCCAAACCTGACTGTTGGGGTATCTCCAGAGGCCACGCTGACCTGATGCAAACAAAGGAAGCCACCAGCGAGGCTCAGC
TCAAGGCTGTCTTTGACAAAGCCGCCCAGAGCCAATGA

>OlaBHMT5 (Japanese medaka; Accession number ENSORLT00000004116)
ATGGCACCAACTAAGAAGGGAATCTTGGAGCGTCTCAATGCTGGAGAGATTGTCATCGGTGATGGGGGCTTTGTGTTTGCCCT
GGAGAAAAGAGGTTACGTGAAGGCTGGTCCCTGGACCCCTGAAGCCGCTAAAGAGTACCCTGAAGCAGTGCGTCAGCTGCACA
GAGAGTTCCTGCGGGCTGGATCCAACGTCATGCAGACCTTCACCTTCTACGCCAGTGATGACAAGCTGGAGAACAGAGGCAAC
AAGCTCAGCTATACTGGCGCTCAGATCAATGAAGCCGCCTGTGATTTGGCTCGGGAAGTCGCCAATGAGGGCGATGCTCTGGT
TGCTGGTGGAGTGTCTCAGACTCCATCCTACCTGAGCTGCAAGAGCGAGAAAGACGTGAAGGCCATCTTCAAGAAGCAGCTGG
ATGTGTTTGTCAAGAAGAACGTGGACTTCCTGATTGCTGAGTACTTTGAGCATGTTGAGGAGGCCGTGTGGGCCGTGGAAATG
CTGAAGGAGACAGGGAAGCCTGTGGCTGCCACTCTGTGCATTGGACCACAGGGGGACATGCATGGTGTCTCACCACAGGAGTG
TGCCGTCAGGCTGGTCACAGCAGGTGCCCAGATTGTTGGAATCAACTGTCACTTTGACCCTGAGACCTGCGTGAAGACTGTGA
AGATGATGAAGGAGGGGGTGGAGAAGGCTGGGCTGAAGGCTCACTACATGGTGCAGCCTCTGGCATTCCACACCCCCGACTGC
AACTGCCAGGGCTTCATCGACCTTCCAGAATTCCCCTTTGGTCTTGAGCCCAGGATCCTGACCCGCTGGGACATGCACAAATA
CGCCAGAGAGGCTTATAACGCTGGAATTCGCTACATTGGTGGGTGCTGTGGATTTGAGCCCTATCACATCAGGGCTGTAGCAG
AGGAGCTGGCCACCGAGAGAGGGTTTTTCCCAGCTGCAACGGAGAAACACGGCCCCTGGGGATCAGGTCTGGAGATGCACACC

AAGCCCTGGGTCAGAGCCAGAGCTCGCCGTGATTATTGGGAGAACCTGAAGCCGTCCTCTGGTCGCCCCTTCTGCCCTTCCTT
GTCTGTCCCCGATGGCTGGGGCGTCACCAGGGGCCATGCCGAGCTCATGCAGCAGAAAGAGGCCACTTCCAAGGACCAACTCA
AACAGCTGTTCGACAGGTCAAAGACGCAATAA

>OmyBHMT3 (rainbow trout; Accession number FR908041)
ATGGCACCAGCTGGAGCTAAGAGAAACATTCTGGAGCGCCTGGATGCAGGCGAGATTGTTATTGGGGATGGAGGTTTCGTCTT
CGCCCTGGAGAAGAGGGGCTATGTGAAGGCGGGTCCCTGGACTCCTGAGGCTGCCGCCGAGCACCCCGAGGCTGTGAGACAGC
TGCACCGAGAATTCCTGCGGGCGGGGTCAAATGTCATGCAGACTTTCACCTTCTACGCCAGCGACGATAAACTGGAGAACAGG
GGTAACGCTCAGCGCTTCACTGGCACACAAATCAACGAGGCCGCCTGTGACCTGGCCAGGGAGGTGGCCAATGAGGGTGACGC
GATGGTGGCGGGCGGGGTCTCCCAGACCCCCTCTTACTTGAGCTGCAAGTGCGAGGACGAGGTCAAGGGCATCTTCAAGAGGC
AGCTTGATGTCTTCGTCAAGAAGAACGTGGATTTCATGATTGCAGAGTACTTTGAGCATGTGGAGGAGGCAGAGTGGGCCGTG
CAGGTGCTGAAGACCTCAGGCAAGCCAGTGTGTGCCTCTCTTTGTATCGGCCCTGACGGAGACCTCAACGGAGTCAGCCCTGG
AGACTGTGCCGTCCAGCTGGTCAAAGCTGGAGCCAATATTGTGGGCATCAACTGTCACTTTGACCCTATGACCTGTGTGAAGA
CGGTGAAAATGATGAAGGAGGGTGTGGAGAAGGTGGGGCTGAAGGCCCATTACATGGTCCAGCCTCTGGCCTTCCACACCCCT
GACTGCAACTGCCAGGGATTCATCGACCTGCCAGAGTTCCCCTTCGGTCTGGAGCCCAGGATTCTGACCAGGTGGGACATGCA
CAAGTACGCCAGAGAGGCCTTCAACGTTGGCATCCGTTTCATCGGTGGCTGCTGCGGCTTCGAGCCCTACCACATCAGGGCTG
TGGCTGAGGAGCTGGCCACAGAAAGAGGCTGCCTGCCCGCTGCCTCAGAGAAGCACGGTAACTGGGGTGCTGGTCTGGAGATG
CACACAAAGCCATGGGTCCGTGCCAGGGCTCGCCGTGACTACTGGGAGAAGTTGAAGCCAGCATCCGGCCGTCCCAGGTGTCC
CTCCATGTCCACCCCTGATAGCTGGGGTGTCACCAAGGGCCATGCCGACCTGATGCAGCACAAAGAAGCCACCAGCCAGGAAG
AGCTCAAGCCCCTGTTTGAGAAGGCCAAGGCCAGCCACTGA

>PchBHMT3 (Victoria big-lipped Hap; Accession numbers BJ670381, BJ679674, BJ685848,
BJ672792 and BJ682950)
ATGGCACCAGCAGGAGAAAAGAAGAACATTTTGAAGCGCCTGGATGCAGGAGAGGTCGTTATTGGGGACGGAGGCTTTGTCTT
CGCGCTCGAGAAGAGAGGTTATGTAAAAGCCGGGCCGTGGACACCTGAAGCTGCTGCCGAGCACCCCGAAGCGGTGCGGCAGC
TGCACAGAGAGTTCCTGAGAGCTGGTTCCAATGTCATGCAGACGTTCACTTTCTACGCAAGCGATGACAAACTAGAGAACAGG
GGCCAGACTCTGAGCTTCACCGGACGCCAAATAAACGAAGCAGCTTGTGACCAGGCCAGAGAGGTGGCCAATGAGGGTGACGC
TCTGGTGGCAGGTGGAGTCTCTCAGACCCCGGCATACCTCAGCGGCAAGACCGAGGAGGAGGTCAAGGCCATCTTCATTAAAC
AGATTGATGTCTTTGTCCAGAAAAATGTCGACTTCTTGATTGCAGAGTACTTTGAGCACGTGGAAGAGGCTGAGTGGGCAGTC
CAGGTTTTGAAGGCTTCCGGTAAGCCAGTCGCTGCAACTCTGTGCATTGGACCTGAAGGGGACCTGAACGGCGTCAGCCCTGG
AGACTGTGCAGTCAGACTTGTTAAAGCTGGTGCTGATATTGTTGGCATTAACTGCCACTTTGACCCAGAGACTTGTGTGAAGA
CAGTGAAACTGATGAAGGAGGGTGTGGAGAAGGCTGGGCTGAAGGCTCACTACATGTCCCAGCCACTGGCCTTTCATACTCCA
GACTGCGGCTGCCAGGGATTCATCGACCTGCCGGAATTCCCGTTTGGTCTGGAGCCAAGGATCCTGACCAGAATGGACATGCA
TAAATATGCCCGGGAAGCATATAATGCTGGGATTCGCTACATTGGAGGCTGCTGCGGATTTGAACCGTATCACATCCGTGCCT
TGGCTGAGGAGCTGGAAGCGGAGCGAGGTTGTCTTCCTGCTGGATCTGAGAAGCATGGCAGCTGGGGTAGTGGCCTGTCATTG
CACACCAAGCCTTGGGTTAGAGCTAGGGCCCGCCGTGACTACTGGGAGAAGCTGAAGCCCGCCTCTGGCCGCCCCTTCTGCCC
CTCCATGTCCACACCTGATAGCTGGGGAGTCACCAAAGGCCACGCTGACCTGATGCAGCAAAAGGAGGCTACCTCCAAGGAGC
AGCTGGAGGCTCTGTTTGACAAGGCCAGCCAAGGCCACTGA

>PflBHMT4 (yellow perch; Accession number EU081830)
ATGGCACCTGGCAAGAAGGGTATCATTGAGCGTCTGAATGCCGGGGAGGTTGTGATCGGCGATGGAGGCTTCGTGTTTGCTCT
GGAGAAGAGAGGCTATGTGAAGGCCGGGCCATGGACGCCCGAGGCTTCTGTCACACACCCTGAGGCCGTGCGGCAGCTGCACA
GGGAGTTCCTGAGGGCAGGATCTAATGTCATGCAGACATTCACCTTCTATGCCAGCGATGACAAACTGGAGAACAGGGGTCAG
AACTTGAGACTCTCTGGGGTACAAATCAACGAGGCCGCCTGTGACCTGGCAAGGGAAGTAGCCAGCGAAGGAGACGCAATGGT
AGCTGGTGGAGTGTGTCAGACTCCATCCTACCTGAGCTGCAAGAGTGAGATAGAAGTGAAGGCCATCTTCAAGAAACAGCTGG
AAGTGTTTATGAAGAAGAATGTAGATTTCCTGATTGCTGAGTACTTTGAGCATGTTGAGGAGGCGGAGTGGGCCGTGCAGGTG
CTGAAGACCAGTGGAAAGCCTGTAGCTGCTTCTTTGTGCATCGGACCAGAGGGAGACATGCACGGCGTCTCACCTGGAGAGTG
TGCTGTCAGGCTGGTGAAGGCCGGTGCCCAGATTGTTGGAGTCAACTGCCACTTTGACCCAATGACCTGCGTGAAGACTGTTA
AGATGATGAAGGAGGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTACAGCCCCTGGCATACCACACTCCTGACTGC
AACTGTCAGGGATTCATCGATCTGCCAGAATTTCCCTTCGCCCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACACGTA
TGCCAGAGTGGCCTACAAGGCTGGCATCCGCTTCATTGGTGGCTGCTGTGGGTTTGAGCCTTATCACATCAGGGCGATAGCGG
AGGAGCTGAGCATCGAGAGAGGGATAATGCCCCCTGCCTCTGAGAAACATGGGATGTGGGGTGCTGGTCTGGAGATGCACACA
AAACCCTGGGTCAGAGCCAGGGCCCGTCGTGACTACTGGGAGCATCTCTCTCCCGCATCTGGCCGTCCCACTTGCCCATCCAT
GTCCACACCAGAGGGCTGGGGTGTGACCAAGGGCCACGCTGACCTGCTGCAGCACAAAGAGGCCACCAGCACCCAGGAAATGA
AGCACGTGCTGGAGATGCAGAAGAAGGCCAAGTCCTCTGCATGA

>PlfBHMT3 (European flounder; Accession number FN432389)
ATGGCACCTGGCAAGAAGGGTATCATTGAACGGCTGAATGCCGGGGAAGTGGTGATTGGCGATGGAGGCTTCGTGTTTGCACT
GGAGAAGAGGGGATATGTGAAGGCTGGGCCCTGGACTCCTGAAGCCACTGTCACTCATCCTGAAGCAGTGCGACAGCTGCACA
GGGAGTTCCTCAGGGCCGGATCCAATGTCATGCAGGCATTTACGTTCTACGCCAGTGATGACAAACTGGAGAACAGGGGTCAG
ACACTCAGTATCACTGGAGTACAAGTCAACGAGGCCGCCTGTGACCTGGCAAGGGAGGTCGCCAACGAGGGCGATGCTCTGGT
CGCTGGTGGAGTGTGTCAGACTCCGTCCTACCTGAGCTGCAAGAGTGAGACGGAAGTGAAGGGCATCTTCAAGAAACAGCTGG
AAGTGTTCGCAAAGAAGAACGTGGACTTCCTTATTGCTGAGTACTTTGAGCACGTTGAGGAGGCGGAGTGGGCCGTGCAGGTG
CTGAAGACCAGCGGGAAGCCTGTGGCTGCCTCTCTGTGCATCGGGCCGGAGGGGGACATGCACGGCGTCTCACCTGGAGAGTG
TGCCGTCAGGCTGGTGAAGGCTGGTGCCCAGATCGTTGGCGTCAACTGCCACTTTGACCCCATGACCTGCGTGAAGGCTGTGA
AGTTGATGAAGGAGGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTGCAGCCACTTGCCTTCCACACTCCCGGCTGC
AACTGTCAGGGATTCATTGACCTACCAGAATTCCCCTTTGGCCTGGAGCCAAGGATCCTGACCCGCTGGGACATGCACAAGTA
CGCCAGAGAGGCCTACAAGGCTGGAATCAAGTTCATTGGAGGCTGCTGTGGCTATGAGCCTTACCACATCAGAGCTGTGGCAG

APPENDIX

AGGAGCTAGCCCTGGAGAGAGGGGTGATGCCCCCCGGCTCGGACAAACATGGAATGTGGGGCGCTGGTCTGGAGATGCACACC
AAGCCCTGGGTCAGAGCCAGGGCCCGCCGTGACTACTGGGAGCATCTCCTGCCCGCATCCGGCCGTCCCCTGTGCGCGTCCCT
CTCCACACCAGAGAGCTGGGGTGTGACCAAGGGCCACACTGACCTGCTGCAGCACATAGAGGCAACCAGCACCCAGGAGATGA
AGCATGTGCTGGAAAAGCAGAAGAAGGCCAAGTCCTCTGCATGA

