UNIVERSIDADE DO ALGARVE

EVOLUTION, GENE REGULATION AND FUNCTIONAL ANALYSIS OF BMP2 IN FISH

Cátia Andreia Lourenço Marques

Tese para a obtenção do grau de Doutor em Ciências Biomédicas

Trabalho efetuado sob a orientação de:

Professora Doutora Leonor Cancela

Doutor Vincent Laizé

2014

UNIVERSIDADE DO ALGARVE

EVOLUTION, GENE REGULATION AND FUNCTIONAL ANALYSIS OF BMP2 IN FISH

Cátia Andreia Lourenço Marques

Tese para a obtenção do grau de Doutor em Ciências Biomédicas

Trabalho efetuado sob a orientação de:

Professora Doutora Leonor Cancela

Doutor Vincent Laizé

2014

Evolution, gene regulation and functional analysis of BMP2 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 referencias incluída.

_____

À Joaninha

**Almost all aspects of life are engineered at the molecular level,
and without understanding molecules we can only
have a very sketchy understanding of life itself**

Francis Crick

AGRADECIMENTOS

A realização e conclusão deste trabalho só foi possível com a ajuda e colaboração de algumas pessoas, às quais gostaria de demonstrar a minha gratidão. A todas elas o meu Muito Obrigada!

Em primeiro lugar gostaria de agradecer à Professora Leonor Cancela pelos conhecimentos transmitidos ao longo de todos estes anos, pelo empenho e preocupação. Obrigada por me ter recebido no seu laboratório e por ter permitido a realização deste trabalho.

Ao Doutor Vincent Laizé, meu co-orientador, a quem não tenho palavras para expressar minha gratidão. Obrigada pela enorme disponibilidade, pela motivação e incentivos constantes e por toda a ajuda prestada ao longo de todos estes anos e que foram fundamentais para a realização deste trabalho. Muito Obrigada!

A todos os meus colegas de laboratório, os de hoje, os de ontem, os de sempre! Obrigada pela partilha, pelos momentos bons e por estarem lá para me aturar nos dias menos bons! À Natércia e ao Gavaia por estarem disponíveis sempre que vos pedi ajuda; ao Daniel por toda a ajuda no esclarecimento de dúvidas e pelas discussões entusiásticas!

À Helena, por me mostrar que ainda existem boas pessoas; à Iris, por ser uma miúda Muito fixe; à Andreia pela boa disposição e alegrias constantes (esse teu sexto sentido mata-me!); ao Mike, pela amizade ao longo de todos estes anos (gosto de ti, pá!); ao Nacho pela partilha de conhecimentos, dentro e fora do lab. Obrigada espanholito pelo incentivo! À Sara por me chamar à terra, por pôr sempre os pontos no "is", pelo incentivo; ao Ricardo Leite, pelos ensinamentos e pela ajuda; à Anabela pelas discussões sempre com um brilhozinho nos olhos. A ciência é linda!

À Vânia, a quem a máxima "primeiro estranha-se, depois entranha-se" se aplica na perfeição. Obrigada por te teres tornado uma presença constante na minha vida, obrigada pelos momentos de partilha e pela ajuda nos momentos menos bons! À Joana R., de quem aprendi a gostar Muito, obrigada pela espontaneidade, disponibilidade, compreensão, partilha e ajuda preciosa ao longo de todos estes anos. À Rafa por estar sempre lá, pela compreensão, pela amizade! À Bri por todos os momentos de partilha ao longo destes anos, pela boa disposição e presença constantes, mesmo apesar da distância. Muito Obrigada meninas, assim é muito mais fácil!

À Bi, amiga de sempre. Obrigada por me ajudares a por as coisas em perspetiva e por me obrigares a relatar o progresso em percentagem (cheguei finalmente aos 100%)!

À Gi pela cumplicidade desde sempre.

À Ondina, pela disponibilidade, pela paciência e por toda a ajuda, em particular nestes últimos tempos.

Ao meu cunhado, obrigada pela ajuda e acima de tudo pela disponibilidade constantes.

À minha irmã, pela ajuda nos momentos mais difíceis (obrigada pelas tuas segundas-feiras), pela cumplicidade, pela partilha, pela presença tão constante na minha vida. É bom sabermos que não estamos sozinhos!

Aos meus pais, pelos valores transmitidos durante toda a vida e pelo apoio e amor incondicionais. Obrigada por acreditarem em mim e por vibrarem com minhas conquistas. A conclusão deste trabalho não seria possível sem vocês.

Ao João pelo amor, disponibilidade e compreensão incondicionais. Obrigada pela enorme ajuda, em especial nestes últimos momentos. Obrigada por estares sempre ai e por amparares todos os golpes. Sem ti tudo teria sido muito mais difícil.

À minha Joana, por tudo! Obrigada por me teres ensinado a viver mais intensamente, por fazeres com que tudo valha sempre muito mais a pena. Obrigada pela alegria, pela espontaneidade. Vamos brincar?!

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/39964/2007.

Muito obrigada!

ABSTRACT

Bone morphogenetic proteins (BMPs) are multifunctional growth factors belonging to the transforming growth factor β (TGFβ) superfamily with a central role in bone formation and mineralization. BMP2, a founding member of this family, has demonstrated remarkable osteogenic properties and is clinically used to promote bone repair and fracture healing. Lack of basic data on factors regulating BMP2 expression and activity have hampered a better understanding of its role in bone formation and bone-related diseases. The objective of this work was to collect new functional data and determine spatiotemporal expression patterns in a fish system aiming towards a better understanding of BMP2 function and regulation. Transcriptional and post-transcriptional regulation of gilthead seabream BMP2 gene was inferred from luciferase reporter systems. Several bone- and cartilage-related transcription factors (e.g. RUNX3, MEF2c, SOX9 and ETS1) were found to regulate BMP2 transcription, while microRNA 20a was shown to affect stability of the BMP2 transcript and thus the mineralogenic capacity of fish bone-derived host cells. The regulation of BMP2 activity through an interaction with the matrix Gla protein (MGP) was investigated *in vitro* using BMP responsive elements (BRE) coupled to luciferase reporter gene. Although we demonstrated the functionality of the experimental system in a fish cell line and the activation of BMP signaling pathway by seabream BMP2, no conclusive evidence could be collected on a possible interaction beween MGP and BMP2. The evolutionary relationship among the members of BMP2/4/16 subfamily was inferred from taxonomic and phylogenetic analyses. BMP16 diverged prior to BMP2 and BMP4 and should be the result of an ancient genome duplication that occurred early in vertebrate evolution. Structural and functional data suggested that all three proteins are effectors of the BMP signaling pathway, but expression data revealed different spatiotemporal patterns in teleost fish suggesting distinct mechanisms of regulation. In this work, through the collection of novel data, we provide additional insight into the regulation, the structure and the phylogenetic relationship of BMP2 and its closely related family members.

Keywords: bone morphogenetic proteins (BMPs), BMP signaling pathway, matrix Gla protein (MGP), transcriptional and posttranscriptional regulation, molecular phylogeny, gene expression patterns, protein structure modelling

RESUMO

O sistema esquelético confere suporte e proteção ao organismo, permite o armazenamento de minerais e desempenha funções hematopoiéticas. Um dos seus principais componentes é o osso, um tecido conectivo especializado e constituído por uma matriz extracelular extensamente mineralizada. O processo de mineralização envolve mecanismos extremamente complexos que estão sujeitos a um rigoroso controlo a nível molecular, no qual estão envolvidas várias proteínas responsáveis pela diferenciação celular e pela síntese de matriz extracelular. Dentro deste conjunto de proteínas, destacam-se alguns fatores de crescimento essenciais ao mecanismo de mineralização tecidular, como é o caso das proteínas morfogenéticas do osso (BMPs). As BMPs pertencem à superfamília de fatores de crescimento de transformação β (TGFβ) e estão envolvidas em vários processos durante a embriogénese, organogénese, proliferação e diferenciação celular e mecanismos de formação óssea. Atualmente estão descritos e caracterizados mais de vinte membros pertencentes a esta família, que foram divididos em várias subfamílias, de acordo com a semelhança da estrutura primária das suas proteínas. A subfamília BMP2/4/16 à qual pertence a BMP2 foi uma das primeiras a ser identificada e caracterizada. A BMP2 é uma das proteínas que possui uma maior capacidade osteogénica e é, desde há muito tempo, considerada um potencial agente terapêutico para o tratamento de doenças relacionadas com o osso, sendo mesmo utilizada em alguns casos clínicos de fraturas ósseas. O conhecimento dos processos de formação óssea, bem como os mecanismos de regulação da BMP2, são por isso de extrema importância para uma melhor compreensão dos processos subjacentes ao desenvolvimento e progressão de algumas doenças ósseas. Assim, a BMP2 tem sido alvo de vários estudos, quer a nível de conhecimento da sua função, quer a nível de mecanismos de ação e processamento. Sabe-se que a BMP2 é uma proteína secretada para a matriz extracelular onde, através de um mecanismo de sinalização molecular, é responsável pela regulação de vários processos. O mecanismo de sinalização inicia-se quando dímeros de BMP2 se ligam aos respetivos recetores, presentes na superfície da célula, ativando assim uma cascata de sinalização molecular. Através de diferentes intermediários intracelulares, envolvidos na cascata de sinalização, a BMP2 é responsável pela regulação transcricional de vários genes-alvo. No entanto, e apesar dos vários estudos que foram feitos nesta área, o conhecimento existente acerca deste assunto é ainda bastante escasso. Neste sentido, o objetivo principal deste trabalho foi a recolha de novos dados funcionais e estruturais, que permitam uma melhor compreensão da função da BMP2. Para tal, e de modo a complementar o conhecimento

existente, utilizámos o peixe como modelo alternativo aos sistemas de mamíferos. O peixe é atualmente reconhecido como um modelo válido para estudos do esqueleto de vertebrados para o qual existem já várias ferramentas que permitem análises in silico, *in vitro* e *in vivo*. Este trabalho envolveu o estudo da regulação do gene da BMP2 de dourada, tanto a nível transcricional como a nível pós-transcricional. Numa primeira fase, foram identificados potenciais reguladores transcricionais da BMP2 de dourada, através da análise in silico da região reguladora do gene. Dentro dos potenciais reguladores transcricionais, foram identificados vários fatores de transcrição com funções descritas ao nível do osso e da cartilagem, nomeadamente o RUNX3, SOX9, MEF2C e ETS1, que foram posteriormente testados a nível funcional através de ensaios repórter de luciferase. Em paralelo, no decorrer da caracterização de reguladores pós-transcricionais da BMP2, através da análise da região 3' não traduzida (3'UTR) do seu mRNA, foi possível identificar um local de ligação para o miR-20a, conservado ao longo da evolução. A fim de melhor compreender os mecanismos de ação da BMP2, neste trabalho investigámos também possíveis parceiros desta proteína. A caraterização da interação entre a BMP2 e a proteína Gla da matriz (MGP), um conhecido inibidor da calcificação, foi avaliada através do uso de um sistema de elementos de resposta às BMPs, acoplado a um gene repórter, a luciferase. Embora tenhamos demonstrado a funcionalidade do sistema através da ativação do mecanismo de sinalização celular pela BMP2 de dourada, não foram obtidos dados conclusivos no que diz respeito à interação entre a BMP2 e a MGP. Finalmente, abordamos o aspecto evolutivo dos membros da subfamília BMP2/4/16 através da análise da sua distribuição taxonómica entre vários organismos vertebrados, bem como as relações filogenéticas existentes entre os vários membros desta subfamília. Foi demonstrado que a BMP16 divergiu antes da BMP2 e BMP4 na linhagem dos vertebrados e foi, provavelmente, o resultado de uma duplicação genómica que terá ocorrido ancestralmente. Dados estruturais sugerem uma conservação funcional das três proteínas, facto que foi confirmado pela capacidade de ativação dos mecanismos de sinalização das BMPs. No entanto, e apesar da conservação ao nível da região codante dos genes das BMP2, BMP4 e BMP16, as regiões não traduzidas são substancialmente diferentes, apontando para uma regulação diferencial dos três genes, como é aliás sugerido pelos distintos padrões de expressão observados para a BMP2, BMP4 e BMP16, tanto em linguado como em peixe zebra.

Ao longo deste trabalho foram recolhidos novos dados que permitem uma melhor compreensão da função e regulação da BMP2. Foram igualmente obtidas informações relevantes acerca da filogenia molecular dos membros da subfamília das BMP2/4/16 que

contribuíram para uma melhor compreensão e interpretação da complexa história evolutiva desta subfamília. No seu conjunto, os resultados deste trabalho contribuem para uma validação do uso dos peixes como um modelo alternativo na investigação de mecanismos moleculares envolvidos no processo de mineralização tecidular.

Palavras-chave: proteínas morfogenéticas do osso (BMPs), proteína Gla da matriz (MGP), Mecanismo de sinalização das BMPs, regulação transcricional e pós transcricional, filogenia molecular, padrões de expressão genética, modelação da estrutura proteica

TABLE OF CONTENTS

ABBREVIATION LIST

BMP	Bone morphogenetic protein
BMPR	BMP receptor
BRE	BMP responsive element
CBFß	Core binding factor ß
DPP	Decapentaplegic
ECM	Extracellular matrix
ETS-1	v-ets avian erythroblastosis virus E26 oncogene homolog 1
Fluc	Firefly luciferase
GFP	Green fluorescent protein
hr-BMP	human recombinant BMP
MAPK	Mitogen-activated protein kinase
MEF	Myocyte enhancer factor
MGP	Matrix Gla protein
MiRNA	MicroRNA
qPCR	quantitative real-time PCR
RA	Retinoic acid
RLuc	Renilla luciferase
R-SMAD	Receptor-regulated SMAD
RUNX	Runt-related transcription factor
SMAD	Contraction of SMA (small body size) and MAD (mothers against decapentaplegic)
SOX	Sex determining region Y box
SPCs	Subtilisin-like pro-protein convertases
TGF-β	Transforming growth factor β
TF	Transcription factor
TSS	Transcription start site
UTR	Untranslated region

PREAMBLE

This thesis is divided into five chapters and a list of references. 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 addresses the regulation of BMP2 gene expression at transcriptional and post-transcriptional levels and is based on manuscripts submitted to Gene and published in the Archives of Biochemistry and Biophysics, respectively. Our attempt to get insights into the residues/domains involved in BMP2-MGP interaction is described in the third chapter (unpublished data). A comparative analysis of BMP2, BMP4 and BMP16 subfamily members giving a molecular and an evolutionary perspective to this work is presented in the fourth chapter, which is based on manuscripts published in the Journal of Applied Ichthyology and submitted to Cellular and Molecular Life Sciences. Finally, chapter five gathers the main conclusions drawn from the data presented in this thesis and presents perspectives for future works. In order to have a better contextualization, a preamble with a brief description of the objectives of the work will be presented at the beginning of chapters 2, 3 and 4.

Chapter 1

GENERAL INTRODUCTION

©Paulo Gavaia

CHAPTER 1. GENERAL INTRODUCTION

1.1 Bone formation and function

Bone is a dense connective tissue that forms part of the vertebrate endoskeleton (it is also present in turtle and armadillo exoskeletons). It is composed by an organic collagenous matrix (mainly type I collagen) extensively mineralized with an inorganic hydroxyapatite lattice (calcium/phosphate mineral) that together contributes to bone elasticity (although to a limited extent) and stiffness. Bone serves multiple functions such as body support, locomotion, internal organ protection, calcium and phosphorus storage and balance, growth factor production and, in mammals, bone is also the primary hematopoietic organ (Kronenberg, 2003; Pirraco et al., 2010; Rameshwar and Stegemann, 2013). Recently, the function of endocrine organ has also be assigned to bone through the secretion of osteocalcin which regulates glucose homeostasis and male fertility in mice (Karsenty and Oury, 2014)

Bone formation occurs through two distinct processes: (1) intramembranous ossification, where bone is formed directly from connective tissue: mesenchymal cells aggregate and differentiate into osteoprogenitor cells then into osteoblasts, and produce 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; Karsenty, 2003); (2) Endo/perichondral ossification, where mesenchymal cells aggregate and differentiate into chondrocytes that produce a matrix rich in type II collagen. After matrix enlargement, chondrocytes become hypertrophic and synthesize a matrix rich in type X collagen; once hypertrophic chondrocytes have produced a net of blood vessels (by secreting vascular growth factors), recruited chondroclast (responsible for collagen matrix degradation) and directed adjacent cells to become osteoblasts, they undergo apoptotic cell death. The cartilaginous matrix left behind will provide a scaffold for the mineral deposition by the bone cells (reviewed in Kronenberg, 2003).

There are three main types of cells in bone (Mackie, 2003; Fig. 1.1): (1) Osteoblasts (bone forming cells), which are specialized mesenchymal cells that undergo a maturation process where transcription factors such as Runt related factor 2 (RUNX2) and Osterix (OSX or SP7) play a determinant role. Osteoblasts are found on bone surfaces and are responsible for the deposition of the osteoid, an unmineralized matrix that will gradually mineralize to form bone. 2) Osteocytes, which are osteoblasts that differentiate into interconnected star-shaped cells when they get entrapped within the mineral matrix. Osteocytes are the most abundant cells in bone, where they act as mechano-sensors and master regulators of bone

remodeling (by secreting factors that regulates the activity of both osteoblasts and osteoclasts); Osteoblasts have also the ability to regulate osteoclast bone resorption activity, through the secretion of specific factors such as receptor activator of NF-kappa-B ligand (RANKL) and osteoprotegerin (OPG), recognized by osteoclasts; (3) Osteoclasts (bone resorbing cells), which are large multinucleated cells with origin in the monocyte-macrophage lineage. They result 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. After completing their function osteoclasts undergo apoptosis; this is a control mechanism to avoid excessive bone resorption (Dallas et al., 2013; Pirraco et al., 2010; Nakamura et al., 2012; Caetano-Lopes et al., 2007).

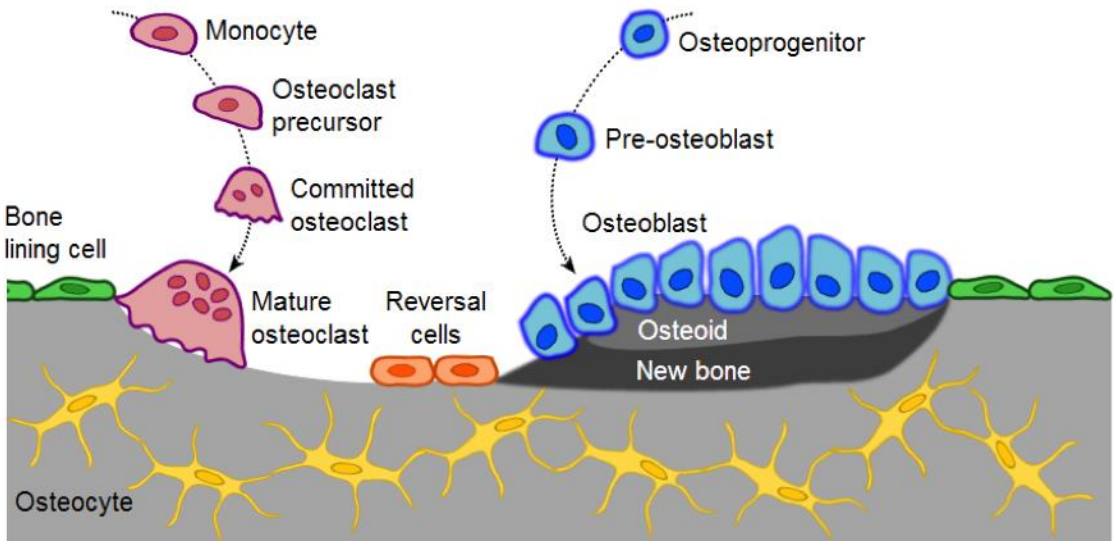


Figure 1.1. Schematic representation of bone cells involved in bone remodeling. Osteocytes (star-shaped yellow cells) are embedded within the mineralized bone matrix and connected through a complex network of cytoplasmic extensions inside lacunae and canaliculi. They are actively involved in bone turnover 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.

Depending on the presence or absence of osteocytes, bone can be classified as cellular/osteocytic bone or acellular/anosteocytic bone, respectively (Horton and Summers, 2009; Fig. 1.2). Osteocytic bone is present in tetrapods, basal teleosts and primitive

osteichthyans, while anosteocytic bone is restricted, with few exceptions, to the skeleton elements of advanced teleosts (Meunier and Huysseune, 1992; Cohen et al., 2012; Kranenbarg et al., 2005). Although they lack osteocytes, responsible for directing bone remodeling by regulating both osteoblasts and osteoclast functions, 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).

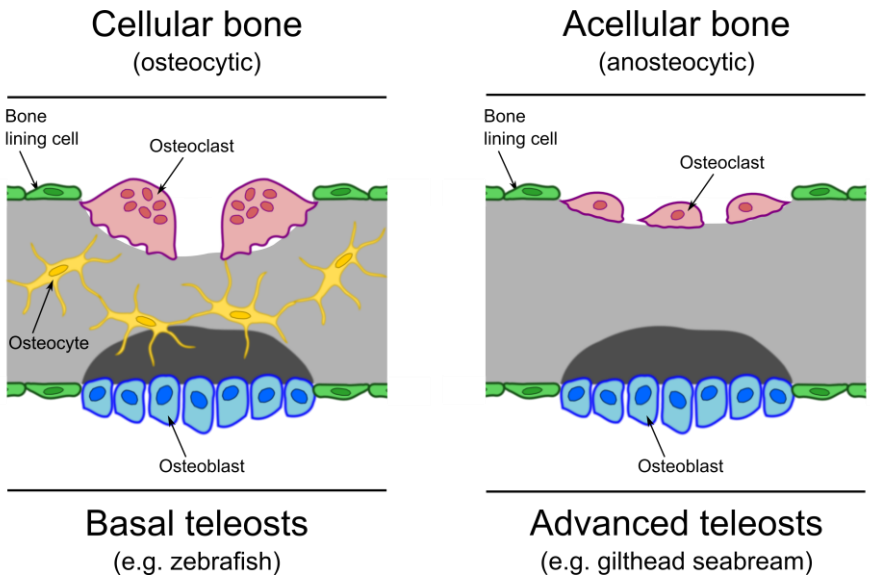


Figure 1.2 Main differences between osteocytic and anosteocytic bone. Osteocytes (star-shaped yellow cells), embedded in bone matrix are only present in osteocytic bone. In organisms that possess anosteocytic bone, osteoclasts (pink) are usually mononucleated and have a limited capacity of bone resorption, creating a shallow lacunae, contrasting with the giant multinucleated cells found in osteocytic bone which produce a 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.

Throughout adult life, and in order to maintain skeletal integrity, the skeleton undergoes continuous remodeling. Bone remodeling is an active and dynamic process and relies on the correct balance between bone resorption by osteoclasts and bone formation by osteoblasts (Crane and Cao, 2014). Osteocytes are also central to bone remodeling by orchestrating the function of osteoblasts and osteoclasts (Bellido, 2014). Bone remodeling is composed of three main phases: (1) the activation phase, where different stimuli (micro-fractures, alteration of mechanical load or the release of some factors – e.g. insulin growth factor-I (IGF1), tumor necrosis factor-α (TNF-α) and parathyroid hormone (PTH) – to bone microenvironment) lead to the activation of the lining cells (resting osteoblasts) that will increase the expression of RANKL and promote osteoclast differentiation; (2) the resorption phase, where active

osteoclasts resorb bone by acidifying bone matrix and secreting proteases that will be responsible for the degradation of the organic matrix of bone; (3) The formation phase starts when growth factors stored in bone – fibroblast growth factors (FGFs), transforming growth factors (TGFs) and bone morphogenetic proteins (BMPs) – are released upon bone matrix degradation and trigger the recruitment of osteoblasts into the area of bone resorption. Active osteoblasts will produce an unmineralized bone matrix (osteoid) and achieve bone remodeling through the mineralization of the osteoid (reviewed in Rucci, 2008). In order to maintain bone homeostasis, the remodeling process has to be spatially and temporally controlled and members of the TGFß protein superfamily are central to this process (Crane and Cao, 2014).

1.1.1 Bone morphogenetic proteins

Bone morphogenetic proteins constitute the largest family of the transforming growth factor β (TGF-β) superfamily and are one of the main classes of multi-faceted secreted factors that drive vertebrate development (Schmierer and Hill, 2007). BMPs were first identified due to their capacity to induce bone formation (Urist, 1965) but they are now known to be involved in several other non-osteogenic processes, including, cell growth, differentiation, matrix production and embryonic development (Gordon and Blobe, 2008); due to their wide range of functions some authors proposed that they should be renamed as Body instead of Bone morphogenetic proteins (Horbelt et al., 2012).

BMPs are synthetized as large precursors composed of a signal peptide, which is responsible for protein secretion, a propeptide involved in protein dimer formation and stabilization and a mature peptide, which is the active form of the protein (Shimasaki et al., 2004); after removal of the signaling peptide, pro and mature peptides undergo dimerization and are sequentially processed into the dimeric mature biological active form by members of the subtilisin-like pro-protein convertases (SPCs) family, that recognize the optimal RXR/KR or the minimal RXXR sequences (Constam and Robertson, 1999; de Caestecker, 2004). The mature and biological active form of the protein is then secreted into the extracellular compartment. BMP secretion as a prodomain-mature complex has also been reported and may function as an additional regulatory mechanism, by inhibiting the binding of the protein to its receptor (Brown et al., 2005; Israel et al., 1992; Liao et al., 2003; Sengle et al., 2008; Wang et al., 1990).

Members of the BMP subfamily can be distinguished from other members of the TGFβ superfamily by the presence, in most cases, of seven cysteine residues instead of the nine. Six

of these cysteines are involved in the formation of three disulfide bridges which are responsible for the folding of the molecule into a unique three-dimensional structure called a cystine knot, while the remaining cysteine is involved in the formation of a disulfide bridge connecting BMP monomers (Shimasaki et al., 2004; Fig. 1.3).