>PmaBHMT4 (red seabream; Accession numbers DC608846, DC606161, DC606490, DC607107, DC606070, DC607807 and DC607837)
ATGGCACCTGTAAAGAAGGGAATCCTTGAGCGTCTGAATGCCGGGGAGGTGGTGATTGGTGATGGAGGCTTCGTGTTTGCCCT
GGAGAAGAGGGGCTACGTGAAGGCCGGGCCATGGACTCCCGAGGCTGCTGTTACACACCCTGAGGCTGTGCGGCAGCTGCACA
GGGAGTTCCTGAGGGCAGGAGCTAACGTCATGCAGACGTTCACCTTCTACGCCAGTGATGACAAACTGGAGAACAGGGGTCAG
ACTCTGAAATACACTGGGGCACAAATCAACGAGGGGGCCTGTGACCTGGCAAAAGAAGTGGCCAGTGAAGGAGACGCTCTGGT
AGCTGGTGGAGTGTGTCAGACTCCATCCTACCTGAGCTGCAAGAGTGAGACGGAGGTGAAGGCCATCTTCAAGAAGCAGATGG
AAGTGTTCATCAAGAAGAACGTGGACTTCCTGATTGCTGAGTACTTTGAGCACGTTGAGGAGGCCGAGTGGGCAGTTCAGGTG
CTGAAGACCAGCGGAAAGCCTGTGGCTGCTTCTCTGTGCATCGGACCGGAAGGAGACATGCACGGCATCTCACCCGCAGAGTG
TGCTGTCAGGCTGGTGAAAGCTGGCGCCCAGATCATTGGAGTCAACTGCCACTTTGACCCCATGACCTGTGTGAAAGCTGTCA
GGATGATGAGGGAGGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTGCAGCCCCTGGCATTCCACACCCCCGACTGC
AACTGTCAGGGATTCATCGATCTGCCAGAATTCCCCTTCGGTCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACAAGTA
CGCCAGAGAGGCCTACAAGGCTGGCATCCGCTTCATTGGTGGCTGCTGTGGCTTTGAGCCTTATCACATCAGGGCTGTGGCAG
AGGAGCTGGCCCCCGAGAGAGGGATAATGCCCCCTGGCTCAGAGAAACATGGAATGTGGGGTGCTGGTCTGGAGATGCACACC
AAGCCCTGGGTCAGAGCCAGGTCCCGTCGTGAGTACTGGGAGAATCTGATGCCCGCATCCGGTCGTCCCAAGTGCCCATCCCT
GTCCACGCCAGAGTGCTGGGGTGTGACCAAGGGCCACGCTGACCTGCTGCAGCACAAAGAGGCCACCAGCAGCCAGGAGATGA
AGCACGTGCTGGAGATGCAGAAGAAGGCCAAGTCCACTGCATAA

>PolBHMT3 (bastard halibut; Accession number EF198069)
ATGGCACCAGCTGGAGCAAAGAGGGGCGTTCTGGAGCGTCTGGATGCGGGAGAGGTCGTCATCGGGGACGGAGGCTTCGTCTT
CGCCCTGGAGAAGAGGGGCTACGTGAAGGCAGGGCCGTGGACACCTGAGGCTGCAGCCGAGAACCCCGAGGCGGTTCGGCAGC
TGCACAGGGAGTTCCTGAGAGCAGGCTCCACTGTCATGCAGACATTCACCTTCTATGCAAGCGATGACAAACTGGTGAACAGA
GGCCACACTCAGCGCTTCACTGGGCAGCAAATAAATGAAGCAGCATGTGACCTGGCCAGAGAGGTGGCCAATGAGGGGGATGC
TCTGGTGGCAGGAGGCATCTCTCAGACTCCCGCTTACCTCAGCCGCAAGAATGAGGGGGAAGTCAAGGCCATCTTTAAGAAGC
AGATCGACGTGTTCATCCAGAAAAATGTGGACTTCTTGATCGCAGAGTACTTTGAGCATGTGGAAGAGGCTGAGTGGGCAGTC
CAGGTTTTAAGGACCACCGGCAAGCCAGTGGCTGCAACTCTCTGTATTGGACCTGAAGGAGACCTGAATGGAGTCAGCCCTGG
AGACTGTGCCGTCAAACTTGTCAAAGCTGGAGCTCACATTGTGGGCATTAACTGCCACTTTGACCCGGAGACCTGCGTGAAGA
CTGTGAAGATGATGAAGGCCGGAGTGGAGAAGGCCGGCCTGAAGGCTCACTACATGAGCCAGCCGCTGGCCTTCCACACTCCT
GACTGCAACCGCCAGGGTTTCATTGATCTGCCCGAGTTCCCCTTCAGTCTGGAGCCGAGGATCCTGATCAGATGGGACATGCA
GAGATACGCCCGGGAAGCGTACATTGCTGGCATCCGCTATATCGGAGGCTGCTGCGGATTTGAACCCTATCACATCCGTGCCC
TGGCTGAGGAGCTGGCACCTGAGAGGGGTTTCCTGCCCGTTGGGTCCGAGAAGCACGGTAACTGGGGAAGTGGCCTGGAAATG
CACACAAAGCCGTGGGTTAGAGCGAGGGCTCGTCGTGACTACTGGGAGAAGCAGAAACCTGCGTCTGGCCGTCCCTACTGCCC
GTCCATGTCCTCACCTGACGGCTGGGGCGTCACCAAGGGCCACGCTGACCTGATGCAGCAGAAGGAGGCCACCTCCAAGGAGC
AGCTGAAGGTTCTCTTTGACAAGGCCAACAAGTGCCACTAA

>PprBHMT3 (fathead minnow; Accession numbers FJ030935, DT240264, DT361111, DT361842, DT361814, DT107683 and DT358512)
ATGGCACCAGTCGGGTCCAAGAGGGGTATTCTGGAGCGCCTGGGCGCGGGTGAGGTGGTGATTGGAGACGGAGGGTTCGTTTT
CGCCCTAGAGAAGCGAGGATACGTGAAGGCCGGACCCTGGACACCAGAGGCCGCAGCTGAGCACCCGGAGGCCGTGCGACAGC
TGCACAGAGAGTTCCTGCGGGCCGGCTCTAATGTCATGCAGACCTTCACTTTCTACGCCAGCGACGACAAGCTGGAGAACAGA
GGGAACAAGCTGACGTACACCGGCCAGCAGATTAACGAGGCGGCGTGTGACCTCGCCAGAGAAGTGGCCAATCAGGGCGATGC
GCTGGTGGCGGGCGGAGTCTCGCAGACACCCTCTTACCTGAGCTGCAAGAGCCAAGAAGAGGTCACCAAGATCTTCAAGAAGC
AGCTCGACGTCTTCATCAAGAAGAATGTGGACTTTCTGATCGCAGAGTACTTCGAGCACGTGGAGGAGGCCGAATGGGCCGTC
CAGGCTCTGAAGGCGACAGGGAAGCCTGTAGCGGCCACGATGTGTATAGGACCTGAAGGAGACATGCACGGCGTGAGCCCGGG
AGATTGTGCCGTCAGACTGGTCAGAGCAGGTGCTGATATCGTGGGTGTGAACTGCCACTTTGACCCCATGACCTGTGTGAAGA
CCGTGGCCCTCATGAAGGCGGCCGTGGAGAAAGCCGGTCTGAAAGCGCATTATATGGCGCAGCCGCTGGCATACCACACACCC
GACTGCAGCTGCCAGGGCTTCATCGACCTGCCCGAGTTCCCCTTCGCCCTGGAGCCGCGGATCCTGACGCGCTGGGAGATGCA
GCAGTACGCCAGGGAGGCCTATAAGGCGGGCATTCGCTTCATCGGCGGCTGCTGCGGGTTCGAGCCATATCACATCCGCGCAG
TAGCCGAGGAGCTCGCGGTCGAGAGAGGAATTCTGCCCGAGGCCTCAGAGAAACACGGCATGTGGGGCGCCGGCCTGGAGATG
CACACCAAACCCTGGGTCAGGGCCAGGGCCCGTCGTGATTACTGGGAGAAGCTGAAGCCCGCGTCGGGACGCCCGCTGTGCCC
GTCCATGTCCACGCCCGACGGCTGGGGGGTCACCAAAGGTCACGCCGAGCTCATGCAGCAGAAGGAGGCCACGACCGAAGACC
AGCTGCGCCCGCTCTTCCAGAAAGCCGACGCCAAAAACTGA

>SauBHMT4 (gilthead seabream; Accession number DQ470488)
ATGGCACCTGTGAAGAAGGGTATCCTTGAGCGTCTGAATGCCGGGGAGGTGGTGATTGGGGATGGCGGCTTCGTGTTTGCCCT
GGAGAAGAGGGGCTATGTGAAGGCCGGACCATGGACCCCTGAGGCTACTGTCACACACCCTGAGGCTGTGCGGCAGCTGCACA
GGGAGTTCCTGAGGGCAGGAGCTAACGTCATGCAGACGTTCACCTTCTACGCCAGTGATGACAAACTGGAGAACAGGGGTCAG
TCTCTGAAATACACTGGGGCACAAATCAACGAGGCGGCCTGTGACCTGGCAAAGGAAGTGGCCAGCGAGGGAGACGCTCTGGT
AGCTGGTGGAGTGTGTCAGACTCCATCCTACCTGAGCTGCAAGAGTGAGACGGAGGTGAAGGCCATCTTCAAGAAGCAGATGG
AAGTGTTCATCAAGAAGAACGTGGATTTCCTGATTGCTGAGTACTTCGAGCACGTTGAGGAGGCGGAGTGGGCCGTTCAGGTG
CTGAAGACCGGCGGAAAGCCTGTGGCTGCTTCTATGTGCATCGGACCGGAAGGAGACATGCACGGCGTCTCACCCGCTGAGTG
TGCCGTCAGGCTGGTGAAGGCTGGTGCCCAGATCATTGGAGTCAACTGCCACTTTGACCCCATGACCTGCGTGCAAGCTGTCA
AGATGATGAAGGCCGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTGCAGCCCCTGGCATACCACACCCCCGACTGC

AACTGTCAGGGATTCATCGATCTGCCAGAATTCCCCTTTGCTCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACAAGTA
CGCCAGAGAGGCCTACAAGGTTGGCATCCGCTTCATCGGTGGCTGCTGTGGGTTTGAGCCCTATCACATCAGGGCTGTGGCGG
AGGAGCTGGCCCCCGAGAGAGGGATAATGCCCCCTGGCTCAGAGAAACATGGAATGTGGGGTTCTGCTCTGGAGATGCACACA
AAGCCCTGGGTCAGAGCCAGGTCCCGCCGTGAATACTGGGAGAATGTTATGCCCGCATCCGGTCGCCCCAAGTGCCCATCCCT
GTCCACGCCAGAGTGCTGGGGTGTGACCAAGGGCCACGCCGACCTGCTGCAGCACAAAGAGGCCACCAGCACCCAGGAGATGA
AGCACGTGCTGGAGATGCAGAAGAAGGCCAAGTCCTCTGCATAA

>SauBHMT6 (gilthead seabream; Accession number GU197551)
ATGGGGAGCAAGAGGAAGGACATCCTGGAGCGACTGAATGCCGGGGAGGTGCTTGTAGGAGATGGGGGTTATGTGATGCAACT
AGAGCGACGAGGTTATGTGAAAGGTGGTAACTGGACACCTGAAGCAGCCGTTGAACACCCTGAAGCAGTGCGGCAGCTGCACA
GGGAGTTTCTGAGAGCAGGAGCCAATGTGATTCAGACATTCACCTTCTACTGCAGTGAGGATAAACTGGAGTTGAGCCATGTC
ACCGGCATCACCGGGGCGCAGATCAATGAGGCAGCCTGTGACCTGGCCAGGGAGGTCGCGAATGAGGGCAATGCATTGGTTGC
CGGGGGTGTTTCCATGACTCCCTGCTATGCGAAGTCTCACAGTGAGACTGAGGTCAAGGCCATCTTTAAGAAACAGATGGATG
ACTTCCTCAAGAAGGACATTGATTTCTTTATAGCTGAGTTCTTCGACCACGTGGAAGAGGCAGTGTGGGCAGTGGAGGTGTTG
AAGACGAGCGGTAAAACGGTGGGTGCAACGCTGTGCATCTCCCCTCATGGAGACATGAAAGGGGTCCCACCTGGAGAGTGCGC
TGTCAGGCTGGTCAAAGCTGGAGCTGACATTGTTGGAATAAATTGCCACTTGGACCCTATGACTGGCGTTCGTACAGTGAAGC
TGATGAAAGAGGGATTAGAGAAAGCAGGTCTCAAAGCCCATCTCATGGTCCAGCCGCTGGGCTTTCACACACCTGAGTGCAAC
CACACTGGATACCTTGCCCTTCCAGAGTTCCCCTTCGCACTGGAAACCAGAGCAATCACCCGCTGGGATATACATAAATACGC
CAGAGAGGCTTATAATGCAGGAATTCGCTACATCGGAGGCTGCTGTGGATTTGAGCCCTATCATATCAGAGCCATAGCAGAGG
AGCTGGCTGCAGAGAGAGGATTCCTCCCACCAGCTTCAGAGAAGCACGGACTCTGGGGAGCTGCTCTGGAGATGCACACTAAA
CCCTGGGTCAGAGCCAGGGCTCGTCGAGAGTATTGGGAAAACCTTTTGCCTGCTTCTGGACGTCCCAAATGCCCTTCCATGGC
CACCCCAGCTGATGAATAA