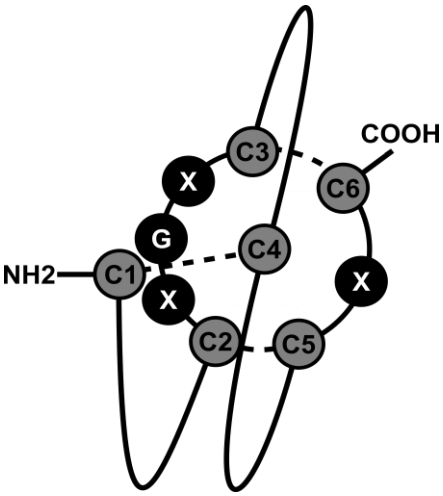


Figure 1.3. Schematic representation of the cysteine knot structure. The six cysteines involved in the knot are labeled from C1 to C6. Cysteines C2 and C3 form disulfide bridges (dashed line) with cysteines C5 and C6, respectively, thus forming a ring. The ring in penetrated by the third disulfide bridge formed between cysteines C1 and C4. Amino acids chains between C1 and C2 and between C4 and C5 form finger-like projections, while the amino acids between C3 and C4 form a helical structure designated by heel. The additional cysteine, involved in covalent dimer formation is localized in front of C4 (not represented). G represents a glycine residue and X represents any residue that is not glycine or cysteine. Adapted from Vitt et al., 2001.

In vitro studies and co-purification of BMPs from tissue extracts evidenced the formation of heterodimers of BMPs, which, in some cases, generated a stronger signal than the one observed from respective homodimers (Aono et al., 1995; Katagiri et al., 2013; Suzuki et al., 1997; Wozney et al., 1988). Signaling by BMPs is mediated through the binding of homo/hetero-dimers to BMP receptors (BMPR type I and II) present on the cell surface and through the intracellular cascade of events responsible for signal transduction (Marcellini et al., 2012). This process has been remarkably conserved from invertebrates to mammals and all the steps of the pathway are tightly regulated at different levels (Gordon and Blobe, 2008). Alterations to normal signaling function, including either germ-line or somatic mutations or changes on gene expression, are often related to developmental disorders, vascular diseases and cancer (Gordon and Blobe, 2008).

Until now, four type I receptors – BMPR1A, BMPR1B, activin receptor-like kinase (ALK) 1 and 2 – and three type II receptors – BMPR2 and type IIA and IIB activin receptors (ActR2A and Actr2B) – have been described to bind BMPs (Tian and Liu, 2013). They have a very flexible oligomerization pattern and BMPs have at least two different options to initiate signal transduction: (1) bind to preformed BMPR1-BMPR2 complexes, inducing conformational changes that will activate the complex and initiate the signal, or (2) bind first to the high affinity receptor (BMPR1) that will subsequently recruit the low affinity receptor

(BMPR2) into the ligand-mediated signal complex (Nohe et al., 2002). Depending on the type of receptor oligomerization and on the different combination of ligands and receptors, different cellular processes and pathways can be activated, allowing a remarkable diversity of BMP signaling (Derynck and Zhang, 2003; Nohe et al., 2004). Furthermore and to increase complexity, BMP signaling can be fine-tuned at different levels: extracellularly by agonists and antagonists that can modify ligand activity, and intracellularly by intermediates that can mediate the crosstalk between different signaling pathways and enhance or inhibit downstream signaling events (Morikawa et al., 2013).

The canonical BMP signaling pathway is activated upon binding of BMP dimers to the high affinity type I receptors, then low affinity type II receptors are recruited and activate type I receptors by phosphorylating specific serine and threonine residues (Fig. 1.4). Activated type I receptors elicit the phosphorylation of intracellular receptor-regulated SMADs (R-SMAD/SMAD1, 5 and 8; SMAD is the contraction of SMA (small body size gene) and MAD (mothers against decapentaplegic gene), homologs of vertebrate SMAD proteins in *Caenorhabditis elegans* and *Drosophila melanogaster*, respectively) that will form a trimeric complex with the common-mediator SMAD (Co-SMAD/SMAD4). The complex will be subsequently translocated into the nucleus where, after interaction with transcription factors, it will regulate the transcription of target genes (e.g. inhibitor of DNA binding 1 (ID1), RUNX2, jun B proto-oncogene (JUNB); Locklin et al., 2001; Marcellini et al., 2012). Inhibitory SMADs (I-SMADs) are responsible for inhibiting BMP signaling by competing with R-SMADs for type I receptors, blocking their recruitment and subsequent phosphorylation, forming a feedback loop mechanism (Imamura et al., 2013).

While the SMAD-dependent pathway is the preferred pathway of BMP signaling, other pathways such as TAK1/p38 mitogen-activated protein kinase (MAPK), Erk MAPK, c-Jun N-terminal kinase (JNK) and the LIM kinase 1 pathways, have been associated with BMP signal transduction (Ramel and Hill, 2012; Nohe et al., 2002). Although mechanisms of activation of these pathways by BMPs are poorly understood, possible crosstalk with the SMAD-dependent pathway have been proposed (Broege et al., 2013; G Chen et al., 2012). Crosstalk between intermediates of SMAD-dependent and -independent pathways can trigger synergistic or antagonizing effects on BMP signaling (Massagué, 1998; Derynck and Zhang, 2003).

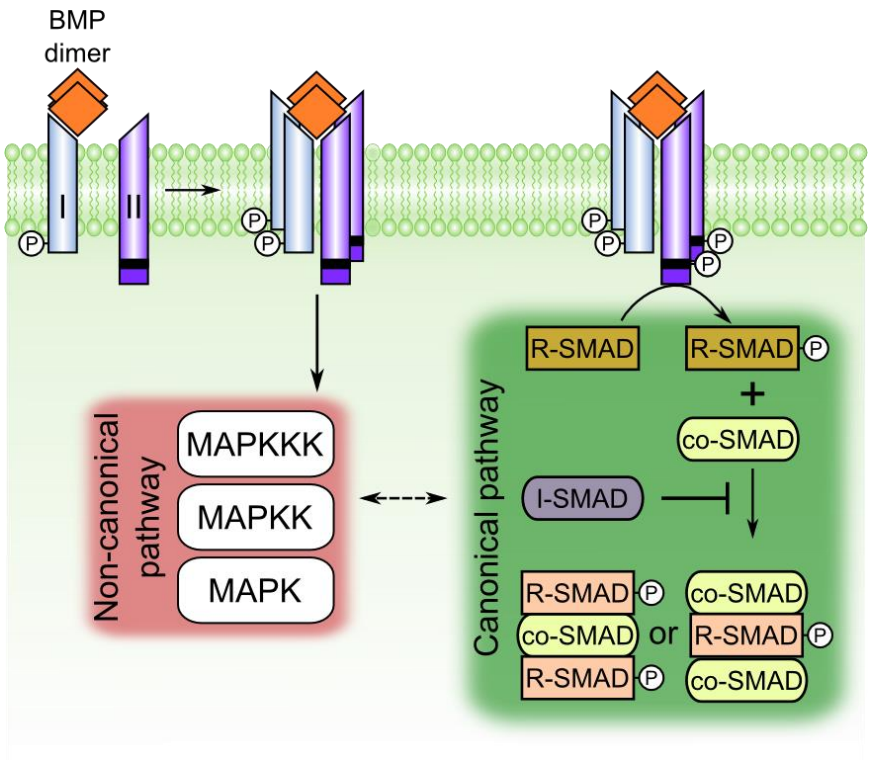


Figure 1.4. Schematic representation of SMAD-dependent and SMAD-independent BMP signaling pathways. Depending on receptor oligomerization, BMPs can activate different signaling pathways. The binding of BMP dimers to preformed receptor complexes, composed of type I and type II BMP receptors (BMPR1 and BMPR2, respectively), leads to the activation of the SMAD-dependent pathway. Activated BMP receptors will recruit and phosphorylate R-SMADs that will associate with co-SMAD and enter the nucleus, where transcriptional regulation of target genes will occur. BMP-induced oligomerization of the receptors – i.e. BMP dimers bind first to the high affinity receptor (BMPR1) that will subsequently recruit the low affinity receptor (BMPR2) – leads to the activation of the non-canonical BMP pathways, e.g. JNK, ERK and p38 MAPK pathways. Transduction of BMP signal can be blocked either by extracellular antagonists (e.g. noggin and chordin or by intracellular proteins such as I-SMADs, which prevent the association between R-SMADs and co-SMADs. I and II indicate type I and type II receptors, respectively. Circled P indicates phosphorylation. Adapted from Bessa et al., 2008; Demers et al., 1999; Liu et al., 2005 and Nohe et al., 2002)

Members of the BMP family have been classified into four subfamilies according to their primary structure and function: BMP2/4, osteogenic proteins (OPs), cartilage-derived morphogenetic proteins (CDMPs) and growth and differentiation factors (GDFs) subfamilies (Kawabata et al., 1998; Li et al., 2003). BMP2/4 subfamily was the first to be identified and characterized; it was initially formed by 2 members, BMP2 and BMP4, which are homologs of the *Drosophila* decapentaplegic (DPP) protein (Padgett et al., 1987; Wozney et al., 1988). BMP16 has recently been identified as a member of this subfamily (Feiner et al., 2009) and we will refer to it as BMP2/4/16 subfamily thereafter. While BMP2 and BMP4 have been extensively studied and shown to be osteo-inductive, critical to early development (deficiency resulted in early lethality and embryos abnormal for gastrulation and mesoderm formation;

Winnier, Blessing, Labosky, and Hogan, 1995) and to be involved in the development of several organ systems, including lungs, limbs, intestine and kidney (Weaver et al., 1999; Selever et al., 2004; Chalazonitis and Kessler, 2012; Takigawa et al., 2010; Nishinakamura and Sakaguchi, 2014), not much is known about BMP16, which was until recently described as a teleost fish specific protein (Feiner et al., 2009). In zebrafish, BMP16 expression was associated to the developing heart, gut epithelium and swim bladder (Feiner et al., 2009).

1.1.1.1 Bone morphogenetic protein 2

The osteogenic capacity of BMP2 were first described by Marshall Urist, in 1965, upon the discovery that demineralized segments of bone had the capacity to induce new bone formation when implanted in rabbit muscle (Urist, 1965). However, it was only in the late 80's that Wozney and co-workers were able to purify it from bovine bone extracts (Wozney et al., 1988). BMP2 was co-purified together with BMP1, BMP3 and BMP4 (that at the time was named BMP2B, due to the high degree of similarity – over 90% – between BMP2 and BMP4 mature peptides) and was shown to induce ectopic bone formation when implanted ectopically into rats (Wang et al., 1990; Wozney et al., 1988). Nowadays, BMP2 is considered to be one of the most potent bone inducers, but despite its role in cartilage and bone metabolism, it has several other non-osteogenic functions in different biological processes (e.g. mesoderm formation and patterning, neural and limb patterning; reviewed by Hogan, 1996).

1.1.1.1.1 Gene and protein structures

Human BMP2 gene is localized on chromosome 20 (Rao et al., 1992) and is composed by 2 coding exons and a 5' non-coding exon (Sharapova et al., 2010; Fig. 1.5A). As for other BMPs, BMP2 is synthetized as a large precursor containing (1) a signal peptide, responsible for directing the protein to the secretory pathway, (2) a pro-peptide, poorly conserved throughout evolution and involved in dimer formation and stabilization that will be cleaved after processing, and (3) a mature peptide which is very well conserved throughout evolution and involved in osteogenic signaling (Sharapova et al., 2010). The folded monomers of BMP2 possess three disulfide bridges, which form the cystine knot, a structure that confers some rigidity to the monomer (Vallejo and Rinas, 2013; Fig 1.3). The monomer fold has been described as a hand, with the fingers being represented by the extended ß-sheets. Thus, the

convex surface of the fingers corresponds to the "knuckles" while the helix region corresponds to the "wrist" (Kirsch et al., 2000; Fig. 1.5B) Monomers of BMP2 are covalently connected by an additional disulfide bridge (Vallejo et al., 2013).

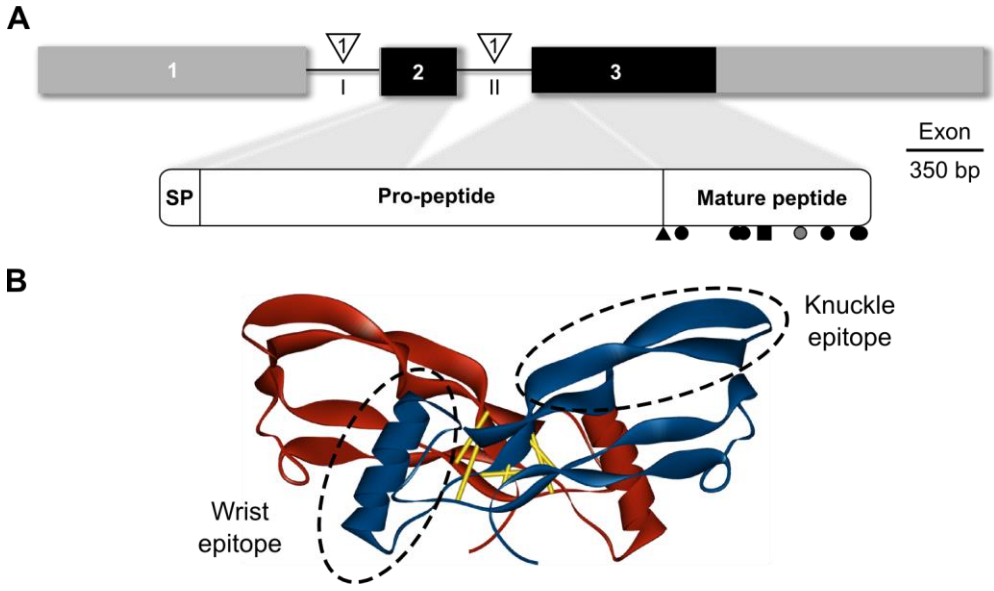


Figure 1.5. (A) Schematic representation of human BMP2 gene and protein structures. In the gene structure, exons are displayed as gray boxes (non-coding exons) and black boxes (coding exons) and indicated with arabic numbers. Introns are displayed as solid lines and indicated with roman numbers. Phase of intron insertion is indicated in white triangles. In the protein structure, SP indicates the signal peptide. Black triangle indicates cleavage site by subtilisin-like pro-protein convertase (SPC). Black circles indicate cysteine residues involved in the cystine knot structure. Gray circle indicates cysteine residue involved in the formation of the disulfide bridge that connects BMP monomers. Black square indicates site of N-glycosylation. Correspondence between coding regions in gene and related domains in protein is indicated with shadowed regions. Adapted from (Rafael et al., 2006) **(B)** Tri-dimensional structure of a human BMP2 dimer. Each BMP2 monomer is indicated with a different color (red or blue). Wrist and knuckle epitopes are indicated with a dashed line. Adapted from Vallejo and Rinas, 2013.