>SauBHMT3 (gilthead seabream; Accession number GU119905)
ATGGCACCAGCTGGACAGAAGAAGGGTATCCTGGAGCGTCTGGATGCGGGAGAGGTCGTCATCGGGGACGGAGGCTTTGTCTT
CGCCCTGGAGAAGAGAGGCTACGTCCAAGCAGGACCGTGGACACCTGAGGCCGCAGCCGAGCACCCTGAAGCGGTGCGCCAGC
TGCACCGCGAGTTCCTGCGTGCGGGCGCCAACGTCATGCAGACCTTCACTTTCTATGCCAGCGACGACAAACTGGAGAACAGG
GGCAACAAACAGCGCTTCACTGGGCAGCAGATCAACGAGGCAGCTTGTGATCTGGCCCGGGAGGTGGCCAATGAGGGAGATGC
TCTGGTGGCGGGCGGAGTCTCTCAGACGCCCTCCTACCTGAGCTGCAAGAGCGAGGATGAGGTCAAAGCCATCTTCATGAAAC
AGATCGAGGTCTTCGTGAAGAAAAACGTGGACTTCCTGATCGCAGAGTACTTTGAGCACGTGGAAGAGGCCGAGTGGGCGGTC
CAAGTCCTGAAGACCACCGGGAAGCCAGTGGCTGCGTCTCTGTGTATCGGACCTGATGGAGATCTGAACAGAGTCAGCCCCGG
AGACTGTGCCGTCAGACTCGTCAACGCTGGAGCTCACATCGTGGGCATCAACTGCCACTTCGACCCGGAGACGTGTCTGAAGA
CGGTGAAGCTGATGATGGAGGGCGTGAAGAAGGCCGGGCTGAAGGCTCACTTCATGATCCAGCCGCTGGCGTACCACACTCCT
GACTGCGGCTGCCAGGGCTTCATCGACCTGCCCGAGTTCCCCTTCAGTCTGGAGCCGAGGATCCTGACCAGATGGGACATGCA
GAAATACGCCCGGGAAGCGTACAACGCCGGCATCCGTTACATCGGAGGCTGCTGTGGGTTTGAGCCCTATCACACCCGGGCGC
TGGCCGAAGAGCTGGCGCCCGAGAAGGGTTTCCTGCCCGCCGGCTCCGAGAAGCACGGCAACTGGGGCAGCAGCCTGGCCCTG
CACACCAAGCCCTGGGTCAGAGCCAGGGCCCGTCGTGACTACTGGGAGACGCTGAAGCCGGCGTCCGGACGTCCCATGTGCCC
CTCGATGTCGACGCCCGACAGCTGGGGCGTCACCAAAGGCCACGCCGACCTGATGCAGCAGAAGGAGGCGACCTCCCAGGAGC
AGCTGAAGGCTCTGTTTGACAAGGCCGCCCACAGCCACTGA

>SauBHMT7 (gilthead seabream; Accession number GU197552)
ATGGAGAGCAAGAAGAGAAGGGGTATCTTGAAGCGCCTGAACGCTGGGGAGGTGATTGTAGGGGACGGAGGTTATGTAATGCA
ACTGGAGCGACGTGGCTATGTGAAGGCGGGACACTGGACCCCTGAGGCTGCTGTCGAACACCCTGAAGCAGTGCGGCAGCTGC
ACAGGGAGTTTCTGAGAGCAGGTGCCAATGTGATTCAGACATTCACCTTCTACTGCAGTGAGGATAAACTGGAGATCAGTGGT
AACGTCCCCAATATCACTGGGGCGCAGATCAATGAGGCGGCCTGTGACCTGGCCAGGGAGGTAGCAAATGAAGGTGGTGCATT
GGTTGCCGGGTGTGTGTCTAAGACTCCGTGTTATGTGACGAGCCACAGTGAGACTGAAGTCAAGGCCATCTTTAAGAAACAGA
TGGATGACTTCCTCAAGAAGGACATTGATTTCTTCATAGTAGAATTTGTTGATCACGTGGAAGAGGCAGTGTGGGCAGTGGAG
GTGTTGAAGACAAGCGGTAAAACGGTGGGTGCAACGCTGTGCATCTCCCCTCATGGAGACATGAAAGGAGTCCCACCTGGAGA
GTGCGCTGTCAGGCTGGTCAAAGCTGGAGCTGACATTGTTGGAATAAATTGCCACTTGGACCCTATGACGTGCGTTCGTACAG
TGAAGCTGATGAAAGAGGGGTTAGAGAAAGCAGGTCTCAAAGCCCATCTCATGGTCCAGCCGCTGGGCTTTCACACACCTGAG
TGCAACCTAGGTGGATACACCAGCCTACCGGAGTACCCCTTCGCAATGGAAACCAGAGCAATCACCCGCTGGGATATACATGA
ATACGCCAGAGAGGCTTATAATGCAGGAATTCGTTACATCGGTGGCTGCTGTGGATTTGAGGCCTATCATATCAGAGCCATAG
CAGAGGAGCTGGCTGCAGAGAGAGGATTCCTCCCACCAGCTTCAGAGAAGCACGGACTCTGGGGAGCTGCTCTGGAGATGCAC
ACTAAACCCTGGGTCAGAGCCAGGGCTCGTCGAGAGTACTGGGAAAACCTTTTGCCTGCTTCTGGACGTCCCAAATGCCCTTC
CATGGCCCCCCCAGCTGATAATTATGGAAAATAG

>SauBHMT5 (gilthead seabream; Accession number GU597054)
ATGGCACCAGTAAAGAAGGGAGTCTTGGAGCGTCTTAACGCTGGGGAGGTCGTGATCGGCGATGGAGGCTTCGTGTTTGCCCT
GGAGAAGAGAGGCTACGTGAAGGCCGGTCCCTGGACCCCTGAGGCCGCCACAGAGCACCCTGAAGCAGTGCGACAGCTGCACA
GGGAGTTCCTGAGGGCCGGGGCCAACGTCATGCAGACCTTCACCTTCTACGCCAGCGATGACAAACTGGAGAACAGGGGCAAC
AAGCTCACCCACACTGGGGCTCAGATCAACGAGGCCGCCTGCGACCTGGCCCGTGAGGTTGCCAACGAGGGCGACGCCCTGGT
TGCCGGAGGAGTGTGTCAGACTCCATCCTACCTGAGCTGCAAGAGTGAGACGGAGGTGAAGGCCATCTTCAAGAAACAGCTGG
ATGTGTTCATGAAGAAGAACGTGGACTTCCTGATTGCTGAGTACTTTGAGCATGTTGAAGAGGCCGTGTGGGCCGTGCAGGTG
CTGAAGGAGACAGGGAAGCCCGTGGCCGCTAATCTGTGCATCGGACCAGAGGGAGACATGCACGGCATCTCACCCGCAGAGTG
CGCCGTCAGGCTGGTGAAGGCTGGTGCCGACATCATTGGAATCAACTGCCACTTTGACCCCATGACCTGCGTGCAAGCTGTCA
AGATGATGAAGGAGGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTGCAGCCCCTGGCATACCACACCCCCGACTGC
AGCTGTCAGGGATTCATCGATCTGCCAGAATTCCCCTTTGGTCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACAAGTA
CGCCAGAGAGGCCTACAATGTTGGCATCCGCTACATTGGTGGCTGCTGTGGATTCGAGCCCTATCACATCAGGGCTGTTGCAG

APPENDIX

AGGAGCTGGCCCCCGAGAGAGGCTTCCTCCCAGACGCATCAGAGAAACACGGCAACTGGGGTGCTGGTCTGGAGATGCACACC
AAGCCCTGGGTCAGAGCCAGAGCCCGTCGTGACTACTGGGAGAACCTGAAGCCAGCCTCTGGCCGTCCTCAGTGTCCCTCCCT
GTCCACCCCTGATGGCTGGGGTGTTACCAAGGGCCACAATGATCTCATGCAGCAGAAAGAGGCCACCTCTCAGGACCAACTCA
AGCAGCTGTTCGACAGGTCAAAGAGCCACTGA

>SmaBHMT4 (turbot; Accession numbers EY454591, EY456005, EY456033, EY456119, EY456251, EY456468 and EY456807)
ATGGCACCTGTCAAGAAGGGTATCATTGAACGTCTGAGCGCTGGGGAGGTGGTGATCGGGGATGGAGGCTTTGTGTTTGCTCT
GGAGAAGAGGGGATACGTAAAGGCTGGGCCCTGGACTCCTGAGGCTGCTGTCACACATCCTGAGGCTGTGCGACAGCTGCACA
GGGAGTTCCTGAGGGCAGGGTCTAACGTCATGCAGGCATTTACATTCTACGCCAGTGATGACAAACTGGAGAACAGGGGTCAG
ATGCTGAAAACCACCGGGGTACAAGTCAACGAGGCAGCGTGCGACCTGGCGAGGGAAGTTGCCAACGAGGGCGATGCCCTGGT
GGCCGGCGGAGTGTCTCAGACTCCATCCTACCTGAGCTGCAAGAGCGAGACAGAAGTGAAGGCCATCTTCAAGAAGCAGCTGG
AAGTGTTCACGAAGAAGAACGTGGACTTCCTAATTGCTGAGTACTTTGAGCATGTCGAGGAGGCGGAGTGGGCCGTGCAAGTG
CTGAAGACGAGCGGGAAGCCTGTCGCCGCCTCTCTGTGCATCGGACCAGAGGGAGACATGCACGGCGTCTCACCTGGAGAGTG
TGCCGTCAGGCTGGTGAAGGCCGGGGCCCAGATCGTAGGAATCAACTGCCACTTTGACCCCATGACCTGTGTGAAGACTGTTA
AGATGATGAAGGAGGGAGTAGAGAAGGCCAAGCTGAAGGCTCACTACATGGTGCAGCCACTTGCCTTCCACACTCCCGACTGC
AACTGTCAGGGATTCATCGACCTGCCAGAGTTCCCCTTTGCCCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACACGTA
CGCCAGAGAGGCCTACAAGGCGGGCATCAGATTCATTGGTGGCTGCTGCGGGTTCGAGCCTTACCACATTAGGGCCGTGGCTG
AGGAGCTGGCCACTGAGAGAGGGGTCATGCCCCCTGGCTCAGACAAACATGGAATGTGGGGCTCTGCTCTGGAGATGCACACC
AAACCCTGGGTTAGAGCCAGGGCCCGTCGTGACTACTGGGAGCATATTTTGCCCGCATCCGGCCGACCCAATTGCCCGTCCGT
TTCCACACCAGAGTCTTGGGGTGTGACCAAGGGCCACACCGACCTGCTGCAGCACAAAGAGGCCACCAGCACCCAGGAAATGA
AGCGCGTGCTGGATATGCAGAAGAAGGCCAAGTCCCCGGCGTGA

>SpaBHMT6/7 (bicolor damselfish; Accession number XM_008281472)
ATGGAGAACAAGAAGAGAAGGGGTATCTTGGAGCGCCTGGATGCTGGGGAGGTGGTTGTAGGAGATGGAGGTTACGTGATGCA
ACTGGAGCGACGTGGCTACGTGACAGCGGGACCCTGGACACCAGAAGCTGCGGTTGAACATCCTGAAGCAGTGCGACAGCTGC
ACAGGGAGTATCTGAGAGCAGGAGCCACCGTGATTCAGACTTTCACCTTCTACTGCAGTGAGGATAAACTGGAGCTCAATGGC
AATGTCACCAACATCACAGGAGCCCAGATCAATGATGCAGCCTGTAAAATGGCCAGAGAGGTCGCTGATGAGGGGGGTGCACT
GGTCGCTGGTGGTGTTTCTCAGACCCCCTGTTATGTGCAGAGTCACAGTGAGACTGAGGTCAAAGCCATCTTTAAGAAGCAAA
TGGATGACTTCCTCAAACAGGACATTGATTTCTTGATTGTGGAGTATTGCGAGCATGTTGAGGAGGCAGTGTGGGCAGTGGAA
GTGCTGAAAACCAGCGGTAAAACAGTGGCCGCAACACTCTGCATCGCCCCGCAAGGAGACATGCATGGAGTCCCACCTGGAGA
GTGTGCTGTCAGGCTGGTCAAAGCTGGGGCGGACATTGTCGGAATAAACTGCCACCTGGACCCACTGACTTGTGTTGAGACGG
TGAAGCTGATGAAAGAGGGATTAGAGAAAGCAGGTCTCAAAGCCCATCTCATGATCCAGCCGCTGGGCTTTCACACACCAGAG
CGCAACATTGATGGCTACCTCAGCCTGCCTGAGTTCCCCTTCGCACTGGAGACTCGAGCAATGACTCGCTGGGACATCCATAA
ATATGCCAGAGAAGCTTACAATGCAGGAATTCGCTACATCGGTGGCTGCTGTGGGTTTGAGCCGTACCATATCAGAGCTATAG
CAGAGGAGCTTTCTGCAGAGAGAGGATTCCTCCCACCAGCTTCAGAGAAACACGGACTTTGGGGCGCTGCCCTGGAGATGCAC
ACTAAACCCTGGGTCAGAGCCAGGGCTCGTCGTGAATATTGGGAAAAGCTTTTGCCTGCTTCTGGACGTCCCAAATGCCCTTC
AATGGCCACTCCAGCTACTGAATAA

>SpaBHMT7/6 (bicolor damselfish; Accession number XM_008281473)
ATGGAGAGCAAGAAGAGAAGGGGTATCTTGGAGCGCCTGGATGCTGGGGAGGTGGTTGTAGGAGATGGAGGTTACGTGTTGCA
GCTCGAGCGACGTGGCTACGTGAAGGCGGGACACTGGACACCAGAAGCTGCGGTTGAACATCCTGAAGCATTGCGACAGCTGC
ACAGGGAGTTTCTGAGAGCAGGAGCCGACATAATGCAGACTTTCACCTTCTACTGCAGTGAGGATAAACTGGATCTCAGTGGC
AATGTCACCAACATCACAGGAGCCCAGATCAATGATGCAGCCTGTAACCTGGCCAGAGAGGTCGCCGATGAGGGGGGTGCACT
GGTCGCTGGTGGTGTTTCTCAGACCCCCTGTTATGTAAAGAATCGCAATGAGACTGAGGTCAAAGCCATCTTTAAGAAGCAAA
TGGATAACTTTCTCAAACAGGACATTGATTTCTTGATCGTGGAGTATTTCGAGTATGTTGAGGAGGCAGTGTGGGCAGTGGAA
GTGCTGAAAACCAGCGGTAAAACAGTGGCTGCAACACTTTGCATCGCCCCACAAGGAGACAGGCATAAAGTCCCACCTGGAGA
GTGTGCTGTCAGGCTGGTCAAAGCTGGGGCGGACATTGTCGGAATAAACTGCCAACTCGACCCAACGATTTGTTTAAAGACGG
TGAAGCTGATGAAAGAGGGATTAGAGAAAGCAGGTCTCAAAGCCCATCTCATGATCCAGCCGCTGGGCTTTCACACACCAGAG
TGCAACATTGGTGGATACGTCAGCCTGCCGGAGTATCCCTTCGCACTGGAGACTCGAGCAATGACTCGCTGGGACATCCATAA
ATATGCCAGAGAGGCTTACAATGCAGGAATTCGCTACATCGGTGGCTGCTGTGGGTTTGAGCCGTACCATATCAGAGCTATAG
CAGAGGAGCTTTCTGCAGAGAGAGGATTCCTCCCACCAGCTTCAGAGAAACATGGACTTTGGGGCGCTGCCCTGGAGATGCAC
ACTAAACCCTGGGTCAGAGCCAGGGCTCGTCGTGAATATTGGGAAAACCTTTTGCCTGCTTCTGGACGTCCCAAATGCCCTTC
AATGGCCACTCCAGCAGCAGACGGACTGGGTGAAAACTGA