Human BMP2 has three N-glycosylated sites confirmed by enzymatic assays; one of those sites is located in the mature peptide, on asparagine (Asn) at position 338 (Hang et al., 2014). N-glycosylation (addition of a glycan to specific Asn residues) is an important feature for protein structure and function (it increases protein stability and half-life; Carreira et al., 2014); nonetheless a functional (although to a much lower extent) non-glycosylated form has been reported (Paulo C Bessa et al., 2008; Ruppert et al., 1996). Western blot analysis of human recombinant BMP2 protein produced by CHO cells revealed the presence of several

forms with different molecular weights in agreement with the predicted weight of the different peptides: BMP2 precursor (including the pro and the mature peptide; 60 kDa), propeptide (40-45 kDa) and mature peptide (18-22 kDa) protein. A precursor of about 45 kDa was detected when cells were pre-treated with an inhibitor of N-linked glycosylation, which is in accordance with the predicted molecular weight based on amino acids composition (Israel et al., 1992).

1.1.1.1.2 BMP2 expression patterns and physiological role

Although its name suggests a role dedicated to bone formation, BMP2 gene is not exclusively expressed in skeletal tissue and has been detected in a wide variety of vertebrate tissues (reviewed by Hogan, 1996). BMP2 expression is first detected at the onset of gastrulation and ventral folding (Madabhushi and Lacy, 2011), supporting a critical role of BMP2 during early development. It is considered a late mesodermal marker and a cardiogenic factor and, in early stages of development, BMP2 expression can be normally detected in lateral plate mesoderm and cardiac-associated pharyngeal endoderm (Ghatpande et al., 2006; Doss et al., 2012). BMP2 gene is described to be ubiquitously expressed and its involvement in several embryonic processes and organogenesis (e.g. gastrulation, neural patterning, gut development, cell proliferation, apoptosis and differentiation; reviewed by Hogan, 1996) evidence its pleiotropic activity (Tseng and He, 2007; Reddi, 1997). The following sections intend to give a brief description of the principal functions associated to BMP2.

1.1.1.1.2.1 Embryonic development

The role of BMP2 in embryonic development was first evidenced from studies in the fruit fly *Drosophila melanogaster*, which demonstrated the role of DPP (the invertebrate homolog of BMP2 and BMP4) in body patterning at several developmental stages (Ferguson and Anderson, 1992; Capovilla et al., 1994). Interestingly, *Drosophila* DPP can induce endochondral bone formation when introduced subcutaneously in mouse, and human BMP4 is able to rescue the dorsoventral defects resulting from the lack of DPP in *Drosophila* (Padgett et al., 1993; Sampath et al., 1993). Because invertebrate and vertebrate homologs are interchangeable, it was proposed that BMP2 may be involved in embryonic development, an hypothesis supported by the presence of BMP2 transcript in early stages of embryo development, before the onset of chondrogenesis and ossification processes (Zhang and Bradley, 1996). The role of BMP2 in vertebrate development is, nowadays, well documented,

and it has been associated with several processes, such as mesodermal patterning, embryonic cardiac development, vasculogenesis regulation, left/right axis patterning, neuronal differentiation, chondrogenic and osteogenic differentiation, limb development and negative regulation of the lymphatic lineage (Bandyopadhyay et al., 2006; Doss et al., 2007; Dunworth et al., 2014; Goldstein et al., 2005; Majumdar et al., 2001; Pal et al., 2006; Martin Raida et al., 2005; Wu et al., 2009).

1.1.1.1.2.2 Cardiac development

Heart is the first organ to form during vertebrate development and knockout (KO) studies in mice (i.e. animals lacking the expression of BMP2) have clearly evidenced the central role of BMP2 in cardiogenesis (Zhang and Bradley, 1996). Embryos of $Bmp2^{-/-}$ mice were non-viable and died between days 7 and 9 after fertilization, possibly from defects in the closure of the pro-amniotic canal and abnormalities in cardiac development (Zhang and Bradley, 1996). BMP2 expression was observed in the promyocardium surrounding mesodermal cells and in the atrioventricular canal of the embryonic heart and was associated with the formation of the heart chambers (Lyons et al., 1990; Zhang and Bradley, 1996).

Ablation of ß-catenin expression in mouse visceral endoderm triggered the formation of multiple hearts at sites of ectopic BMP2 expression, evidencing the important role of BMP2 for the onset of heart formation (Lickert et al., 2002). Expression of BMP2 gene in mesenchymal cells also suggests a role in the mesenchyme development and in the differentiation of mesenchymal cells into myocardial tissue (Abdelwahid et al., 2014). BMP signaling pathway and BMP2 have also been clearly associated with mechanisms of cardiogenesis through their regulatory action on cardiomyocytes marker gene expression (e.g. GATA4, 5 and 6, NKX2.5, MEF2C and TBX[1]; Wang et al., 2013; Andrée et al., 1998; Chen et al., 2012; de la Pompa and Epstein, 2012; Ghosh-Choudhury et al., 2003; Monzen et al., 1999; Pucéat, 2007; Schultheiss et al., 1997). Moreover, several functional assays also support the role of BMP2 in heart valve formation, cardiac contractility, stimulation of cardiomyocyte shortening, formation of heart jelly, induction of endocardial epithelial to mesenchymal transition (EMT) and coordination and development of atrioventricular region (Abdelwahid et al., 2014; Ma et al., 2005; Rivera-Feliciano and Tabin, 2006; Sugi et al., 2004; Wang et al., 2007; Ghosh-Choudhury et al., 2003).

[1] GATA (GATA-binding protein); NKX2.5 (NK2 homeobox 5); MEF2C (myocyte enhancer factor 2C); TBX (T-box transcription factor)

1.1.1.1.2.3 Bone and cartilage formation

Because of their osteogenic properties, BMP2, BMP4, BMP6, BMP7 and BMP9 have been considered as candidates for a role in the induction of bone formation (Wozney, 2002; Luu et al., 2007). Among those, BMP2 demonstrated a remarkable osteogenic capacity, evidenced by its ability to induce ectopic bone formation when implanted into rodents (Chen et al., 2004; Wang et al., 1990; Wozney et al., 1988). In combination with other factors, BMP2 delivers one of the most important signals that drive osteoblast maturation and differentiation, from the precursor mesenchymal cells, and the crosstalk between BMP and Wnt/β-catenin signaling pathways is thought to be crucial in this process (Friedman et al., 2009; Chen et al., 2014; Bais et al., 2009; Mbalaviele et al., 2005). Additionally, BMP2 has been associated in chick embryos with control of size and shape of developing bones (Karsenty and Wagner, 2002; Reddi, 1997).

The lethality of BMP2 inactivation prevented the study of its role during bone formation. To overcome this problem, the use conditional KO animals, where BMP2 expression is ablated in specific cell types using the Cre-Lox recombination system, has been considered (Mi et al., 2013). Bones from animals where osteoblastic expression of BMP2 was ablated were proved to be of poor quality, thinner, fragile and to have an increased risk of fracture along with a deficient healing capacity (Yang et al., 2013). The absence of BMP2 in osteoblasts mimics what often happens with aging: a reduced bone vascularization and a decrease in osteoblasts number and differentiation (Yang et al., 2013). BMP2 has been reported as an essential factor not only in post-natal bone formation and maintenance, but also in bone regeneration and fracture repair (Rosen, 2009; Alam et al., 2009; Tsuji et al., 2006; Bais et al., 2009). Mice lacking limb-specific expression of BMP2 exhibited bone micro-fractures and an increased risk of fracture (Rosen, 2009). Osteoprogenitor cells of mice lacking BMP2 expression exhibited a defective differentiation and proliferative capacity as well as an impaired ability to fully regenerate bone tissue (Granero-Moltó et al., 2009).

Besides its role in osteoblast differentiation and bone formation, BMP2 has been associated with chondrocyte maturation and endochondral bone formation, during early stages of mammalian skeleton development (Shu et al., 2011; Karsenty and Wagner, 2002; Bandyopadhyay et al., 2006). Conditional ablation of BMP2 expression in chondrocytes led to the development of a severe chondrodysplasia phenotype in mice, resulting from defective mechanisms of chondrocyte proliferation, differentiation and apoptosis and a disorganized growth plate region (Shu et al., 2011). BMP2 expression by chondrocytes is also central to the

promotion of cartilage callus maturation in healing bone fractures and animals lacking BMP2 expression in chondrocytes leads to a prolonged cartilage phase (Mi et al., 2013). The role of BMP2 in bone repair was further confirmed by high levels of BMP2 transcript in periosteal cells and hypertrophic chondrocytes during fracture healing, suggesting that the first mesenchymal cells expressing BMP2 will be the cells committed to the osteoblastic and chondrocytic lineages (Matsubara et al., 2012; Yu et al., 2010).

1.1.1.1.2.4 Vascularization

Several reports support the central role of BMPs in the regulation of mechanisms at the origin of vascularization during development and bone formation (Hogan, 1996), through the direct activation of BMP signaling pathway or by interacting with other pathways (e.g. VEGF signaling pathway) also involved in the process of vascularization (Moser and Patterson, 2005; He and Chen, 2005; Deckers et al., 2002). BMP2 was shown to accelerate bone healing process by supporting vascularized bone regeneration and also to regulate the expression of angiogenic growth factors, such as placental growth factor (PlGF) and members of the family of basic-helix-loop-helix transcription factors, known to play a vital role in angiogenesis (Hollnagel, 1999; He et al., 2013; Raida et al., 2005). Besides being involved in physiological vascularization, BMP2 function has been associated with pathological conditions, such as tumor angiogenesis in several types of cancer (Raida et al., 2006; Langenfeld and Langenfeld, 2004; Raida et al., 2005). Under experimental conditions, BMP2 could increase the size and the number of tumor blood vessels in mice, an effect that could be reversed in the presence of specific BMP2 inhibitors (Langenfeld and Langenfeld, 2004).

1.1.1.1.2.5 Odontogenesis

BMPs are also central to odontogenesis (Feng et al., 2011; Mu et al., 2012; L-A Wu et al., 2010) and BMP2 is a decisive contributor to tooth formation during early mouse development. During this developmental time window, BMP2 gene expression is localized in the dental epithelium while it is transferred at later stages to the mesenchymal dental papilla (Chen et al., 2008). After birth, BMP2 transcript is detected in odontoblasts and ameloblasts involved in tooth cytodifferentiation (Yang et al., 2012). Besides a role in odontoblast differentiation (Yang et al., 2012), BMP2 has proven to be an important factor in the vascularization of dental pulp, in the differentiation of ameloblast and dental pulp cell and in the formation and mineralization of dentin and enamel (Miyoshi et al., 2008; Feng et al.,

2011; Chen et al., 2008; Wu et al., 2010). Furthermore, when removed from osteoblast precursors, BMP2 was shown to be involved in the development of tooth roots as well as in their supporting structures, again due to a decrease on vascularization (Rakian et al., 2013). A role of BMP2 in root development is also supported by the reduction of the amount and quality of the dentin produced upon the inactivation of BMP2 gene expression in mice early odontoblasts (Feng et al., 2011). Delayed amelogenesis and disorganized enamel were also observed in these animals, although no significant difference was reported in the overall quantity of enamel formed (Feng et al., 2011; Yang et al., 2012). The transcriptional up-regulation of transcription factors involved in the process of tooth formation (NFY, CEBP, DLX2 and 3, MSX1 and LEF-1)[2] and of marker genes involved in odontoblast and ameloblast differentiation (DSPP, p75NGFR and AMELX)[3] by BMP2 through the SMAD-dependent BMP signaling pathway further demonstrate a role of BMP2 during odontogenesis (Dassule and McMahon, 1998; Fan et al., 2009; Harris et al., 2003; Hassan et al., 2004; Hoeppner et al., 2009; Chen et al., 2008; Miyoshi et al., 2008).

1.1.1.1.2.6 Other functions

During the development of the nervous system of several vertebrates, BMP2 has been associated with glial and central nervous system fates and retinal patterning. BMP2 was shown to promote survival and differentiation of neurons, neural crest cell induction and migration, neural tube patterning and brain regionalization, but also to play an important role in practically every stages of gut and enteric nervous system (ENS) formation (Chalazonitis and Kessler, 2012; LaBonne and Bronner-Fraser, 1999; Sela-donenfeld and Kalcheim, 1999; Sato et al., 2010; Sailer et al., 2005; Sakuta et al., 2006). Roles in cell proliferation and apoptosis have also been described, and depending on the cellular context and/or morphogen concentration, BMP2 can either promote or inhibit cell differentiation and apoptosis (Kim et al., 2013). For instance, transduction of BMP2 signal through the TAK1-p38 kinase pathway had pro-apoptotic effects in mouse MH60 cells (Kimura et al., 2000), while it had anti-apoptotic effects in the mouse chondrocytic N1511 cell line via Akt-mediated NF-kB activation (Sugimori et al., 2005).

[2] NFY (heterotrimeric transcription factor Y); CEBP (CCAAT enhancer-binding protein); DLX (distaless); MSX (meshless); LEFT1 (lymphoid enhancer binding factor)
[3] DSPP (dentin sialophosphoprotein); p75NGFR (outer enamel epithelium p75); AMELX (amelogenin)

1.1.1.1.3 BMP2 and human diseases

As described above, BMP2 is a multifunctional growth factor that plays important roles in many biological processes and whose abnormal expression or signaling disruption is linked to several pathological processes (Liu et al., 2008). Osteoporosis, a disease characterized by a decrease in bone mineral density and an increased fracture risk, was linked to chromosome 20p12.3, the same region containing BMP2 locus (Styrkarsdottir et al., 2003). Variants of BMP2 resulting in reduced activity or gene expression, have been associated to osteoporotic phenotypes (Fritz et al., 2006) and several polymorphisms in BMP2 gene have been identified and related to the disease (Li et al., 2011; Styrkarsdottir et al., 2003; Wang et al., 2012). Osteoarthritic tissue has also a reduced BMP2 gene expression (Nakase et al., 2003) and single nucleotide polymorphisms (SNPs) in BMP2 have been linked to an increased susceptibility to osteoarthritis (Valdes et al., 2004). Brachydactyly, a condition characterized by a shortening of the fingers and toes, was associated to a duplication in a conserved sequence, downstream BMP2 gene, which normally functions as a regulatory element controlling BMP2 expression in the limbs (Dathe et al., 2009).

Besides bone and cartilage metabolism, BMP2 has also been associated with other pathologies. For example, BMP2 expression was shown to be significantly increased in colon tissues from patients with Hirschsprung disease and has been associated, along with other two members of the BMP family (BMP5 and BMP10) to the development of the disease (Wu et al., 2014). Comparative genomic hybridization studies have also linked the genomic region containing BMP2 locus to the predisposition to develop pre-excitation syndrome (Lalani et al., 2009), a condition where heart ventricles contract prematurely. High levels of BMP2 transcript and protein were observed in hypercalciuric rat kidney tissue, possibly suggesting a role of BMP2 in renal stone formation in idiopathic hypercalciuria (Jia et al., 2014).

Dysregulations of BMP2 signal have also often been associated to cancer, and BMP2 gene expression was found to be altered in many types of tumors (e.g. breast cancer, pancreatic cancer, osteosarcoma and prostate cancer; Horvath et al., 2004; Kleeff et al., 1999; Sadikovic et al., 2010; Yamamoto et al., 2002). Nevertheless, mechanisms underlying BMP2 role in tumorogenesis are not completely understood and there is some controversy regarding its protective or stimulating effect on tumor development (Langenfeld et al., 2006; Wang et al., 2012). On one hand, BMP2 has shown anti-proliferative effects in breast, lung, colon adenocarcinoma, gastric, osteosarcoma and non-small cell lung cancer cell lines (Wang et al., 2011; Wen et al., 2004; Beck et al., 2006; Pouliot and Labrie, 2002; Soda et al., 1998), and in

prostate cancer the loss of BMP2 signaling was associated to cancer progression and aggressiveness (Dumont and Arteaga, 2003). In the other hand, a role in cell invasion and cancer progression has also been proposed, and a BMP2-dependent increase of blood vessels formation was observed in tumors formed by A549 cells in nude mice (Langenfeld and Langenfeld, 2004). Migration and an increased invasiveness capacity were also associated with BMP2 expression in several cancers, including chondrosarcomas, breast cancer, pancreatic cancer, lung cancer, gastric cancer and melanomas (Kang et al., 2010; Hsu et al., 2011; Rothhammer et al., 2005; Kleeff et al., 1999; Jin et al., 2012; Fong et al., 2008).

1.1.1.1.4 Clinical use of BMP2

Due to its remarkable osteogenicity, BMP2 was soon considered a potential therapeutic agent to reverse or alleviate bone disorders. Its capacity of interacting with VEGF signaling pathway also provide BMP2 with the striking capacity of promoting angiogenesis, a critical step on the success of bone regeneration process (Deckers et al., 2002). After several successful pre-clinical tests realized in several mammalian species (Bax et al., 1999; Li and Wozney, 2001; Brown et al., 2005; Zhao et al., 2002), BMP2 was tested in humans and shown to positively affect fracture repair (Valentin-Opran et al., 2002; Li and Wozney, 2001). Since 2002, human recombinant BMP2 (hr-BMP2) is commercially available for clinical use as a treatment to replace invasive autogenous bone graft surgery (Hoffmann et al., 2013). The use of rh-BMP2 is nowadays a common practice during orthopedic surgery and is increasingly applied to treat spinal fusion, open tibia fracture, non-union, sinus lift and alveolar ridge augmentation (Abd-El-Barr et al., 2011; Govender et al., 2002; Tressler et al., 2011; Jung et al., 2003). Off-label trials have also successfully used recombinant BMP2 in the reconstruction of mandibular segments and teeth (Carter et al., 2008; Cochran and Wozney, 1999).

Although the clinical use of rh-BMP2 is very well tolerated (Szpalski and Gunzburg, 2005), several secondary effects have been reported. Heteropic ossification, inflammation, excessive bone resorption and nerve compression were observed following the local administration of rh-BMP2 to improve bone healing at fracture site (Chen et al., 2010; Shah et al., 2010; Woo, 2013). Recently, higher cancer risk in patients treated with rh-BMP2 has also been reported (Carragee et al., 2013). One of the aspects that may influence the incidence of these secondary effects may be related to the need of using very high therapeutic doses, i.e.

almost 100 times higher than physiological doses, due to the reduced half-life of BMP2 (Kim et al., 2013).

1.1.1.1.5 Transcriptional regulation of BMP2 gene

The pleiotropic role of BMP2 in animal physiology implies a tight regulation of the gene expression, both spatially and temporally (Jiang et al., 2010). Promoter regions of mouse and human genes (among others) have been analyzed and, although it remains controversial, two transcription start sites have been identified in both species (Abrams et al., 2004; Ghosh-Choudhury et al., 2001; Helvering et al., 2000; Sugiura, 1999). Binding sites for bone and cartilage related transcription factors, including RUNX2, and SOX9[4], have also been identified in human gene promoter (Helvering et al., 2000). Several other common *cis*-regulatory elements, such as SP1, CREB and RARE[5] were found in the BMP2 promoter regions of human and mouse (Helvering et al., 2000; Zhang et al., 2011; Goto et al., 2006). The presence of long-range regulatory elements, i.e. outside of the promoter region, have been evidenced and shown to control mouse BMP2 gene transcription (Chandler et al., 2007). This, together with a capacity of auto regulation, that functions as a negative feed-back loop decreasing BMP2 levels, and the presence of a complex promoter organization, with suppressing elements located contiguously to enhancer elements, are some of the evidences for the complex transcriptional regulation to which BMP2 gene is subjected (Sugiura, 1999; Ghosh-Choudhury et al., 2001).

1.1.1.1.6 Post-transcriptional regulation of BMP2 transcript

BMP2 gene expression is also regulated at the post-transcript level, i.e. after transcription of the gene (Jiang et al., 2008). Three prime untranslated regions (3'UTRs) of transcripts are known to often contain elements important for the post-transcriptional regulation of gene expression (Mignone et al., 2002). Interestingly, 3'UTR of BMP2 transcript has been highly conserved throughout evolution and AU-rich sequences commonly associated with the mechanisms underlying post-transcriptional regulation are similarly positioned in the 3'UTR of *Drosophila* DPP and human BMP2 transcript, suggesting the conservation of an ancient regulatory system (Fritz et al., 2004). A single nucleotide

[4] SOX9 (sex determining region Y box 9)
[5] CREB (cAMP responsive element binding protein); RARE (retinoic acid responsive element)

polymorphism (SNP rs235764) has been identified in the AU-rich element of the human gene leading to alterations in RNA stability, decay and consequently RNA levels (Fritz et al., 2006). This ultra-conserved region, function as a control mechanism, regulating BMP2 levels according to tissue and cell types specificities, normally acting as a post-transcriptional repressor. Disruptions on this post-translation mechanism are often associated to pathological calcifications (Kruithof et al., 2011). Similarly, multiple polyadenylation signals, which are commonly associated with the modulation of RNA abundance, were evidenced in the 3'UTR of human BMP2 transcript suggesting another way of post-transcriptional regulation of BMP2 gene expression (Fritz et al., 2004).

When talking about post-transcriptional regulation it is common to refer to microRNAs (miRNAs), which due to their ability to regulate mRNA translation, are potent regulators of protein availability (Shivdasani, 2006). Although the role of miRNAs in bone formation, mineralization and homeostasis is still largely understudied, there are some evidences of miRNAs involved in the regulation of bone metabolism. Recently miR-20a was shown to promote osteogenic differentiation of human mesenchymal stem cells by targeting PPARγ, BAMBI and CRIM1[6], negative regulators of BMP signaling pathway (Zhang et al., 2011). Overexpression of miR27a, was also described to decrease protein levels of BMP2, while overexpression of miR378 increases the levels of BMP2 available for osteoinduction, showing that the modulation of BMP2 protein levels can be dependent on the cellular context and intermediates present (Gong et al., 2014; Hupkes et al., 2014). Thus, understanding the mechanisms through which post-transcriptional regulation of BMP2 is controlled will greatly help to uncover how repressor and/or activators are able to modulate BMP2 synthesis in different cell types.

1.1.1.1.7 BMP2 signaling modulation

Due to its involvement in diverse activities from embryonic patterning to skeleton formation and homeostasis, BMP2 activity has to be controlled at various levels. Besides being controlled at gene (e.g. transcriptional and post-transcriptional regulation) and protein levels (proteolytic cleavage, post-translational modification, homo/hetero dimerization and differential receptor binding specificity), BMP2 signal is also modulated intra and extracellularly through the binding of specific molecules (Yanagita, 2005). The different

[6] PPAR (peroxisome proliferator-activated receptor); BAMBI (BMP and activin membrane-bound inhibitor); CRIM1cysteine rich transmembrane BMP regulator 1

levels of regulation include: (1) extracellular binding of proteins to BMP2, preventing the interaction of BMP2 dimers with specific receptors; (2) docking of BMP2 to co-receptors and dominant negative pseudoreceptors competing with the regular BMP receptors; (3) intracellular binding of I-SMADs to R-SMADs, preventing BMP2 signal transduction; (Gazzerro and Canalis, 2006; Fig 1.6).

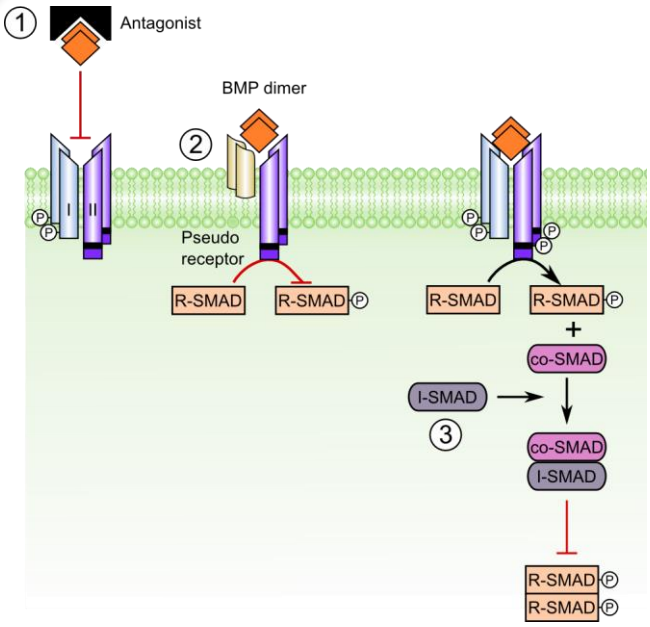


Figure 1.6 Schematic representation of the different levels of regulation of the BMP signaling. (1) Extracellular antagonists bind to BMP dimers and prevent the interaction with type I and type II receptors; (2) Pseudoreceptors modulate BMP signaling at membrane level, preventing the signal to be transduced; (3) Inhibitory SMADs regulate intracellular signaling by binding to co-SMAD and avoiding complexation with R-SMAD and the consequent regulation of target genes. I and II indicate type I and II receptors, respectively. Adapted from Balemans and Van Hul, 2002.

1.1.1.1.7.1 Extracellular modulation of BMP2 activity

Several proteins are known to bind extracellularly to BMP2 and to reduce its activity or prevent the transduction of the signal. These molecules are commonly known as BMP2 antagonists (Yanagita, 2005), possess a structure analogous to the one of BMP2 and have been classified into several families, based on the number of cysteines involved in the formation of the cystine knot: Dan family (eight-membered ring), twisted gastrulation (nine-membered ring), Noggin and chordin (ten-membered ring; Avsian-Kretchmer and Hsueh, 2004; Vukicevic and Sampath, 2008).

A list presenting some of BMP2 interacting partners is presented in Table 1. It is worth to note that the inhibitory effect is, in some cases, not restricted to BMP2 but also directed to other BMPs or TGF-β family members, and may have different functions unrelated to the antagonistic effect exerted in the BMP signaling (Walsh et al., 2010).