>SpaBHMT4 (bicolor damselfish; Accession number XM_008294071)
ATGGCACCGTCAAAGAAGGGAATCATGGAGCGTCTCAACGCTGGGGAGGTGGTGATTGGCGATGGAGGTTTCGTGTTTGCTCT
GGAGAAGAGAGGCTACGTGAAGGCCGGTCCCTGGACCCCTGAAGCTGCTGCAGAGTACCCCGAAGCAGTGCGACAACTGCACA
GGGAGTTCCTGAGGGCTGGATCCAACGTCATGCAGACCTTCACCTTCTACGCCAGTGATGACAAGCTGGAGAACAGAGGCAAC
AAGCTCACCTTCACCGGAGCTCAGATCAACGAGGCCGCCTGTGACCTGGCCCGTCAGGTCGCCAACGAGGGCGACGCTCTGGT
CGCCGGTGGAGTGTGTCAGACTCCCTCCTACCTGAGCTGCAAGAGCGAGACAGAAGTGAAGGCCATCTTCAAGAAACAGCTGG
ATGTGTTCGTGAAGAAGAACGTGGACTTCCTGATTGCTGAGTACTTTGAGCACGTTGAGGAGGCCGTGTGGGCTGTGGAGGTG
CTGAAGGCGACGGGGAAGCCTGTGGCGGCTTGTATGTGCATCGGACCACAGGGAGACATGCACGGCGTTTCCCCTGCAGAGTG
CGCTGTCAGGCTGGTCAAGGCCGGTGCCCAGATTGTGGGAATCAACTGCCACTTTGACCCCATGACCTGCGTGGAGGCCGTCA
GGATGATGAAGGAGGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTGCAGCCCCTGGCGTACCACACCCCCGACTGC
AACTGTCAGGGATTCATCGATCTGCCAGAATTCCCCTTCGGCCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACAAGTA
CGCCAGAGAGGCCTACAACGTTGGCATTCGCTTCATCGGTGGCTGCTGTGGCTTTGAGCCTTATCACATCAGGGCTGTGGCAG
AGGAGCTGGCTGCTGAGAGGGGCATCAAGCCCGCTGCGTCAGAGAAACACGGCATGTGGGGCGCTGGTCTGGAGATGCACACC

AAGCCCTGGGTCAGAGCCAGAGCCCGCCGTGACTACTGGGAGAACCTGAAGCCAGCCTCCGGTCGTCCTCAGTGCCCCTCCCT
GGCCACCCCCGACGGCTGGGGTGTCACCAAGGGCCACGCTGACCTCATGCAGCAGAGAGAGGCCACTTCTAAGGACCAACTCA
AACAGCTGTTCGACAGGTCAAAGAGCCACTGA

>SpaBHMT5 (bicolor damselfish; Accession number XM_008294072)
ATGGCACCTGTTAAGAAGGGTATCCTTGAACGGCTCAATGCTGGCGAGGTGGTGATTGGCGATGGAGGCTTTGTGTTCGCTCT
GGAGAAGAGGGGATACGTGAAGGCTGGGCCTTGGACTCCTGAGGCTACTGTCACACACCCTGAGGCCGTGCGACAGCTGCACA
GGGAGTTCCTGAGAGCCGGAGCTAATGTCATGCAGACACTGACGTTCTATGCCAGTGATGACAAACTGGAGAACAGGGGTCAG
ACAGTGAAGTATACTGGAGTACAAATCAACGAGGCCGCCTGCGACCTGGCGAAGCAAGTAGCCAGCGAGGGCGACGCTCTGGT
GGCCGGTGGAGTGTGTCAGACTCCCTCCTACCTGAGCTGCAAGAGCGAGGCGGAAGTGAAGGCCATCTTCAAGAAACAGCTGG
ATGTGTTTGTGAAGAAGAACGTGGACTTCCTGATTGCTGAGTACTTTGAGCACGTCGAGGAGGCAGAATGGGCCGTGCAGGTG
CTGAAGGCCACCGGAAAACCCGTAGCTGCTTCCATGTGCATCGGACCAGACGGAGACTTGCACGGTGTTTCTGCTGGAGAGTG
CGCCGTCAGGCTGGTGAAGGCTGGTGCCCAGATTGTGGGAGTCAACTGCCACTTTGACCCCATGACCTGCGTGAAGACCGTGA
AGCTGATGAAGGAGGGAGTGGAGAAGGCCGGGCTGAAGGCTCACTACATGGTGCAGCCGCTGGCGTACCACACCCCCGAGTGC
ACCTGTCAGGGATTCATCGAGCTGCCAGAATTCCCCTTCGCCCTGGAGCCCAGGTTGCTGACCCGCTGGGACATGCACAAGTA
CGCCAGAGAGGCCTACAACGTTGGCATTCGCTTCATCGGTGGCTGCTGTGGCTTTGAGCCGTATCACATCAGGGCCTTGACAG
AGGAGCTATCTGCCGAGAGAGGCATCCTGCCCCCTGGAGCCGACTACCACGGAATGTGGGGTTCTGGTCTGGAGATGCACACC
AAGCCCTGGGTCAGAGCCAGGTCCCGCCGTGAGTACTGGGAGAACATCTTGCCATCTTCTGGTCGTCCCCTGTGCGCCTCCCT
GTCCAAGCCGGAGCACTGGGGCGTGACCAAGGGCCACGCCGACCTGCTGCAGCACAAAGAGGCCACCAGCAGCCAGGAAATGA
GGCACGTGCTGGAGATGCAGAAGAAGGCCAAGTCTTCTTCGTGA

>SraBHMT6/7 (grass rockfish; Accession numbers EW979074, EW986546, EW976038, EW979075, EW986547, EW976039, EW979368 and EW979367)
ATGGAGAGCAAGAAGAGAAGGGGTATCTTGGAGCGACTAAATGCTGGGGAAGTGGTTGTAGGAGATGGAGGTTATGTGATGCA
GCTAGAGCGGCGTGGCTATGTGAAGGCTGGACATTGGACACCTGAAGCTGCTGTTGAACATCCTGAAGCAGTGCGGCAGCTGC
ACAGGGAGTTTCTCAGAGCAGGAGCCAATGTGATTCAGACTTTTACCTTCTACTGCAGTGAGGATAAACTGGACATCAGTGGC
AATGTCACCAACATCACTGGGGCCCAGATCAATGAGGCAGCATGTGACCTGGCCAGAGAGGTAGCCAATGAGGGTGATGCGTT
GGTCGCTGGGTGTGTGTCTAAGACTCCCTGTTATGTGGAGACTCACAGTGAGACTGAGGTCAAGGCCATCTTTAAGAAACAGA
CGGATGACTTCCTCAAGAAGGACATTGATTTCTTTATAGTGGAGTTCGTTAATCACGTGGAAGAGGCAGTGTGGGCAGTGGAG
GTGCTGAAGACCAGTGGTAAACCAGTGGGTGCAACACTGTGCATCTCCCCTCACGGAGACATGAACGGAGTCCCACCTGGAGA
GTGTGCTGTCAGGCTGGTCAAAGCTGGAGCCGACATTGTTGGAATAAACTGCCACTTGGACCCTCTGACGTGTGTTCGTACGG
TGAAGTTGATGAAAGCGGGATTAGAGAAAGCTGGTCTCAAAGCCCATCTCATGATCCAGCCGCTGGGCTTTCACACACCTGAG
TGCAACTTTGGTGGATACACCAGTCTACCTGAGTACCCCTTTGCAATGGAGACCAGAGCAATAACCCGCTGGGACATCCATAA
ATACACCAGAGAGGCTTACAATGCTGGTATTCGCTACATTGGTGGCTGCTGTGGATTTGAGCCCTACCATATCAGAGCTATAG
CAGAAGAGATGGCTGAAGAGAGAGGATTCCTCCCACCAGCTTCAGAGAAGCACGGACTCTGGGGAGCTGCTCTGGAGATGCAC
ACTAAGCCCTGGGTCAGAGCCAGGGCTCGTCGAGAGTACTGGGAAAACCTTTTGCCTGCTTCTGGACGTCCCAAATGCCCTTC
CATGGCCACACCAGCTGCTGATTATGAAAAAACAGATGGAATCTAA

>SsaBHMT3 (Atlantic salmon; Accession number NM_001139685)
ATGGCACCAGCTGGAGCTAAGAGGAACATTCTGGAGCGCCTGGATGCAGGCGAGATTGTTATTGGGGATGGAGGTTTCGTCTT
CGCCCTGGAGAAAAGGGGCTATGTGAAGGCGGGTCCCTGGACTCCTGAGGCTGCCGCCGAGCACCCCGAGGCTGTGAGACAGC
TGCACCGTGAATTCCTGCGGGCGGGGTCAAATGTCATGCAGACTTTCACCTTCTACGCCAGCGATGATAAACTGGAGAACAGG
GGTAACGCTCAGCGCTTCACCGGCACACAAATCAACGAGGCCGCCTGTGACCTGGCCAGGGAGGTGGCCAATGAGGGCGACGC
GATGGTGGCGGGCGGGGTCTCCCAGACCCCCTCTTACTTGAGCTGCAAGTGCGAGGACGAGGTCAAGAGCATCTTCAAGAGGC
AGCTTGATGTCTTCGTCAAGAAGAACGTGGATTTCATGATTGCAGAGTACTTTGAGCACGTGGAGGAGGCAGAGTGGGCCGTG
CAGGTGCTGAAGACCTCGGGCAAGCCAGTGTGTGCCTCTCTGTGTATCGGCCCTGACGGAGACCTCAACGGAGTCAGCCCTGG
AGACTGTGCCGTCCGGCTGGTCAAAGCTGGAGCCAATATTGTGGGCATCAACTGTCACTTTGACCCCATGACCTGTGTGAAGA
CGGTGAAAATGATGAAGGAGGGTGTTGAGAGGGCGGGGCTGAAGGCCCACTACATGGTCCAGCCTCTGGCCTTCCACACCCCT
GACTGCAACTGCCAGGGATTCATCGACCTGCCAGAGTTCCCCTTCGGTCTGGAGCCCAGGATTCTGACCAGGTGGGACATGCA
CAAGTACGCCAGGGAGGCCTTCAACGTTGGCATCCGTTTCATTGGCGGCTGCTGCGGCTTCGAGCCCTACCACATCAGGGCTG
TAGCTGAGGAGCTGGCCACAGAGAGAGGCTACCTGCCCGCTGCCTCAGAGAAGCACGGTAACTGGGGTGCTGGCCTGGAGATG
CACACCAAGCCATGGGTCCGTGCCAGGGCTCGCCGTGACTACTGGGAGAAGTTGAAGCCAGCATCCGGCCGTCCCAAGTGTCC
CTCCATGTCCACCCCTGATAGCTGGGGTGTCACCAAGGGCCATGCCGACCTGATGCAGCACAAAGAGGCCACCAGCCAGGAAG
AGCTCAAGCCCCTGTTTGAGAAGGCCAAGGCCAGCCACTGA

>TniBHMT5 (green spotted pufferfish; Accession number CR636934)
ATGGCACCGGGAAAGAAGGGAATTTTGGAGCGTCTCAACGCAGGAGAGATCGTGATCGGCGATGGAGGGTTTGTGTTTGCGCT
GGAGAAGAGAGGCTACGTGAAGGCCGGTCCCTGGACCCCTGAAGCCACAACGGAACACCCCGAAGCAGTGCGCCAGCTGCACA
GGGAGTTTCTGAGGGCAGGGGCCAACGTCATGCAGACCTTCACCTTCTACGCCAGCGACGACAAGCTGGAGAACAGAGGCAAC
AAAGTCACCTACACTGGAGTTCAGATCAACGAGGCCGCCTGCGATCTGGCCCGGGAGGTGGCCAATGAAGGCGATGCCCTGGT
GGCCGGAGGAGTGTGTCAGACACCATCTTACCTGAGCTGTAAGAGCGAGAAGGAAGTGAAGGACATCTTCAAGAGACAACTGG
ACGTGTTTGTGAAGAAGAACGTGGACTTCCTGATTGCTGAGTACTTTGAGCATGTTGAAGAGGCCGTGTGGGCTGTGGAGGTG
CTGAAGACCACAGGAAAGCCTGTCGCTGCTTCTCTGTGCATCGGACCCAAGGGTGACATGCACGGCGTGTCCCCAGGAGAGTG
TGCTGTCAGGCTGGTCAAAGCTGGTGCCCAGATTGTCGGAGTCAACTGCCACTTTGACCCCATGACCTGCGTGGAAACTGTCA
AGCTGATGAAAGAGGGAGTGGAGAAGGCCGGGCTCAAGGCCCACTACATGGTGCAGCCCCTGGCGTTCCACACCCCCGACTGC
AACTGTCAGGGCTTCATCGACCTGCCAGAATTTCCTTTCGGTCTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACAAGTA
CGCCAGGGAGGCCTACAACGCAGGGATCCGATTCATCGGAGGCTGCTGTGGATTTGAGCCGTATCACATCAGAGCCTTGGCAG
AGGAGCTGGCCCCTGAGAGAGGCTTCCTCCCCGTCGGATCAGACAAACATGGCAACTGGGGTGCTGGTTTGGAGATGCACACC

APPENDIX

AAGCCCTGGGTCCGAGCCAGAGCCCGTCGTGACTACTGGGAGGCCCTGAAGCCAGCCTCCGGTCGGCCCCTGTGCCCCTCCCT
GTCCACTCCAGATTGCTGGGGTGTCACAAAGGGTGATACTGAGCTCATGCAGCAGAAAGAGGCCACCTCAAAGGAACAACTTA
AGCAGCTGTTCGACCGCTCAAAGAGCCACTAA

>TruBHMT3 (Japanese pufferfish; Accession number XM_003978117)
ATGGCACCCGCGGCGAACAAGAAGGGCATTTTGGAGCGTCTGGATGCAGGAGAGATCGTTATCGGTGACGGAGGCTTCGTGTT
TGCACTTGAGAAGAGAGGTTACGTAAAAGCAGGACCGTGGACACCTGAAGCCACCGTGGAGTACCCCGAAGCCGTGCGCCAGC
TGCACAGAGAGTTCCTGAGGGCGGGTTCTGATGTCATGCAAACGTTCACTTTTTACGCCAGCGATGACAAACTAGAGAACAGG
GGCCACGCTCAGCGCTTCACCGGGACACAAATAAATGAAGCAGCCTGTGACCTGGCCAGGGAGGTGGCCAATGAAGGCAATGC
CCTGGTGGCAGGAGGAGTCTCTCAGACCCCAGCTTACCTTAGCTGCAAGAGTGAAGATGAGGTCAAGGCTATCTTCAAGAAGC
AGCTTGACGTCTTTGTGAAGAAAGACGTGGACTTCTTGATTGCAGAGTACTTTGAGCATGTGGAAGAGGCTGAATGGGCCGTA
CAGGTTCTGAAGGCTACTGGGAAGCCTGTGGCCGCCACCCTGTGTATCGGACCAGAAGGTGATCTGAATGGAATCAGCCCCGG
GGAATGTGGTGTCAGACTTGTCAAAGCTGGAGCTCAGATTGTGGGCATCAACTGCCATTTTGACCCAGAGACCTGTGTGAAGA
CTGTGAAGATGATGAAGGAGGGCGTGGAGAAGGCTGGACTAAAGGCTCATTACATGTCCCAGCCGCTGGCTTATCACACTCCT
GATTGCAACTGCCAGGGCTTCATTGATCTCCCCGAATTCCCCTTTAGTCTGGAGCCCAGGATACTGACCAGATGGGACATGCA
AAAATATGCCCGGGAGGCCTATAACGCTGGTATCCGTTACATCGGTGGCTGCTGCGGATTTGAAGCCTATCATATCCGTGCCC
TGTCTGAGGAACTGCAGGCTGAGAGGGGTTTTTATCCTGCTGGCTCAGAGAAACATGGCAACTGGGGCAGCGGTTTGGAGATT
CACACAAAGCCTTGGGTCAGAGCCAGGGCTCGTCGTGACTACTGGGAGAAGCTGAAACCTGCTTCTGGTCGTCCCTTCTGCCC
CTCTTTGTCAGAACCCGATAGCTGGGGTATCACCAAAGGCCATGCTGACCTGATGCAGAAGAAGGAGGCCACCTCACAGGAGG
AAATGAAAGCTCTTTTTCAAAAGGCTGACAAGTGTTAG

>TruBHMT5 (Japanese pufferfish; Accession number XM_003974601)
ATGGCACCAGGAAAGAAGGGAATTTTGGAGCGTCTCAACGCAGGCCAGGTTGTGATCGGCGATGGAGGGTTCGTGTTTGCTCT
GGAGAAGAGAGGCTACGTGAAGGCCGGTCCCTGGACCCCAGAAGCCACAAAGGAACATCCTGAAGCAGTGCGGCAGCTGCACA
GGGAGTTCCTGAGGGCGGGGGCCAACGTCATGCAGACCTTCACCTTCTACGCCAGCGATGACAAGCTGGAGAACAGGGGCAAC
AAACTCACCTACACTGGAGCTCAGATCAACGAGGCCGCCTGCGATCTGGCCCGGGAGGTTGCCAATGAGGGCGATGCCCTGGT
TGCTGGGGGAGTGTCTCAGACCCCATCTTACCTGAGCTGTAAGAGTGAGAAAGAAGTGAAGGACATCTTCAAGAGACAACTGG
ATGTGTTTGTGAAGAAGAACGTGGACTTCCTGATCGCTGAGTACTTTGAGCATGTTGAAGAGGCCGTGTGGGCTGTGGAGGTG
TTGAAGGCGACAGGAAAGCCTGTTGCCGCTTCTCTGTGCATTGGACCCAAGGGAGACATGCACGGCGTGTCCCCAGCAGAGTG
CGCCGTCAGGCTGGTCAAAGCTGGTGCCCAGATCGTTGGAGTCAACTGCCACTTTGACCCCATGACCTGCGTGGAAGCTGTCA
AGATGATGAAAGAGGGAGTGGAGAAGGCCGGGCTCAAGGCTCACTACATGGTGCAGCCCCTGGCCTTCCACACCCCCGACTGC
AACTGTCAGGGATTCATCGATCTGCCAGAATTTCCTTTTGGTTTGGAGCCCAGGATCCTGACCCGCTGGGACATGCACAAGTA
TGCCAGAGAGGCCTACAACGCTGGGATCAGATTCATTGGAGGCTGCTGTGGATTTGAGCCTTATCACATCAGAGCTTTGGCAG
AGGAGCTGAGCCCTGAGAGAGGCTGCCTCCCAGTCGGATCAGAGAAACACGGCAGCTGGGGTGCTGGTCTGGAAATGCACACC
AAGCCCTGGGTCAGAGCCAGAGCCCGTCGTGACTACTGGGAGAACCTGAAGCCAGCCTCTGGCCGGCCCCTGTGCGCCTCCCT
GTCTGCTCCCGATTGCTGGGGTGTCACAAAAGGTGACACTGCCCTCATGCAGCAGAAAGAAGCAACCACAAAGGACCAACTGA
AGCAGCTGTTCGACCGCAAAAGCAACTAA

>LchBHMT1 (West Indian Ocean coelacanth; Accession number XM_005990710)
ATGTCACCGCCTGGAAAAGCCAAGAGAGGTATTCTGGAGCGCTTGGATGCCGGAGAGGTGGTGGTAGGAGATGGTGGGTTTGT
GTTTGCTCTGGAGAAGAGGGGTTGGGTCCGAGCAGGACCATGGACTCCGGAGGCCACTGTGGAGCATCCAGAGGCAGTCCGGC
AGCTTCATCGAGAGTTCCTTCGAGCCGGAGCAAATGTCATGCAGACTTTCACATTTTATGCTAGTGAGGACAAGCTGGAGAAT
AGAGGCAACTATGTCTCTCAGAAAATATCTAGTCAAAAGGTGAACGAAGCTGCCTGTGATCTTGCCCGAGAAATAGCCCAGGA
AGGCGATGCACTGGTAGCCGGCGGCGTGTGCCAATCTCCCTCCTACCTCAGCTGCAAGAGTGAAGTGGAGGTGAAAAACATTT
TCCGGAAACAGCTGGACGTTTTTGTCAAGAAAAACGTGGACTTCCTGATTGCTGAGTACTTTGAACACGTGGAGGAAGCTGAA
TGGGCAGTTCATGTTCTCAAGGAATCTGGAAAACCAATCGCTTCCTCTCTCTGCATTGGTCCACAAGGAGACCTGAATGGAGT
TCCACCTGGCGAGTGTGCAGTCAAACTAGTAAAAGCAGGTGCATCCATTGTGGGAGTTAATTGCCATTTTGACCCAGCAACTT
GCTTGAAGACTGTGAAGCTCATGAAAGAGGGTTTGCAGGCTGCTGGTGTTAAGGCCCATCTAATGATTCAGCCCCTTGCTTTT
CACACCCCTGATTGTAGCCAGCAAGGCTTCATTGATCTACCTGAATTTCCTTTTGCTTTGGAACCTAGGATCCTTACTAGGTG
GGATGTACAGAAGTATGCAAGAGAAGCCTACAACTTGGGAATCCGCTATATTGGAGGCTGCTGTGGATTTGAACCATATCACA
CTCGAGCCATATCTGAAGAACTGGCAACAGAAAGGGGCTTTTTACCACCAGCTTCTGAGAAACATGGCTTGTGGGGCATTGCT
CTAGCAATGCACACCAAGCCCTGGGTTAGAGCAAGGTCCAGAAGGGAACACTGGGAGAAAATGTGCCCTGCTTCCGGCAGGCC
ATATTGTGCTTCAATGTCCAAACCAGATACCTGGGGTGTGACCAAAGGAGATGCTGCGCTGATGCAGCAGAAAGAGGCAACAA
GTGAACAGCAGCTGCAGGAACTCTTTGCCAGGCAAATGCAACAGGCCGGCAACGTGGTGTAA

>CpyBHMT1 (Japanese firebelly newt; Accession numbers FS312172, FS293370, FS307658,
FS306025, FS305377 and FS300617)
ATGGCAGCAAAGAAAGGTATTCTGGCCCGTCTGGATGCTGGAGAAATAGTAATTGGAGATGGCGGTTTTGTGTTTGCATTGGA
GAAGAGGGGTTACGTGAAGGCTGGACCATGGACCCCAGAAGCTGCAGTGGAGCACCCAGAAGCAGTACGACAGCTTCATCGAG
AGTTTCTTCGTGCTGGATCAAATGTCATGCAGACATTCACTTTCTATGCGAGCGAGGACAAGCTGGAAAACAGAGGCAACTAC
GTAGCATCTAAAATTTCTGGGCGGAAGGTGAACGAAGCGGCCTGTGACATTGCGAGAGAGGTGGCCGCGGAGGGCGACGCTTT
AGTGGCCGGAGGGGTCAGCCAGACCCCATCTTACCTCAGTAACAAGAGCGAGACCGAGGTGAAGGGCATCTTCCGGCAGCAGC
TTGGAGTCTTCACTTGTAAAAAAGTGGATTTCTTAATTGCAGAGTACTTTGAACACGTGGAAGAAGCCGCCTGGGCTGTTGAG
GTCCTGAAGGAGTCTGGGCTGCCAGTGGCTGCTACCATGTGCATCGGTCCGGAGGGGGACATGAATGGGGTGTCGCCTGGCGA
ATGTGCAGTAAGGCTGGTCAAGGCTGGTGCAGCAATTGTCGGCGTGAACTGCCACTTTGACCCAACGTTTTCTCTGAAAACCA
TGAAACTCATGAAACAGGGCCTGGAAGCAGAAAACATCAAAGCACATCTGATGGTCCAACCACTCGCTTACCACACCCCCGAC
TGTGGCAAACAGGGATTTATTGATCTTCCAGAGTTTCCTTTTGCTTTGGAGCCAAGGATAGTCACCAGGTGGGATATTCAGAA
ATATGCGAGGGATGCCTACAATCTTGGCATCCGGTACATCGGCGGTTGCTGTGGATTTGAACCTTACCACATCCGAGCAATAG

CAGAGGAGCTTGCTCCAGAAAGAGGATTTTTGCCAAAAGGATCTGAAAAACATGGCAGCTGGGGCAGTGGTCTGGACATGCAC
ACGAAACCCTGGGTTAGAGCAAGAGCACGAAGGGACTACTGGGAGAAAATTGCCCCTGCATCTGGCAGACCATATTGTCCTTC
AATGTCCAGGCCTGATGCCTGGGGAGTGACCAAAGGAGCAGCTGAGCTGATGCAGCAGAAGGAGGCCACCACTGAGGACCAAC
TCAAGGAGCTGTTCGCAAAACAGAAACTGTGA

>XlaBHMT1 (African clawed frog; Accession number NM_001094947)
ATGGCACCGGTTGGAGCAAAGAAAGGTCTTCTGGAGCGCCTAGATGCTGGAGAAGTTGTCATTGGAGATGGTGGCTTTGTGTT
TGCTTTAGAGAAACGAGGGTATGTCAAGGCTGGACCTTGGACGCCTGAAGCTGCAGTGGAACATCCAGAAGCAGTACGGCAGC
TTCATCGAGAATTTCTCCGAGCTGGAGCCAACGTCATGCAGACCTTTACCTTCTATGCTAGTGATGATAAGCTGGAGAACAGA
GGCAACTATGTGGCTGAGAAAATCTCTGGCCAGAAAGTAAATGAGGTGGCCTGTGATATAGCACGAGAAGTGGCCAATGAAGG
CGATGCATTAGTGGCAGGAGGGGTGAGCCAGACTCCATCGTACCTTAGCTGCAAAAGTGAAGTCGAGGTGAAAGGCATCTTCA
GGAAGCAACTGGATGTCTTCATTAAGAAGAACGTGGATTTTCTCATTGCCGAGTATTTTGAACATGTTGAAGAAGCTGTTTGG
GCTGTTGAGGTACTGAAAGAATCCGGGAAGCCAGTTGCTGCTACGCTCTGCATTGGTCCTGAGGGGGATTTGAATGGAGTTTC
ACCTGGAGAGTGTGCTGTTCGCTTGGCCAAAGCTGGTGCCTCGGTGGTGGGAGTCAACTGCCATTTTGACCCAATGACCTGTG
TTGCAACAGTGAAGCTTATGAAAGAAGGCCTGGTAGCTGCTAAGGTGAAGGCCCATCTAATGACCCAGCCACTTGCTTATCAT
ACGCCAGACTGTGGAAAGCAAGGTTTTATTGACCTGCCTGAGTTTCCATTTGCTTTGGAACCAAGAATTGTCACTAGGTGGGA
CATACACAAGTACGCTCGGGCGGCGTATGACCTTGGAGTGCGCTATATTGGAGGCTGCTGTGGCTTTGAGCCCTATCACACTA
GAGCAATAGCAGAAGAACTGGCTCCAGAAAGAGGTTTCCTACCGAAAGGTTCTGAGAAGCACGGAAGCTGGGGAAGTGGCTTA
GAAATGCATACAAAACCCTGGGTCAGAGCAAGGGCTAGGCGAGACTACTGGGAGAAGCTACCTCCAGCCTCAGGTAGACCTTA
CTGTCCTTCTATGTCCAAGCCAGATGAATGGGGTGTGACCAAAGGAGATGCTGATCTAATGCAGCAAAAAGAGGCAACTACAG
AACAGCAACTGAAAGATCTAATTGCAAAGCAAGGCATCAAATCCAACTAG