Table 1. BMP2 interacting partners

Family name	Protein name	Function	References
Dan	Differential screening-selected gene aberrative in neuroblastoma	Tumor suppressor	(Ozaki and Sakiyama, 1994; Hung et al., 2012)
	Cerberus	Head organizer	(Silva et al., 2003; Piccolo et al., 1999)
	Caronte	Left-right asymmetry	(Schlange et al., 2002)
	Protein related to Dan and Cerberus	BMP signaling regulation in ovary, brain and bone	(Ideno et al., 2009; Sudo et al., 2004)
	Gremlin	Apical ectodermal ridge and epithelial-mesenchymal feedback signaling; early limb outgrowth; kidney patterning; kidney and lung morphogenesis	(Merino et al., 1999; Michos et al., 2004)
	Uterine sensitization associated gene-1	Hair follicles induction; kidney patterning; tooth development	(Yanagita, 2005)
Twisted gastrulation	Twisted Gastrulation	Dorso-ventral axis formation; brain development; BMP2 agonist (depending on the concentration)	(Sun et al., 2010; Gazzerro et al., 2005; Larraín et al., 2001)
Noggin and Chordin	Noggin	Neural tissue, bone and joint formation;	(Brunet et al., 1998; McMahon et al., 1998)
	Chordin	Morphogenesis; body axis and neural tissue formation	(Londin et al., 2005; Piccolo et al., 1996)
	Chordin-like 1	Dorso-ventral and anterio-posterior patterning; retina development	(Nakayama et al., 2004; Branam et al., 2010)
	Chordin-like 2	Joint specification; osteoarthritic cartilage regeneration	(Nakayama et al., 2004)
	Brorin	Neural development	(Koike et al., 2007)
Other	Follistatin	Neural induction	(Zhang et al., 1997)
	Follistatin-related gene	Activins modulator in hematopoiesis	(Maguer-Satta and Rimokh, 2004)
	Crossveinless	Wing vein development in *Drosophila*; Agonist/antagonist of BMP2 signaling (depending on proteolytic cleavage)	(Rentzsch et al., 2006)
	Matrix Gla protein	Calcification inhibitor; BMP2 modulator	(Boström et al., 2001)

The large majority of the proteins interacting with BMP2 are negative regulators of its activity. Few of them are, however, capable of enhancing the biological effect of BMP2, and those are known as BMP2 agonists. Depending on their concentration and on the presence of other modulators, some proteins can be either antagonists or agonists (e.g. Twisted

Gastrulation, crossveinless/BMPER members, that increase or decrease signaling in a context dependent manner (Umulis et al., 2009; Zhang et al., 2010; Rentzsch et al., 2006).

1.1.1.1.7.1.1 Matrix Gla protein as an extracellular regulator of BMP2 activity

Matrix Gla protein (MGP) is a small, γ-carboxylated, vitamin K-dependent protein (Price and Williamson, 1985) that was initially found to be tightly associated with BMP2 (so tightly that the use of strong denaturing agents was needed to separate both proteins; Urist et al., 1984). MGP has been described as a calcification inhibitor and some groups speculated that it may be through its direct interaction with BMP2 and the subsequent attenuation of BMP2 signaling (Boström et al., 2001; Xue et al., 2006; Sweatt et al., 2003). The interaction between the two proteins was shown through the use of co-immunoprecipitation studies, however residues involved in the binding were not identified. Several pieces of evidence have been collected in the last decade toward Gla residues of MGP being important for the interaction with BMP2, although conclusive data to support this evidence, as well as the tridimensional structure/model of BMP2-MGP complex, are still missing. In 2008, Yao and co-workers reported the interaction between human BMP4 and MGP and identified the Gla residues and a specific proline residue (Pro64) as important for the binding between the two proteins (Yao et al., 2008). Whether these residues are directly involved in the interaction between BMP2 and MGP remains to be determined.

1.1.1.2 Evolutionary relationship of BMP2/4/16 subfamily members

Organism evolution is closely related to gene duplication and divergence events which are essential mechanisms to create gene novelty (Force et al., 1999). Whole genome duplication (WGD) events had an important role in the construction of vertebrate genomes; it is commonly accepted that two rounds of WGD events (referred as 1R and 2R) have occurred during vertebrate evolution, besides the fish-specific WGD event (3R) that affected teleost fishes (Goode et al., 2011; Amores et al., 1998; Dehal and Boore, 2005; Kuraku et al., 2009).

After duplication, genes can evolve in three possible ways: (1) one of the duplicates is lost or degenerates throughout organism evolution; gene loss occurs at a high frequency (e.g. it is estimated that more than 80% of zebrafish genes duplicated through the 3R were lost later on; Woods et al., 2000); (2) one of the duplicates acquires a complementary function of the original gene; this process is known as subfunctionalization, and both gene copies are maintained, but differentially regulated; and (3) one of the duplicates gain a new and different

function from the ancestor gene; this process is known as neofunctionalization, and both gene copies are maintain but code for proteins with different functions (Fritsch et al., 2010; Fig. 1.7). Interpretation of molecular phylogenies are complicated by gene duplication/deletion and chromosomal rearrangements events that can occur in the genome (Chapman et al., 2004).

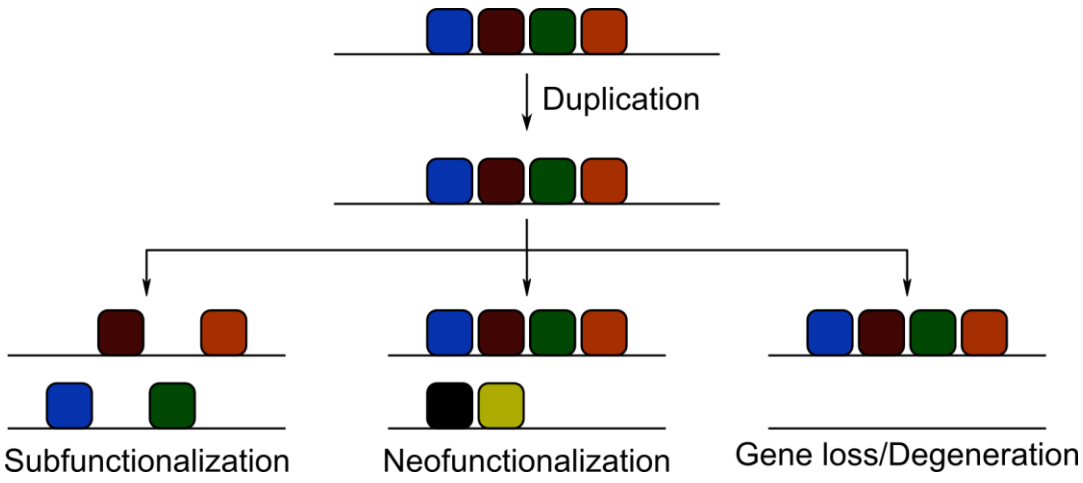


Figure 1.7. Schematic representation of the possibilities after a gene duplication event. A gene subfunctionalization occurs when both copies of the gene are maintained and evolve towards a complementation of the function. Alternatively, if one of the duplicates evolves towards a different function we say that a neofunctionalization have occurred. However, the most common fate of a gene after a duplication event is the loss or degeneration of one of the duplicates Adapted from http://commons.wikimedia.org/wiki/File%3AEvolutionOfDuplicateGenes.png By Veryhuman (Own work).

The BMP family is one of the several families of genes that evolved through gene duplication and divergence from a small number of ancestors and the diversity of members belonging to this family are a consequence of this evolutionary mechanism (Fritsch et al., 2010).

1.2 Fish as a suitable model to study skeletogenesis

Fishes form the largest, most successful and diverse class of vertebrates and represent slightly more than 50% of the extant vertebrate species, most of them belonging to the ray-finned fish class (Actinopterygii; Nelson, 2006). By sharing with mammals a number of important characteristics (e.g. organ systems, gene functions, developmental organization and physiological/biochemical mechanisms, including mechanisms of cartilage and bone formation), fishes have been considered a suitable model organism to study mechanisms

underlying vertebrate development, including skeletogenesis (Berghmans et al., 2005; Kabashi et al., 2011). Among teleost fish, zebrafish *Danio rerio* has specific characteristics and technical advantages that favor its use as a model organism for developmental biology and biomedical research: (1) external fertilization and embryonic development, which facilitate embryo manipulation and visualization; (2) rapid development, almost every body structure are visible 48 hours post fertilization; (3) embryos transparency, which allow a direct observation of internal developing organs and tissues, in particular the skeleton; (4) small size and easy maintenance under laboratory conditions; (5) short generation times, 3 months to get an adult fish; (6) large progeny number, hundreds of eggs per spawning; and (7) availability of various genomic tools, zebrafish genome is almost completely sequenced and annotated and the majority of the genes have assigned human orthologs (reviewed in Laizé et al., 2014). Additionally, the highly developed methodology for genetic manipulation and transgenic availability (reviewed in Laizé et al., 2014) made zebrafish a suitable and valuable tool, that has already been used successfully, to study molecular and cellular mechanisms underlying several human pathologies (e.g. cancer, cardiac disorders, hemophilia, osteoporosis, kidney and liver diseases, central nervous system disorders) and can thus be considered a promising tool on the development and validation of new therapeutics (Berghmans et al., 2005; Kabashi et al., 2011; Huttner et al., 2013; Jagadeeswaran and Liu, 1997; Barrett et al., 2006; Swanhart et al., 2011; Sadler et al., 2005). Anatomical and developmental features of human and fish skeletal elements are remarkably conserved making fish an interesting group to study the evolution of bone tissue (Laizé et al., 2014). Also, and because of their evolutionary position, fish are important models to better understand evolution of gene families (Guo et al., 2011).

Although zebrafish is the number one fish model, other fish have been extensively used in research, in particular bone research, e.g. Japanese medaka, green-spotted pufferfish, Atlantic salmon and Senegalese sole. Gilthead seabream *Sparus aurata* is one of the most important aquaculture species in Portugal and south Europe and because of the high rate of skeletal abnormalities that these animals develop when cultured under intensive farming conditions, the study of skeleton and muscle development are of particular importance. To address these issues, during the last years, an effort to develop biochemical, molecular and cellular tools have been made. Several cell lines have been developed from calcified tissues of the gilthead seabream. These cells have the capacity to mineralize their extracellular matrix and have been characterized in relation to their gene expression patterns and transfectability (Marques et al., 2007; Rafael et al., 2010; Pombinho et al., 2004). This together with the

increasing availability of gene and transcripts sequences available in public sequence databases, makes seabream a suitable model to study skeletogenesis and mechanisms of tissue mineralization, and it has been already used successfully in several genetic and functional studies (Conceição et al., 2008; Ferraresso et al., 2008; Rosa et al., 2010; Rafael et al., 2006; Fonseca et al., 2007)

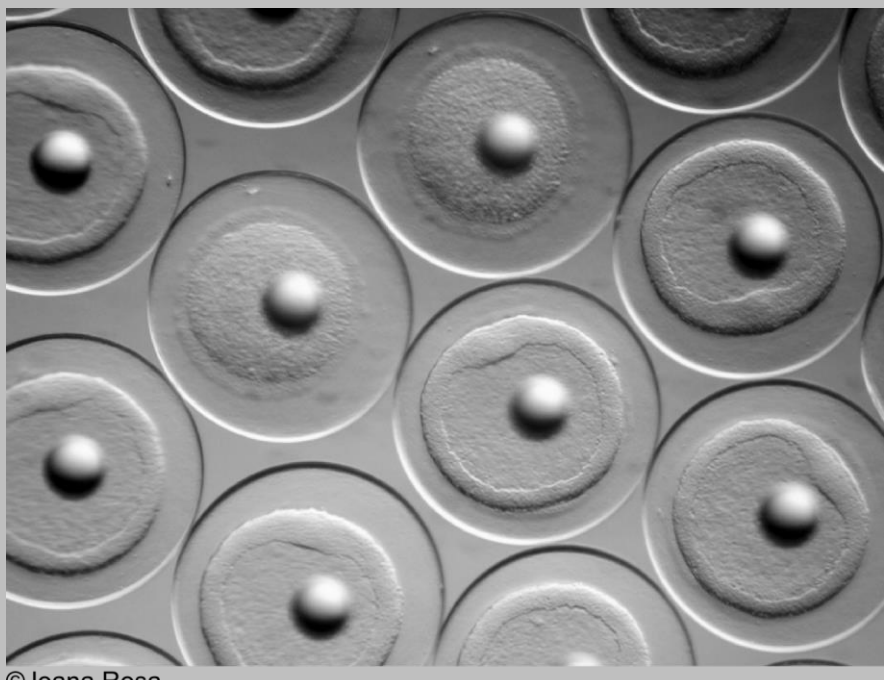
1.3 Objectives of this work

As evidenced above, BMP2 is an important multifunctional growth factor, central to many physiological processes from embryonic development and throughout adulthood. It has a major signaling activity during bone formation and metabolism and defective gene expression or disrupted protein signaling has been associated with pathological conditions. To better understand mechanisms underlying BMP2-related bone disorders it is therefore critical to collect data on bone-related BMP2 function and mechanisms of regulation. The main objective of this work was to investigate the regulation of BMP2 gene expression at transcriptional and post-transcriptional levels but also to study the negative regulation of BMP2 signaling through BMP2-MGP interaction. Because fish is a suitable model organism to study vertebrate development and a promising alternative to mammals to get insights into bone formation, experimental systems based on the gilthead seabream *Sparus aurata*, the zebrafish *Danio rerio* and the Senegalese sole *Solea senegalensis* were used to collect expression and functional data toward these objectives. The 5' flanking region of seabream BMP2 gene (including the proximal promoter region) and the 3' untranslated region of BMP2 transcript were analyzed for the presence of binding sites for transcription factors and microRNAs, respectively, then inserted into luciferase reporter vectors to test the functionality of potential binding sites. Interaction between BMP2 and MGP were explored in a fish-based system using a reporter assay where BMP responsive elements have been coupled to luciferase gene. A secondary objective of this work was to get insights into the evolutionary relationship between members of the BMP2/4/16 subfamily and collect basic data for comparative analysis of BMP2, BMP4 and BMP16 expression and function in fish. To achieve these goals, the taxonomic distribution and molecular phylogeny were inferred from genomic data collected from sequence databases over most vertebrate taxa. Patterns of gene expression during development, in adult tissue and upon retinoic acid exposure were determined through qPCR analysis. Protein tri-dimensional structure was modeled based on the crystallographic structure available for human BMP2. Data collected were compared to

evidence potential similarities/differences in the expression and function of BMP2, BMP4 and BMP16.

Chapter 2

REGULATION OF GILTHEAD SEABREAM BMP2 GENE



©Joana Rosa

Preamble

Bone morphogenetic protein 2 is a morphogen critical to several developmental processes, in particular bone formation. Due to its pleiotropic action it has been proposed that BMP2 gene expression is under tight regulatory mechanisms. Thus, understanding mechanisms underlying transcriptional and post-transcriptional regulation of BMP2 gene is important to increase the knowledge on the physiological role of BMP2, and much remains to be done regarding this particular point. This chapter is divided in two parts based on manuscripts that address the transcriptional and post-transcriptional regulation of gilthead seabream BMP2 gene. In the first manuscript, submitted to Gene, the activity of seabream BMP2 promoter regions and the functionality of bone and cartilage related transcription factors are evaluated using luciferase reporter assays. In the second manuscript, published in the Archives of Biochemistry and Biophysics and resulting from a collaborative work with Daniel Tiago (first author) and Vânia Roberto (third author), the mineralogenic activity of miR20a in fish bone-derived cells and its contribution to the post-transcriptional regulation of gilthead seabream BMP2 is evaluated. As second author, I was directly involved in the functional analysis of the conserved miR20a binding site that was predicted in gilthead seabream BMP2 3'UTR and I mainly contributed to data presented in figures 2.2.2, 2.2.4 and 2.2.7.

CHAPTER 2. GILTHEAD SEABREAM BMP2 GENE REGULATION

2.1 Transcriptional regulation of gilthead seabream bone morphogenetic protein (BMP) 2 gene by bone- and cartilage-related transcription factors

Cátia L. Marques

M. Leonor Cancela

Vincent Laizé

2.1.1 Abstract

Bone morphogenetic protein (BMP) 2 belongs to the transforming growth factor β (TGFβ) superfamily of cytokines and growth factors. While it plays important roles in embryo morphogenesis and organogenesis, BMP2 is also critical to bone and cartilage formation. Protein structure and function have been remarkably conserved throughout evolution and BMP2 transcription has been proposed to be tightly regulated, although few data is available. In this work we report the cloning and functional analysis of gilthead seabream BMP2 promoter. As in other vertebrates, seabream BMP2 gene has a 5' non-coding exon, a feature already present in DPP gene, the fruit fly ortholog of vertebrate BMP2 gene, and maintained throughout evolution. In silico analysis of seabream BMP2 promoter revealed several binding sites for bone and cartilage related transcription factors (TFs) and their functionality was evaluated using promoter-luciferase constructions and TF-expressing vectors. Runt-related transcription factor 3 (RUNX3) was shown to negatively regulate BMP2 transcription and combination with the core binding factor β (CBFβ) further reduced transcriptional activity of the promoter. Although to a lesser extent, myocyte enhancer factor 2C (MEF2C) had also a negative effect on the regulation of BMP2 gene transcription, when associated with SRY (sex determining region Y)-box 9 (SOX9b). Finally, v-ets avian erythroblastosis virus E26 oncogene homolog 1 (ETS1) was able to slightly enhance BMP2

transcription. Data reported here provides new insights toward the better understanding of the transcriptional regulation of BMP2 gene in a bone and cartilage context

2.1.2 Introduction

Bone morphogenetic proteins (BMPs) are multifunctional growth factors that belong to the TGFβ superfamily and form a subfamily with more than 20 members (Bragdon et al., 2011). BMP2 was first identified in bone and later associated with the control of osteogenesis and chondrogenesis through BMP signaling pathway (reviewed by Carreira et al., 2014; Rosen, 2009). Beside its critical role during skeletogenesis, BMP2 is also involved in many other physiological processes, such as embryonic patterning and organogenesis (reviewed by Hogan, 1996). BMP2 gene is flanked by regions classified as gene deserts (long regions without nearby genes) that may contain important regulatory elements, and the presence of long-range elements controlling BMP2 transcription was reported in mammals (Chandler et al., 2007; Dathe et al., 2009). The remarkable conservation of protein structure and function (Carreira et al., 2014) conjugated with its crucial role during development, maintained throughout vertebrate evolution, suggest that BMP2 transcription may be tightly controlled (Sugiura, 1999). The conservation of BMP2 gene, in particular its promoter region, has been reported in mammals (i.e. mouse and human; Abrams et al., 2004; Sugiura, 1999) and binding sites for several bone- and cartilage-related transcription factors (TFs), such as RUNX and SOX9, were predicted. Although the activation of RUNX2 on human BMP2 promoter has not been proved (Helvering et al., 2000), RUNX2 was shown to effectively increase BMP2 gene transcription while BMP2 was also able to regulate RUNX2 transcription in a feedback regulatory mechanism (Choi et al., 2005).

Surprisingly, not much more is known about transcriptional regulation of BMP2 by bone- and cartilage-related TFs and thus much remains to be done regarding this question. The aim of this work is to evaluate the activity of gilthead seabream BMP2 promoter and get insights into its transcriptional regulation by bone- and cartilage-related transcription factors. The presence of *cis*-regulatory elements will be predicted in silico and their functionality will be accessed through luciferase reporter assays.

2.1.3 Materials and methods

2.1.3.1 Amplification of genomic DNA

5' flanking region and intron I of gilthead seabream BMP2 gene were amplified by PCR from a *Sca*I GenomeWalker library (Clontech) using Advantage Polymerase Mix (Clontech), 0.2 µM of Adaptor Primer 1 (AP1; initial PCR) or AP2 (nested PCR) and gene-specific primers SauBMP2_1Rv or SauBMP2_2Rv (initial PCR), and SauBMP2_3Rv or SauBMP2_4Rv (nested PCR), respectively. Nested PCR was performed using a 1:50 dilution of the initial PCR. DNA fragments were separated on agarose gel, purified using GeneJET Gel Extraction kit (Thermo Scientific), cloned into TOPO vector (Life Technologies) and sequenced on both strands. Gene-specific primers were designed according to the sequence available in GenBank (accession no. AY679787) and are listed in Table 1.

Table 2.1.1. PCR primers used in this study

Name	Sequence (5'-3')*
SauBMP2_1Rv	AAGTCTTGGTCAGCGCGGAAACGAA
SauBMP2_2Rv	GGCCTGCGCCTCAGTCCAAACATATT
SauBMP2_3Rv	GGCTCGCGTAGGCAGGACCGATCTA
SauBMP2_4Rv	GGATAAGTCCCGTGGCACCTTCCAGC
SauBMP2_*Hin*dIII_5Rv	CACGCAAGCTTCACTGGTGCTGAGGTATT
SauBMP2_*Hin*dIII_6Rv	CACGCAAGCTTGGCTCGCGTAGGCAGGACCGATCTA
SauBMP2_*Xho*I_1Fw	CCGGAGCTCGAGCCGCCTGCCCCACCATCAT
SauBMP2_*Kpn*I_2Fw	CCGGAGGGTACCGCCTCCTCGTCAGGTAAA
SauBMP2_*Xho*I_3Fw	CCGGAGCTCGAGTCACGTTCACGGGGAAGCATTGC
SauBMP2_*Xho*I_4Fw	CCGGAGCTCGAGGTGTGAGTTTCCAGGATGTGTA
SauBMP2_*Xho*I_5Fw	CCGGAGCTCGAGTTTGTTTGACATGAGAAGGGG
SauBMP2_*Xho*I_6Fw	CCGGAGCTCGAGGAGGTGCTTTATCGCGGACA
SauBMP2_*Xho*I_7Fw	CCGGAGCTCGAGACTGCTCTCTCTCGTGTTC

*Underlined sequences indicate recognition site for endonucleases cited in primer name

2.1.3.2 Preparation of promoter-luciferase and deletion constructs

Constructs containing 5' flanking region and intron I (construct C1, -1531/+53) or only intron I (construct C2, +301/+1282) of gilthead seabream BMP2 gene were amplified using reverse primer SauBMP2_*Hin*dIII_5Rv in combination with forward primers SauBMP2_*Xho*I_1Fw and SauBMP2_*Kpn*I_2Fw, respectively. Deletion constructs of 5' flanking region were amplified using reverse primer SauBMP2_*Hin*dIII_6Rv in combination with forward primers SauBMP2_*Xho*I_1Fw (construct C3, -1531/+53), SauBMP2_*Xho*I_3Fw (construct C4, -842/+53), SauBMP2_*Xho*I_4Fw (construct C5, -656/+53), SauBMP2_*Xho*I_5Fw (construct C6, -367/+53), SauBMP2_*Xho*I_6Fw (construct C7, -294/+53) and SauBMP2_*Xho*I_7Fw (construct C8, -59/+53). DNA fragments were digested with *Hin*dIII and *Xho*I or *Kpn*I endonucleases and directionally cloned into pGL3 vector (Promega) upstream the firefly luciferase gene. All construct were sequenced on both strands to confirm direction and absence of mutations. Primers used for PCR amplification of these constructs are listed in Table 1.

2.1.3.3 In silico sequence analysis

Presence of *cis*-regulatory elements, i.e. transcription factor binding sites, in the 5' flanking region and intron I of gilthead seabream BMP2 gene, were predicted using MatInspector (V7.1; Cartharius et al., 2005) at www.genomatix.de and PATCH (Vpublic 1.0; Chekmenev et al., 2005) at www.gene-regulation.com. Sites with scores below 0.75 (MatInspector) and 0.85 (PATCH) were not considered. Repetitive sequences were identified using RepeatMasker software (Vopen-4.0.5) at www.repeatmasker.org.

2.1.3.4 Cell culture and transient transfection assays

Human embryonic kidney (HEK) 293 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Life Technologies) supplemented with 10% fetal bovine serum (FBS; Sigma-Aldrich), 1% of penicillin-streptomycin (Life Technologies) and 1% of L-glutamine (Life Technologies), and maintained at 37ºC in a 5% CO_2 humidified atmosphere. Sub-confluent cultures were trypsinized every three days and cells seeded in a 10-cm plate at a density of 8.7×10^5 cells/plate. The day before the transfection, cells were seeded in a 24-well plate at a density of 5×10^4 cells/well, then further cultured for 16 h. Cultures at 50-60%

confluence were transfected with 250 ng of each of the DNA constructs using 1 µl of X-tremeGENE HP DNA transfection reagent (Roche). When appropriate, expression vectors (50 ng; pCMX backbone) containing the coding sequence of zebrafish ETS1 (KF774190), CBFβ (KF709197), RUNX3-MASN (AB043789), MEF2Ca (BC059188), MEF2Cb (EU825718) and SOX9b (NM_131644), under the control of CMV promoter, were co-transfected with selected constructs of BMP2 promoter. pRL-null vector (Promega), which express *Renilla* luciferase (Rluc) but lacks promoter and enhancer elements, was used in all the transfections (25 ng) to normalize the firefly luciferase (Fluc) activity. After 48 h, cells were lysed and luciferase activity was measured using Dual-Luciferase Reporter Assay system (Promega) in a BioTek Synergy 4 plate reader. Luciferase activity was determined from the ratio Fluc/Rluc.

2.1.4 Results and Discussion

2.1.4.1 BMP2 gene structure has been conserved throughout evolution

Sequences of BMP2 gene were collected from several vertebrate species using on-site Blast facilities of Ensembl database and their structure was compared with that of the gilthead seabream gene, recently cloned in our lab (GenBank accession no. AY679787) and with fruit fly DPP gene, the ancestor of vertebrate BMP2 genes (Kingsley, 1994; Fig. 2.1.1).

Gene structure has been remarkably conserved throughout evolution and orthologs in vertebrates and fruit fly share the same simple structure: three exons – including a 5´non-coding exon – and two introns inserted within the same phase in all species evaluated. Although its size is variable (from 246 to 1212 nucleotides), BMP2 5' non-coding exon has been conserved throughout evolution and may work as an additional mechanism of regulation (Barrett et al., 2013). The presence of 5' non-coding exons has been observed in genes which transcription is under the control of alternative promoters (Conceição et al., 2008; Banday et al., 2012). In mammals, while some studies report the existence of at least two major transcription start sites (TSSs) driven by two alternative promoters (Feng et al., 1997; Ghosh-Choudhury et al., 2001; Sugiura, 1999), others defend the occurrence of a single TSS regulated by a single promoter located in the 5' flanking region of the gene (Helvering et al., 2000; Heller et al., 1999).