>XtrBHMT1 (Western clawed frog; Accession number NM_001011217)
ATGGCACCGACTGGAGCAAAGAAAGGCCTTCTAGAGCGCCTAGATGCCGGTGAGGTTGTCATCGGAGACGGCGGATTTGTGTT
TGCTTTAGAAAAACGAGGCTATGTCAAGGCCGGACCTTGGACGCCAGAAGCTGCAGTGGAACACCCAGAAGCAGTACGGCAGC
TTCACCGAGAATTCCTGCGAGCTGGAGCCAATGTCATGCAGACATTTACCTTCTATGCTAGTGATGACAAGCTGGAGAACAGA
GGCAACTATGTGGCTAAGAAAATCTCTGGCCAGAAAGTAAATGAAGCTGCCTGTGACATAGCCCGAGAAGTGGCCAACGAGGG
TGATGCATTAGTAGCAGGAGGGGTGAGCCAGACACCATCGTACCTCAGCTGCAAAAGTGAAGTCGAGGTGAAAGGCATCTTCA
GGAAGCAACTGGATGTCTTCATTAAGAAAAACGTGGATTTCCTCATAGCTGAGTATTTTGAACATGTTGAAGAAGCTGTTTGG
GCCGTTGAGGTACTGAAAGAGTCTGGAAAACCAGTTGCTGCTACGCTCTGCATTGGTCCTCAGGGAGATTTGAATGGAGTTAC
ACCTGGAGAGTGTGCTGTGAGGTTGGCCAAAGCTGGTGCCTCTGTGGTGGGAGTCAACTGCCATTTTGACCCAATGACCTGTA
TTGCCACAGTGAAGCTTATGAAGGAAGGCCTGGTAGCTGCTAAGGTGAAGGCCCATTTAATGACCCAGCCATTAGCTTATCAC
ACACCAGACTGCGGAAAGCAAGGTTTTATTGACCTGCCTGAGTTTCCATTTGCTTTGGAACCAAGAATTGTCACCAGGTGGGA
CATACACAAGTACGCTCGGGAGGCTTATAACCTCGGAGTGCGCTACATTGGAGGCTGCTGTGGCTTTGAGCCCTACCACACCA
GAGCAATAGCCGAAGAGCTGGCCCCAGAGAGAGGTTTCCTACCGCCAGGTTCTGAGAAGCATGGAAGCTGGGGAAGCGGCTTA
GAAATGCATACAAAACCCTGGGTTAGAGCAAGGGCTAGGCGGGACTACTGGGAGAAGCTACCTCCAGCCTCGGGTAGGCCTTG
CTGTCCTTCTATGTCCAAGCCAGATGCATGGGGCGTGACCAAAGGAGATGCTGATCTGATGCAGCAGAAAGAGGCAACAACAG
AACAACAACTGATCGATCTATTTGCAAAGCAATGCATCAAATCCAATTAG

>AcaBHMT1 (green anole; Accession number XM_003216351)
ATGGCTCCTGTTGGAGGGACAGCAAAAGCCAAGAAGGGCATTTTAGAACGTTTAAATGCAGGAGAAGTTGTGATTGGTGACGG
AGGATTTGTCATTGCTCTTGAAAAGAGAGGTTATGTAAAAGCTGGACCATGGACTCCAGAAGCTACAGTAGAACACCCAGAAG
CAGTTCGTCAGCTGCACCGAGAATTCCTAAGGGCTGGAGCAAATGTTTTACAAACATTCACCTTTTATGCCAGTGATGATAAA
TTAGAGAACAGGGGAAACTATGTAGCTGATAAAATTTCTGGCCAGAAGGTGAATGAAGCTGCTTGCAATATTGCAAAAGAAGT
AGCTGAAGAAGGAGATGCTTTGGTGGCTGGAGGTGTCAGTCAAACACCTTCATATCTAAGCAGCAAGAGTGAAGCTGAAATTA
AATCAATATTTCGAAAGCAGCTAAACATTTTTATGAAACAAAAGGTGGATTTCTTAATTGCAGAGTATTTTGAACATGTTGAA
GAAGCTGTGTGGGCAGTTGAGACATTAAAGGAATCTGGGCTGCCCATTGCCGCTAGTCTGTGCATTGGTCCAGAAGGAGACAT
GCATGGTATATCACCTGGAGAATGTGCTGTTCGGCTTGTAAAGGCTGGTGCTTCTATTGTTGGAGTAAACTGCCATTTTGATC
CAACCACTTGCCTCAAAACTGTGAAACTCATGAAAGAAGGTTTGGCAGCTGCTAAACTAAAGGCACATTTGATGTCACAACCA
CTGGCTTTTCATACTCCCGATTGTGGAAAGCAGGGTTTCATTGATCTGCCAGAATTTCCTTTTGCTTTGGAGCCAAGAATCCT
GACCAGATGGGATATCCATAAATATGCAAGGGAAGCTTATAACTTGGGAATCAGATACATTGGAGGATGCTGTGGATTTGAAC
CATATCACATCAGAGCCATTGCTGAGGAGCTGGCCCCTGAAAGAGGCTTCTTACCACATGGTTCTGAAAAACATGGCAGCTGG
GGAAGTGATCTGAGCATGCATACCAAACCATGGGTCAGAGCGAGGGCCAGGAAAGAGTACTGGGAGAATCTGCTGCCTGCTTC
AGGCCGCCCATATTGCCCTTCACTGTCAAAACCAGATGACTGGGGAGTGACCAAAGGAGATGCGGAGCTGATACAGCAGAAAG
AAGCAACATCTGAGCAACAGCTAAAAGAGCTGTTTAACAAACAGAAATTAAAATCCAAAGTTGTGGCTTAA

>AsiBHMT1(Chinese alligator; Accession number XM_006021415)
ATGGCACCGGTCGGAGGGACCCCCAAGCCGCCCAAGAAGGGTATCCTAGAGCGTCTAGATTCTGGAGAAATTGTGATTGGAGA
TGGAGGATTTGTCATTGCCCTTGAAAAGAGGGGTTATGTAAAAGCTGGGCCGTGGACTCCAGAAGCTGCAGTAGAACACCCAG
AAGCAGTCCGCCAGCTTCACCGGGAGTTCCTCAGAGCTGGATCAAATGTTTTGCAGACATTCACCTTTTATGCCAGTGAAGAC
AAACTAGAGAACCGGGGAAACTATGTATCTGACAAAATATCTGGCCAGAAAGTAAATGAAGCTGCTTGTGACATTGCACAGGA
AGTATCCCAAGAAGGAGATGCTTTGGTGGCAGGAGGAGTCAGTCAGACACCATCTTACCTCAGCTGCAAGGATAAAGCAGAGG
TTAAAGCAATTTTCCAAAAGCAGCTAGAGGTCTTTGTTAAAAAGAATGTGGACTTCTTAATTGCAGAGTATTTTGAACACGTT
GAAGAAGCTTCCTGGGCAGTTGAAGTCTTAAAAGAATCTGGGAAGCCAGTTGCAGTTTCTTTGTGCATTGGTCCAGAGGGAGA
TATGCATGGTATACCTCCTGGAGAATGTGCTGTCCAGCTGGTAAAGGCAGGTGCTTCTATTGTTGGGGTGAACTGCCATTTTG
ACCCAGAGACTTGCCTAAAAACTGTGAAGCTCATGAAAGAGGGCTTACAGGCTGCTAAACTGAAAGCACACCTGATGTCTCAA
CCACTTGCTTTCCATACCCCTGATTGTGGGAAGCAGGGTTTTATTGATCTTCCAGAATTTCCTTTCGGTCTGGAACCAAGAAT
TCTTACCAGATGGGATGTTCAAAAATATGCAAGAGAAGCCTATAACTTGGGAATCCGGTACATTGGAGGCTGCTGTGGATTTG

APPENDIX

AGCCCTATCACATCAGAGCAATAGCTGAGGAGCTGGCCCCTGAAAGAGGATTTTTGCCAGAGGCTTCTGAGAAACATGGTAGC
TGGGGAAGTGGTTTGAGCATGCATACCAAACCTTGGGTCAGAGCAAGGGCCAGAAAAGAGTACTGGGAGAATCTGCTGCCTGC
TTCAGGCAGGCCATACTGTCCTTCAATGTCAAAACCAGATGCCTGGGGAGTGACAAAAGGAGCTGCAGAGCTAATGCAACAGA
AAGAGGCAACAACTGAACAACAGCTGAAAGAGCTCTTTGAGAAACAGAAGTTCAAATCCAGTATGGTTGCCTAA

>CbeBHMT1 (painted turtle; Accession number XM_005286438)
ATGCAGTCGCCGCCTGTCGGAGCGAGCCCCCAGCACGCCAAGAAGAAGGGTATTCTAGAACGTCTAGATGCTGGAGAAATTGT
GATTGGAGATGGAGGATTTGTGATAGCTCTTGAAAAGAGAGGTTATGTAAAAGCTGGACCCTGGACTCCAGAAGCTGCAGTAG
AACACCCAGAAGCAGTTCGTCAACTTCATCGGGAATTCCTCAGAGCTGGATCAAATGTTTTACAGACATTCACCTTTTATGCC
AGTGAAGATAAATTAGAGAATAGGGGCAACTACGTAGCTGAGAAAATATCTAGCCAGGCAGTGAATGAAGCTGCTTGTGACAT
TGCAAAGGAAGTGGCTAAAGGAGGAGATGCTTTGGTAGCTGGAGGAGTCAGTCAGACACCATCATACCTGAGTGGCAGGGATG
AAGCTGAGGTTAAAGCAATTTTTCGAAAGCAGCTGGATGTCTTTGCCAAAAAGAATGTGGACTTCTTAATTGCGGAGTATTTT
GAACATGTTGAAGAAGCTGCATGGGCTGTTGAAACCTTAAAAGAATCTGGGAAGCCAATCGCAGCTACCATGTGCATTGGTCC
AGAGGGAGACTTGCATGGTGTACCACCTGGAGAATGTGCTGTCCAACTGGTAAAGGCTGGTGCTTCTATTGTTGGAGTGAACT
GCCATTTTGATCCAGAGACTTGCCTGAGAACTGTGAAGCTCATGAAAGAGGGCTTGGAGGCTGCTAAACTGAAAGCACACCTG
ATGTCTCAGCCACTTGCTTTCCATACCCCTGACTGTGGGAAGCAGGGCTTTATTGATCTTCCAGAATTTCCTTTTGCTCTGGA
GCCAAGAATTGTCACCAGATGGGACGTTCAAAAATATGCGAGAGAGGCCTATAACCTAGGGATTCGATACATTGGAGGCTGCT
GTGGATTTGAGCCTTATCACATCAGAGCAATAGCTGAGGAGCTGGCCCCTGAAAGAGGATTTTTGCCACTGGCTTCTGAGAAA
CATGGTAGCTGGGGTAGTGATTTGAGCATGCATACCAAACCTTGGGTCAGAGCAAGGGCAAGAAAAGAGTACTGGGAGAATCT
GTTCCCTGCTTCAGGCAGGCCATACTGTCCTTCAATGTCAAAACCAGATAACTGGGGAGTGACTAAAGGAGCTGCAGAGCTGA
TGCAGCAGAAAGAAGCAACAACTGAACAACAGCTGAAGGAGCTCTTTAAGAAACAGAAGTTCAAATCCAATATAGTGGCTTAA

>CmyBHMT1 (green sea turtle; Accession number XM_007066694)
ATGGCTCCCGGAAGCAGGAGCAGCCAGAGGGTTCCGCGTGCTGCCCTCACCGCAAGCAGCAGCTCTGCAGCTCCCATTGGCCG
GGAACCATGGCCAGTGGTACCTGCAGGGGTGGTGCCTGTGGATAGGGCAGAGCAGCATGGCCCTGCCTCCAGAGTGGGAAGAG
GCCACTGTTTCCAGGAGCTGTTTCAGGGTATTCTAGAACGTCTAGATGCTGGAGAAATTGTGATTGGAGATGGAGGATTTGTG
ATAGCTCTTGAAAAGAGAGGTTATGTAAAAGCTGGACCTTGGACCCCAGAAGCTGCAGTAGAACACCCAGAAGCAGTTCGTCA
ACTTCATCGGGAATTCCTCAGAGCTGGATCAGATGTTTTACAGACATTCACCTTTTATGCCAGCAAAGATAAATTAGAGAATA
GGGGCAACTACGTAGCTGAGAAAATATCTGGCCAGACAGTGAATGAAGCTGCTTGTGACATTGCAAAGGAAGTGGCTAAAGAA
GGAGATGCTTTGGTAGGTGGAGGAGTCAGTCAGACACCATCATACCTGAGTGGCAAAGATGAAGCTGAAGTTAAAGCAATTTT
TCGACAGCAGCTGGATGTCTTTGCCAAAAAAAAATGTGGACTTCTTAATTGCAGAGTATTTTGAACATGTTGAAGAAGCTGCAT
GGGCTGTTGAAACCTTAAAAGAATCTGGGAAGCCAGTCGCAGCTACCTTGTGCATTGGCCCAGAGGGAGACTTGCATGGTGTA
CCACCTGGAGAATGTGCTGTCCAACTCGTAAAGGCTGGTGCTTCTATTATTGGAGTGAACTGCCATTTTGATCCAGAGACTTG
CCTGAGAACTGTGAAGCTCATGAAAGAGGGCTTGGAGGCTGCTAAACTGAAAGCACATCTGATGTCTCAGCCTCTTGCTTTCC
ATACCCCTGACTGTGGGAAGCAGGGCTTTATTGATCTTCCAGAATTTCCTTTTGCTCTGGAGCCAAGAATTCTCACGAGATGG
GATGTGCAAAAATATGCAAGAGAGGCCTATAACTTAGGGATTCGATACATTGGAGGCTGCTGTGGATTTGAGCCTTATCACAT
CAGAGCAATAGCTGAGGAGCTGGCCACTGAAAGAGGATTTTTGCCACCGGCTTCTGAGAAACATGGTAGCTGGGGTAATAACT
TGAGCATGCATACCAAACCCTGGGTCAGAGCAAGGGCAAGAAAAGAGTACTGGGAGAATCTGTTGCCTGCTTCAGGCAGGCCA
TACTGTCCTTCAATGTCAAAACCAGATAACTGGGGAGTGACCAAAGGCGCTGCAGAGCTGATGCAGCAGAAAGAAGCAACAAC
TGAACAACAGCTGAAGGAGCTCTTTAAGAAACAGATGTTCAAATCCAATATAGTAGCTTAA

>GgaBHMT1 (chicken; Accession number XM_414685)
ATGCTGCCGATCGCGAAGAGCCTGAAACAGGGCAAAAGGGGCATTCTAGAGCGCTTGGATGCTGGAGAGATTGTGATTGGAGA
TGGGGGATTTGTCTTTGCCCTAGAAAAGAGGGGATATGTAAAAGCTGGGCCTTGGACTCCCGAAGCAACAGTAGAACACCCAG
AAGCAGTTCGTCAGCTTCACAGGGAATTCCTCAGAGCTGGATCCAATGTTCTGCAAACATTCACCTTTTATGCCAGTGAGGAC
AAACTAGAGAACAGAGGCAACTACGTAGCTGAGAAAATAACTTGCCAGAAAGTGAATGAAGCTGCTTGTGACATTGCCAGAGA
AGTGGCCAATGAAGGTGATGCTTTAGTAGCTGGAGGAGTTAGTCAGACACCATCATACTTAAGCTGCAAGGACAAAGCAGAGG
TTAAAGCAGTCTTTCGAAAGCAGCTAGATATCTTCATGAAGAAGAATGTGGACTTCCTCATTGCAGAGTATTTTGAACATGTT
GAAGAAGCTGTCTGGGCAGTTGAAGTCCTAAAAGAATCCGGCAAGCCAGTTGCAGCTACAATGTGCATAGGTCCAGAAGGAGA
CATGCATGGTGTGCCTCCTGGACAATGTGCTGTCCAACTGGTGAAGGCTGGTGCTTCCATTGTTGGAGTCAACTGTCATTTTG
ACCCAGATACTGTACTGGAAACTGTGAAACTGATGAAAGAGGGTTTGCAAGCTGCTAAGCTGAAAGCCCACCTGATGTCTCAA
CCACTTGCTTTCCATACACCTGACTGTGGGAAGCAGGGATTTATTGATCTTCCTGAATTTCCCTTTGGTCTGGAACCAAGAAT
CATCACTAGATGGGATGTTCAAAAATACGCAAGAAAGGCTTATGACTTAGGAATCCGCTATATTGGAGGCTGCTGTGGATTTG
AGCCATATCATGTCCGAGCAATAGCTGAAGAGCTGGCTCCTGAAAGAGGATTTTTACCAGAAGCTTCTGAGAAACATGGTAGC
TGGGGCAATAGCCTGAGCATGCATACTAAACCCTGGGTCAGAGCAAGAGCAAGAAAAGAATACTGGGAGAATCTGAAGCCTGC
TTCAGGCAGGCCATATTGTCCTTCAATGTCAAAGCCAGATGGCTGGGGAGTGACCAAAGGAGCCAGGGAGCTGATGCAGCAGA
AAGAAGCAACAACTGAACAGCAGCTGAAGGAGCTCTTCCAGAAACAGAAAGTCTAA

>PsiBHMT1 (Chinese softshell turtle; Accession number XM_006119143)
ATGCAGGCGCCACCTGTGGCAGAGAGCCGCCAGCACGCCAAGAAGGGTATTCTAGAACGGCTAGATGCTGGAGAAATTGTGAT
TGGAGATGGAGGCTTTGTGATAGCTCTTGAAAAGAGAGGTTATGTAAAAGCTGGACCCTGGACTCCAGAAGCTGCAGTAGAAT
ATCCAGAAGCAGTTCGTCAACTCCATCGGGAATTCCTCAGAGCTGGATCAAATGTTTTACAGACATTCAGCTTTTATGCCAGC
GAAGATAAATTAGAGAATAGGGGCAACTACGTAGCTGAGAAAATATCTAGCCAGAGTGTGAATGAAGCTGCTTGTGACATTGC
AAAGGAAGTGGCTAAAGAAGGAGATGCTTTGGTAGCTGGAGGAGTTAGTCAAACACCATCATACCTGAGTGGCAAGGATGAAG
TTGGGGTTAAAGCAATTTTTCAAAAGCAGTTAGAGGTCTTTGTCAAAAAGAATGTGGACTTCTTAATTGCAGAGTATTTTGAA
CATGTTGAAGAAGCAGTATGGGCCGTTGAAACCTTAAAAGTATCTGGAAAACCAGTTGCAGCTACCATGTGCATTGGTCCAGA
GGGAGACTTGCATGGTGTACCACCTGGAGAATGTGCTGTCCAACTGGTTAAAGCTGGTGCCTCTATTATTGGAGTGAACTGCC
ATTTTGACCCAGAGACTTGCCTGAGAACTGTAGAGCTCATGAAAAAGGGCTTAGAGTCTGCTAAATTGAAAGCACACCTGATA

TCTCAACCACTTGCTTACCATACTCCTGACTGTGGGAAGCAGGGATTCATTGATCTCCCAGAATTTCCTTTTGCTCTGGAACC
AAGAATCGTTACCAGATGGGATGTTCAGAAATATGCACGAGAAGCCTATAATCTAGGGATTCGATACATTGGAGGCTGCTGTG
GATTTGAGCCTTATCACATCAGAGCAATAGCTGAGGAGCTGGCCACTGAAAGAGGATTTTTGCCACTGGCTTCTGAGAAACAT
GGTAGCTGGGGTAGTGGTTTGAGCATGCATACCAAACCCTGGGTCAGAGCAAGGGCAAGAAAAGAGTACTGGGAAAATTTGTT
GCCTGCTTCAGGCAGGCCATACTGTCCTTCAATGTCAAAGCCAGATGGCTGGGGAGTGACCAAAGGAACTGCAGAGCTGATGC
AGCAGAAAGAAGCAACAACTGAACAACAGCTGAAAGAGCTCTTTAAGAAACAGAAGTTCAAGTCCAATATAGTAGCTTAA

>HsaBHMT1 (human; Accession number NM_001713)
ATGCCACCCGTTGGGGGCAAAAAGGCCAAGAAGGGCATCCTAGAACGTTTAAATGCTGGAGAGATTGTGATTGGAGATGGAGG
GTTTGTCTTTGCACTGGAGAAGAGGGGCTACGTAAAGGCAGGACCCTGGACTCCTGAAGCTGCTGTGGAGCACCCAGAAGCAG
TTCGCCAGCTTCATCGAGAGTTCCTCAGAGCTGGCTCAAACGTCATGCAGACCTTCACCTTCTATGCGAGTGAAGACAAGCTG
GAGAACAGGGGCAACTATGTCTTAGAGAAGATATCTGGGCAGGAAGTCAATGAAGCTGCTTGCGACATCGCCCGACAAGTGGC
TGATGAAGGAGATGCTTTGGTAGCAGGAGGAGTGAGTCAGACACCTTCATACCTTAGCTGCAAGAGTGAAACTGAAGTCAAAA
AAGTATTTCTGCAACAGTTAGAGGTCTTTATGAAGAAGAACGTGGACTTCTTGATTGCAGAGTATTTTGAACACGTTGAAGAA
GCTGTGTGGGCAGTTGAAACCTTGATAGCATCCGGTAAACCTGTGGCAGCAACCATGTGCATTGGCCCAGAAGGAGATTTGCA
TGGCGTGCCCCCCGGCGAGTGTGCAGTGCGCCTGGTGAAAGCAGGAGCATCCATCATTGGTGTGAACTGCCACTTTGACCCCA
CCATTAGTTTAAAAACAGTGAAGCTCATGAAGGAGGGCTTGGAGGCTGCCCGACTGAAAGCTCACCTGATGAGCCAGCCCTTG
GCTTACCACACTCCTGACTGCAACAAGCAGGGATTCATCGATCTCCCAGAATTCCCATTTGGACTGGAACCCAGAGTTGCCAC
CAGATGGGATATTCAAAAATACGCCAGAGAGGCCTACAACCTGGGGGTCAGGTACATTGGCGGGTGCTGTGGATTTGAGCCCT
ACCACATCAGGGCAATTGCAGAGGAGCTGGCCCCAGAAAGGGGCTTTTTGCCACCAGCTTCAGAAAAACATGGCAGCTGGGGA
AGTGGTTTGGACATGCACACCAAACCCTGGGTTAGAGCAAGGGCCAGGAAGGAATACTGGGAGAATCTTCGGATAGCCTCAGG
CCGGCCATACAACCCTTCAATGTCAAAGCCAGATGGCTGGGGAGTGACCAAAGGAACAGCCGAGCTGATGCAGCAGAAAGAAG
CCACAACTGAGCAGCAGCTGAAAGAGCTCTTTGAAAAACAAAAATTCAAATCACAGTAG

>HsaBHMT2 (Human; Accession number NM_017614)
ATGGCACCTGCTGGACGCCCGGGGGCCAAGAAGGGGATTTTGGAGCGCCTGGAGAGTGGGGAGGTTGTGATTGGAGATGGCAG
CTTTCTCATTACTCTGGAGAAGAGAGGCTATGTGAAGGCTGGGCTCTGGACTCCAGAGGCAGTGATAGAACACCCAGACGCAG
TTCGTCAACTTCACATGGAATTCTTGAGAGCAGGATCAAATGTCATGCAGACTTTTACCTTTTCTGCCAGTGAGGACAATATG
GAAAGCAAGTGGGAAGATGTAAATGCTGCTGCCTGTGACCTCGCCAGGGAAGTGGCTGGCAAAGGTGATGCTTTGGTAGCAGG
GGGGATCTGCCAGACATCAATATACAAATACCAGAAGGATGAAGCTAGAATTAAAAAACTTTTTCGACAACAGCTAGAAGTTT
TTGCCTGGAAAAATGTGGACTTCTTGATTGCAGAGTATTTTGAGCACGTTGAAGAAGCTGTGTGGGCTGTGGAAGTCTTAAAA
GAATCAGATAGACCCGTGGCAGTTACCATGTGCATAGGCCCAGAGGGAGACATGCATGATATAACCCCCGGAGAATGTGCTGT
GAGGCTGGTGAAGGCAGGGGCTTCCATCGTTGGCGTGAACTGCCGCTTTGGGCCCGACACCAGCTTGAAGACGATGGAGCTCA
TGAAGGAGGGTCTTGAGTGGGCAGGGCTGAAAGCGCACCTCATGGTGCAGCCTCTGGGGTTCCACGCGCCTGACTGTGGCAAA
GAGGGGTTTGTGGATCTCCCAGAATATCCCTTTGGACTGGAGTCCAGAGTTGCCACCAGATGGGATATTCAAAAATACGCCAG
AGAGGCCTACAACCTGGGGGTCAGGTACATTGGCGGGTGCTGTGGATTTGAGCCCTACCACATCAGGGCAATTGCAGAGGAGC
TGGCCCCAGAAAGGGGCTTTTTGCCACCAGCTTCAGAAAAACACGGCAGCTGGGGAAGTGGTTTGGACATGCACACCAAACCC
TGGATTAGAGCAAGGGCTCGAAGGGAGTATTGGGAGAATCTGCTGCCAGCTTCAGGCAGACCTTTCTGTCCTTCGCTGTCAAA
GCCAGACTTCTAA

>MdoBHMT1 (gray short-tailed opossum; Accession number XM_001381549)
ATGGTACCTGCTGGAGGCAAAAAGATGAAGAAGGGAATTCTCCAACGTTTGGATTCTGGAGAGATTGTAATTGGAGATGGAGG
ATTTGTCTTTGCCCTTGAGAAGAGAGGATATGTGAAAGCTGGGCCCTGGACCCCAGAAGCTGCAGTTGAACACCCAGAAGCAG
TTCGACAGCTTCATCGGGAATTCCTTCGAGCAGGCTCAAATGTCATGCAGACCTTCACCTTCTATGCTAGTGAGGACAAGTTG
GAGAACAGGGGAAATTATGTTGCAGAGAAAATATCTGGGCAGAAAGTGAATGAAGCTGCCTGTGACATTGCCCTGCAAGTGGC
TGCAGAAGGTGATGCTTTGGTTGCTGGTGGAGTTAGTCAGACACCATCATACCTGAGTTGCAAGAGTGAAACTGAAGTGAAAA
AAATTTTCCGTCAACAGTTAACAGTCTTTATGAAAAAGAATGTGGACTTTCTGATTGCAGAGTATTTTGAACATGTTGAAGAA
GCCGTGTGGGCAGTTGAAGCCTTAAAAGAGTCTGGAAAGCCAGTGGCTGCTACAATGTGCATTGGCCCAGAAGGAGATCTGCA
TGGTGTGACGCCTGGAGAATGTGCAGTGCGGCTGGTTAAAGCAGGAGCTTCCATTGTTGGTGTGAATTGCCATTTTGATCCTA
CAATTAGTTTAAAAACTGTGAAACTCATGAAAGAGGGATTGGAGGCTGCCAAATTGAAAGCCTATTTGATGACACAGCCCTTG
GCTTACCATACTCCAGATTGTGGCAAACAAGGATTTATTGATCTTCCAGAATTTCCTTTTGGTCTGGAACCCAGAGTTGCTAC
CAGATGGGATATTCAAAAATATGCCAGGGAAGCCTACAAGCTGGGTGTCAGGTACATCGGGGGCTGCTGCGGATTTGAGCCTT
ATCACATCAGAGCGATTGCCGAGGAGCTGGCTCCTGAAAGGGGCTTCTTGCCTGATGCTTCTGAAAAACATGGCAGCTGGGGG
AGTGGCTTGAATATGCACACCAAGCCCTGGGTTAGAGCAAGGGCAAGGAGAGAGTACTGGGAGTCACTACCGCTGGCTTCTGG
TAGGCCATACTGTCCTTCTATGTCCAAACCTGATGCCTGGGGAGTTACCAAAGGAACTGCAGAGCTCATGCAGCAGAAGGAAG
CCACAACTGACCAACAACTGAAGGAACTCTTTGAGAAACAGAAGTTCAAATCTTCGGTAGCCTAA

>MdoBHMT2 (gray short-tailed opossum; Accession number XM_001367487)
ATGGCACCCGTCGGAGGGAAAAGGGCAAAGAAGGGCATTCTAGAACGGCTGCATAGTGGAGAGGTTGTGATTGGTGATGGAAG
TTTTCTCTTTACCCTTGAGAAGAGGGGGTATGTGAAAGCTGGACTTTGGACCCCAGAGGCAACTGTAGAGCATCCAGAAGCAG
TACGCCAGCTTCATACAGAGTATCTGAGAGCTGGATCCAATGTCATGCAGACCTTTACCTTCTTTGCTGGCAAGGACAATTTG
GAAAGCAAGTGGAAAGAAGTAAACGAAGCTGCCTGCGACCTTGCCAGGGAAGTGGCTGAAAAGGGAGATGCCTTGGTAGCAGG
TGGAATCAGTCAAACGTCATTGTACAAAAGTCATGGCAGTGAACCTGAAATTAAAGAAAATTTCCAACTACAACTAGAAGTTT
TTACTAGAAAGAATGTGGACTTCTTGATTGCAGAGTATTTTGAGTATTGTGAAGAAGCTGTATGGGCTGTTGAAGTTTTGAAA
GGGTCTGGGAAGCCAGTGGCAGCTACCATGTGCATTGGTCCAGAGGGAGACATGAATAATCTAACACCTGGAGAATGCGCAGT
TCGACTGGTCAAAGCAGGAGCTTCAATCGTGGGTGTGAACTGTCGCTTTGGACCAGAGATTAGCTTGGCCACTGTGAAGCTAA
TGAAAGAGGGTCTAGAAGCCTCTGGCCTGAAAGCTTACTTGATGGTACAGTCACTGGGGTTCCACACACCTGACTGTGGCAAG
GGAGGCTTTGTAGACCTCCCCCAGTACCCTTTTGGTATGGAACCCAGAGTTGCTACCAGATGGGATATTCAAAAATATGCCAG

APPENDIX

GGAAGCCTACAAGCTGGGTGTCAGGTACATCGGGGGCTGCTGCGGATTTGAGCCTTATCACATCAGAGCGATTGCCGAGGAGC
TGGCTCCTGAAAGGGGCTTCTTGCCTGATGCTTCTGAAAAACATGGCAGCTGGGGGAGTGGCTTGAATATGCACACCAAGCCT
TGGGTTAGAGCAAGGGCACAGAAAGAATACTGGAAGAATTTGTTGCCAGCTTCAGGAAGGCCATACTTTCCTTCTTTGTCAAA
ACCAAATGCCTGA

>MmuBHMT1 (mouse; Accession number BC037004)
ATGGCACCAGTTGCTGGCAAGAAGGCCAAGAAGGGGATCTTAGAACGCTTAAATGCCGGAGAAGTTGTGATTGGAGATGGAGG
ATTTGTCTTTGCACTGGAAAAGAGGGGCTATGTAAAGGCTGGACCCTGGACCCCAGAAGCTGCCGTGGAGCATCCTGAGGCAG
TTCGTCAGCTTCATCGGGAGTTCCTCAGAGCTGGATCGAACGTCATGCAGACCTTCACTTTCTATGCAAGTGAAGACAAGCTG
GAAAACAGAGGAAACTATGTGGCAGAGAAGATTTCTGGGCAGAAAGTCAACGAAGCTGCTTGTGACATTGCACGGCAAGTGGC
TGATGAAGGAGACGCTTTGGTTGCAGGAGGCGTGAGCCAGACGCCTTCATACCTTAGCTGCAAGAGTGAGGTAGAAGTGAAAA
AGATATTTCGCCAACAGCTAGAGGTGTTCATGAAGAAGAACGTGGACTTCCTCATTGCAGAGTATTTTGAACATGTTGAAGAA
GCCGTGTGGGCAGTGGAAGCCTTAAAAGCATCTGGTAAGCCCGTAGCAGCTACCATGTGCATTGGGCCCGAGGGAGATCTGCA
TGGCGTGCCCCCTGGAGAGTGTGCCGTGCGTCTGGTGAAAGCAGGCGCCTCCATTGTCGGCGTGAACTGCCACTTCGACCCCA
GCGTCAGCTTACAGACTGTGAAGCTCATGAAGGAGGGTTTGGAGGCTGCGCGGTTGAAAGCTTACCTGATGAGCCAGCCCCTG
GCCTACCATACCCCTGACTGTGGCAAACAGGGATTTATTGATCTCCCAGAATTCCCCTTTGGATTGGAACCCCGAGTTGCCAC
TAGATGGGATATTCAAAAATATGCCAGAGAGGCCTACAACCTGGGGGTTAGGTACATTGGCGGCTGCTGCGGATTTGAGCCCT
ACCACATCAGGGCGATTGCAGAGGAGTTGGCCCCAGAAAGGGGATTTTTGCCACCGGCTTCAGAAAAACATGGCAGCTGGGGA
AGTGGTTTGGACATGCACACCAAACCCTGGATCAGAGCAAGGGCCAGAAAGGAATACTGGCAGAATCTGCGAATAGCTTCCGG
CAGGCCGTACAACCCTTCCATGTCCCGGCCAGATGCTTGGGGCGTGACTAAGGGAGCAGCCGAGCTGATGCAGCAGAAGGAGG
CCACTACTGAGCAGCAGCTGAGAGAGCTCTTTGAAAAACAAAAATTCAAGTCTGCACAGTAG

>MmuBHMT2 (mouse; Accession number BC013515)
ATGGCACCAGCTGGAAGCACTCGAGCCAAGAAGGGCATCTTGGAGCGTCTGGACAGCGGGGAGGTGGTGGTTGGGGACGGCGG
CTTTCTCTTCACTCTGGAAAAGAGAGGCTTCGTGAAGGCGGGGCTTTGGACTCCAGAAGCAGTGGTAGAACATCCAAGTGCAG
TTCGTCAACTTCACACAGAATTCTTGAGAGCGGGAGCTGATGTCTTGCAGACATTCACCTTTTCTGCTACTGAAGACAATATG
GCCAGCAAGTGGGAAGCTGTGAATGCAGCTGCCTGTGACCTGGCCCAGGAGGTGGCTGGTGGAGGGGGTGCTTTGGTGGCAGG
GGGCATCTGCCAGACATCACTCTACAAGTACCACAAGGATGAAACTAGAATTAAAAACATTTTTCGACTACAGCTAGAAGTTT
TTGCCAGGAAAAATGTGGACTTCTTGATTGCAGAGTATTTTGAGCATGTGGAAGAAGCTGTATGGGCTGTCGAAGTCTTGAGA
GAGGTGGGGGCACCTGTGGCTGTGACCATGTGCATCGGCCCAGAGGGGGACATGCACGACGTGACACCTGGAGAGTGTGCGGT
GAAGCTGGCGCGTGCAGGGGCGGACATCATTGGGGTCAACTGCCGGTTTGGGCCCTGGACTAGCTTGCAGACCATGAAGCTCA
TGAAAGAGGGCCTCAGAGATGCCAGCCTGCAGGCTCACCTTATGGTCCAGTGCTTGGGCTTCCACACACCGGACTGTGGCAAG
GGAGGGTTTGTGGACCTCCCAGAGTATCCCTTTGGCCTGGAGCCAAGAGTTGCCACCAGATGGGATATTCAAAAATACGCCAG
AGAGGCCTACAACCTGGGGATCAGGTACATTGGCGGCTGCTGCGGATTTGAGCCCTACCACATCAGGGCGATTGCAGAGGAAC
TGGCCCCAGAAAGGGGATTTTTGCCACCGGCTTCAGAGAAACATGGCAGCTGGGGAAGTGGTCTGAACATGCACACCAAACCC
TGGATCAGAGCAAGGGCTAGACGAGAATACTGGGAGAATCTGCTGCCAGCCTCAGGAAGACCTTTCTGTCCTTCCCTATCAAA
GCCAGATGCATAA

>OanBHMT1 (Platypus; Accession number XM_007662293)
ATGGGTGACAACCACCTGGTCAAGACCATGGGTCTTCCTCGCTGGTTCTACATAATGCCACTCTCTAGGCCTCCAAAAGAAAA
TACCTCCAGGGGTCTGAAGCTGAGTAGGGGGACTTTCCTCACAGAGAAAGGTGGAGAACGCCTGAATGGCCGACCCCTGCGTC
GGCTGAGCACGGACCGCCCCGACCTGCGGCCCGGGCGCTCGGAATGTATTCCCACCCGGGTGACTCTGGGGGCTTCTTTCCCG
CTTCCTTTGCAGGGGATTCTAGAACGGTTGAATGGGGGTGAGGTGGTGATTGGAGATGGAGGGTTTGTCTTTGCCCTGGAGAA
GAGGGGGTATGTGAAAGCTGGACCCTGGACCCCGGAAGCGGCTGTAGAACACCCGGAAGCTGTTCGTCAGCTTCACCGAGAGT
TCCTCAGATCTGGATCAAATGTTATGCAGACTTTCACCTTTTATGCAAGTGAGGACAAGTTGGAGAATAGAGGAAATTATGTT
GCGCAGAAAATATCTGGGCAGAAGGTGAATGAAGCTGCGTGTGACATTGCCCGGCAAGTGGCTGATGAAGGAGACGCCTTGGT
AGCCGGAGGCGTCAGTCAGACCCCATCCTACCTCAGTTGCAAGAGTGAAACTGAAGTGAAAAAAATTTTTCGGCAGCAGCTAG
AGGTCTTCACCAAAAAAAATGTGGACTTCTTGATTGCAGAGTATTTTGAGCACGTCGAAGAAGCCGTGTGGGCAGTTGAAGCT
TTAAAGGAGTCTGGGAAGCCGGTGGCGGCTACCATGTGCATCGGTCCGGAGGGAGATCTGCACGGAATTCCGCCCGGAGAATG
TGCTGTCCGGCTGGTCAAGGCTGGAGCTTCTATTGTGGGCGTGAACTGCCATTTTGATCCTACGACTAGCCTGAAAACCGTGA
AGCTTATGAAAGAGGGGTTGGAGAGAGCAAAATTGAAAGCTCACCTGATGGCCCAGCCACTGGCTTACCACACTCCTGACTGT
GGCAAACAGGGCTTTATTGACCTCCCAGAGTTTCCTTTTGGCCTGGAACCCAGAGTCACCACCCGGTGGGACATCCAGAAATA
CGCCCGCGAGGCCTACGAGCTGGGCGTGAGGTACATCGGCGGCTGCTGTGGGTTTGAGCCCTACCACGTGCGAGCCATCGCCG
AGGAGCTGGCCCCGGAGCGGGGGTTTCTGCCGCCGGCGTCCGACAAGCACGGGAGCTGGGGCAGCGGCCTGGACATGCACACC
AAACCCTGGATTAGGGCCAGGGCGAGGAAGGAGTATTGGCAGAATCTGCCGCTCGCTTCAGGCAGACCGTACTGTCCTTCAAT
GTCTCGGCCAGATGCCTGGGGCGTGACCAAAGGGACCGCTGAGCTGATGCAGCAGAAAGAAGCAACCACTGACCAACAGCTGA
AGGACCTCTTTCAGAAACAGAGGTTCAAGTCGGGAGTAGTTTAA

>OanBHMT2 (Platypus; Accession number XM_007662286)
ATGCAGACATTTACCTTTTTCGCTGGGGAGGACTCTCTGGACTGCAAGTGGAAGAAAATCAACGAAGCGGCGTGTGACCTCGC
CAGAGAAGTAGCCAAAGACGGAGATGCCCTGGTAGCTGGGGGGATCAGCCAGACCCCATCGTACAAAAGTCGGAAGGGCGAAG
CGGAAATTAAAAAAGTTTTTCAACAGCAACTGGAGGTCTTCGTCCGAAAGAACGTGGACTTCCTGATTGCGGAGTATTTTGAG
CACGTCGAAGAAGCCGTGTGGGCTGTTGAAGTCTTAAAGGCGTCTGGGAAGCCAGTGGCGGCTACCATGTGCATCGGTCCGGA
GGGAGATATGCACGGAATTCCTCCCGGAGAATGTGCCGTCCGGCTGGTCAAGGCTGGAGCCTCGCTCGTTGGAGTGAACTGCC
GTTTCGGGCCCGAGACTAGCTTGAAAACCGTGTCGCTCATGAAGCAGGGCTTGAAGTCTGCTGGCTTGAAAGCTCACCTCATG
GTCCAGACGCTGGCATTCCACACCCCTGACTGTGGCAAAGCCGGATTTGTTGATCTCCCGGAATATCCTTTCGGCCTGGAACC
CAGAGTCGCCAGCCGATGGGACATCCAGAAATACGCCCGCGAGGCCTACGAGCTGGGCGTGAGGTACATCGGCGGCTGCTGTG
GGTTTGAGCCCTACCACGTCCGAGCCATCGCCGAGGAGCTGGCCCCGGAGCGGGGGTTTCTGCCGCCGGCGTCCGACAAGCAC

GGGAGCTGGGGCAGCGGCCTGGACATGCACACCAAACCCTGGATTAGGGCCAGGGCGAGGAAGGAATACTGGCAGAATCTGCC
GCTGGCTTCAGGCAGACCGTACTGTCCTTCACTGTCGAGGCCAGATGCCTGA