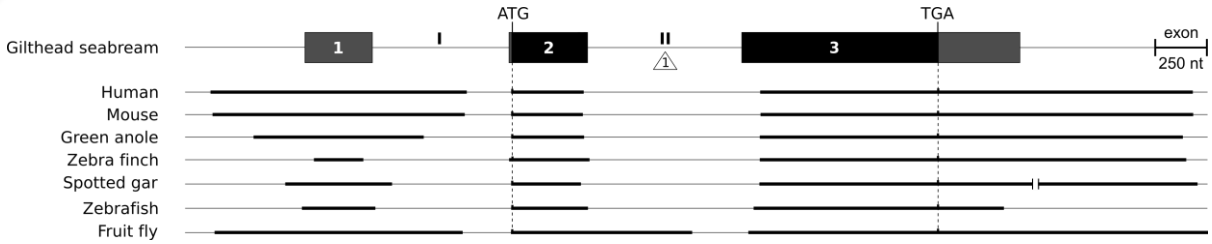


Figure 2.1.1. Schematic representation of BMP2 gene structure. Exons are displayed as *gray boxes* (non-coding exons) and *black boxes* (coding exons) and indicated with arabic numbers in gilthead seabream scheme, and as *thick lines* in other species. Introns are displayed as *solid black lines* and indicated with roman numbers in gilthead seabream scheme, and as *thin lines* in other species. Phase of intron insertion is indicated in *triangles*. *Dashed lines* indicate translation initiation and termination sites (aligned according to gilthead seabream sites ATG and TGA, respectively). Accession numbers of BMP2 gene sequences in Ensembl database: Human (*Homo sapiens*, ENSG00000125845); Mouse (*Mus musculus*, ENSMUSG00000027358); Green anole (*Anolis carolinensis*, ENSACAG00000003113); Zebra finch (*Taeniopygia guttata*, ENSTGUG00000006434); Spotted gar (*Lepisosteus oculatus*, ENSLOCG00000016442); Zebrafish (*Danio rerio*, ENSDARG00000041430; isoform 2b) and fruit fly (*Drosophila melanogaster* DPP, FBgn0000490).

Our survey of GenBank and Ensembl databases for BMP2 transcripts with different 5'ends failed to provide conclusive evidence of alternative promoter usage (results not shown). Thus, both the 5' flanking region and intron I were analyzed in silico for the presence of regulatory elements and further tested, independently or in combination, for their capacity of regulating luciferase gene transcription.

2.1.4.2 In silico analysis and basal activity of gilthead seabream BMP2 promoter and intron I

RepeatMasker software revealed the presence of several repetitive sequences, namely T-rich regions, tri and tetra nucleotide repeats, in both 5´flanking region and intron I of gilthead seabream BMP2 gene (Fig. 2.1.2). Repetitive sequences, also known as DNA satellites, are commonly found in vertebrate genomes (Tomilin, 2008). When discovered, repetitive sequences were described as "junk" or "parasitic" DNA (Doolittle and Sapienza, 1980; Orgel and Crick, 1980), but nowadays they are widely recognized as essential for genome function (Shapiro and von Sternberg, 2005). Tandem repeats can vary both in number and in length of the repeat unit (Gemayel et al., 2010) and certain types of repeats have been associated with protein binding sites, interaction with transcription factors or disease development (Baldi and Baisnee, 2000). T-rich regions have been described as

transcriptional activators of both prokaryotic and eukaryotic promoters (Haque et al., 2004; Sohaskey et al., 1999; Nishi and Itoh, 1986; Kube et al., 1999) while some classes of triplet DNA repeats were associated with the development of neurodegenerative disorders, such as Huntington's disease and fragile X syndrome (Walker, 2007).

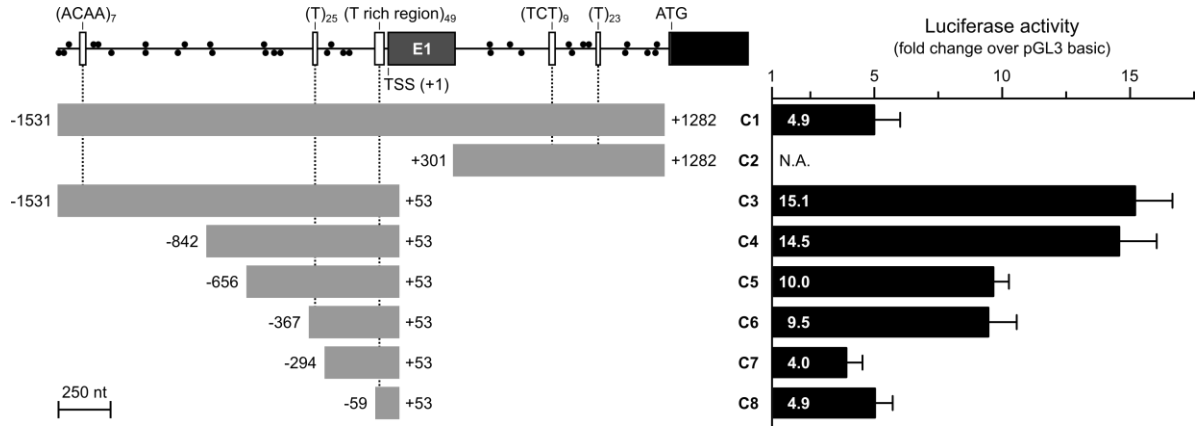


Figure 2.1.2. Basal activity of gilthead seabream BMP2 gene promoter constructs in HEK-293 cells. Promoter deletion constructs (C1 to C8) are presented on the left side as *light gray boxes*. Nucleotide positions are given according to currently known transcription start site (*TSS*). *White boxes* indicate repetitive sequences (motif and number of repetitions are indicated on the top of each box). Non-coding and coding exons are displayed as *dark gray* and *black boxes*, respectively. *ATG* indicates translation initiation. *Black circles* indicate *in silico* predicted SP1 sites. Luciferase activity (Fluc/Rluc; n = 4) is presented as fold change over the activity of promoter less pGL3 basic vector. *N.A.*, not active.

Tetranucleotide repeats are among the most common repeats found in the genome of vertebrate species, and elevated microsatellite alterations at selected tetranucleotide repeats were observed in several types of cancer (Bacolla et al., 2008; Katti et al., 2001). No TATA or CAAT consensus sequences were predicted in gilthead seabream BMP2 promoter, upstream the TSS, using PATCH software. TATA-less promoters, also known as dispersed promoters, are very common in vertebrate genes (Barrett et al., 2013). Several BMP genes have been reported to have TATA-less promoters (Kawai and Sugiura, 2001; Hino et al., 1996; Shore et al., 1998; Simon et al., 2002; Tamada et al., 1998), including human and mouse BMP2 genes (Sugiura, 1999; Ghosh-Choudhury et al., 2001). A survey of other BMP2 genes (i.e. zebra finch, spotted gar, zebrafish and fruit fly; Fig. 2.1.1) failed to identify TATA boxes upstream TSS, suggesting that the presence of TATA-less promoters in BMP2 gene is a common feature. To further confirm the TATA-less nature of BMP2 promoter, various SP1 binding sites, GC-rich regions normally present in dispersed promoters (Pugh and Tjian,

1991; Smale and Kadonaga, 2003), were predicted using PATCH software in the 5' flanking region and intron I of gilthead seabream BMP2 gene (Fig. 2.1.2).

In order to evaluate the functionality of these regions in regulating BMP2 transcription, 5' flanking region and intron I or only intron I (C1 and C2, respectively; Fig.2.1.2) were cloned into pGL3 vector, upstream from the firefly luciferase gene. With the purpose of identifying regulatory regions in the 5' flanking region, several deletion constructs (C3 to C8; Fig.2.1.2) were individually transfected into HEK-293 cells and their activity determined from firefly luciferase activity measurements normalized with the activity of the promoter less pGL3 basic vector. While 5' flanking region (C3) increased 15 times the luciferase activity in HEK-293 cells, intron I (C2) failed to trigger any transcriptional activity, indicating that (1) intron I does not have any transcriptional activity, (2) does not function as a stand-alone and alternative promoter but as an enhancer/silencer of the main promoter (i.e. the 5' flanking region), or (3) HEK-293 cells are not suitable to study its transcriptional activity (e.g. they do not express the necessary factors). Although this should be further confirmed, the 3-fold decrease in luciferase activity observed when 5'flanking region and intron I were combined (C1) suggest that intron I may serve as a silencer of BMP2 transcription and in fact, several TF-binding sites were predicted in this region (results not shown).

Transfections of deletion constructs related to the 5' flanking region (C4-C8) resulted in a gradual decrease of luciferase activity that may be linked to the presence of binding sites for positive regulators in the deleted promoter regions. At least two enhancers may be present in BMP2 promoter, as suggested by the two drops in luciferase activity observed from C4 to C5 and from C6 to C7. Although additional studies are required to confirm the presence of enhancers in these regions (e.g. site-directed mutagenesis), we propose that the T-rich sequence identified through *in silico* analysis (present in C6 but absent in C7) may be responsible for the halving of the luciferase activity observed in C7.

2.1.4.3 RUNX3 and CBFß interact to regulate seabream BMP2 promoter

The 5´flanking region of gilthead seabream BMP2 gene was analyzed using MatInspector and PATCH online tools, to identify *cis*-regulatory elements that may be involved in its regulation. Several putative binding sites for Runt-related transcription factors (RUNX) were identified (Fig. 2.1.3A), similarly to what has been described in human BMP2

and BMP4 gene promoters (Helvering et al., 2000). To evaluate the functionality of these sites, promoter constructs were co-transfected with zebrafish RUNX3 (MASN isoform) in HEK-293 cells either alone or in combination with zebrafish CBFβ, a transcriptional co-regulator of RUNX factors (Fig. 2.1.3B; Warren et al., 2000).

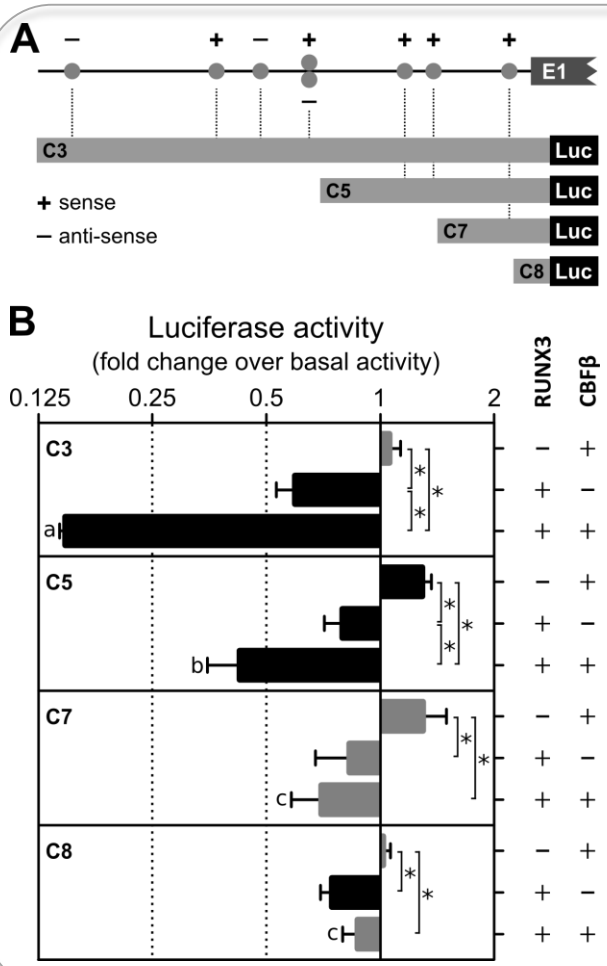


Figure 2.1.3. Effect of RUNX3 and CBFβ on the transcriptional activity of gilthead seabream BMP2 gene promoter, determined from luciferase activity of promoter constructs. **A**, Schematic representation of BMP2 promoter constructs with putative RUNX binding sites indicated as *gray circles*. *E1*, exon 1. **B**, Luciferase activity (Fluc/Rluc) of promoter constructs co-transfected with expression vectors carrying zebrafish RUNX3 or CBFβ in HEK-293 cells. Values are presented as fold induction over the basal activity of each construct. *Gray bars* represent values that are not significantly different from the basal activity. *Asterisks* and *letters* indicate values significantly different for the same construct and between different constructs, respectively (one-way ANOVA followed by Tukey's post-test; $P<0.05$).

RUNX3 expression decreased luciferase activity (1.7 fold in C3, and 1.3 fold in C5, C7 and C8), suggesting that it may work as a negative regulator of BMP2 transcription. While CBFβ expression did not significantly change luciferase activity, the co-expression of RUNX3 and CBFβ strongly decreased luciferase expression in C3 (6.7 fold) and in C5, although to a lesser extent (2.3 fold). CBFβ cannot bind to DNA and was therefore not expected to trigger any change in BMP2 promoter activity. But it is known to efficiently mediate the interaction of RUNX family members with the transcription machinery (Blake et al., 2000; Wang et al., 1993) and the potentiation of the transcriptional regulation of BMP2 gene by RUNX3 further confirm the capacity of CBFβ in co-regulating gene transcription. Data reported by Kundu and co-workers (Kundu et al., 2002) revealed that, in bone and

cartilage tissues, CBFβ interacts with RUNX2, enhancing its transactivation capability. Similarly RUNX genes were shown to be susceptible to auto- and cross-regulation by RUNX family members (Drissi et al., 2000; Spender et al., 2005), an effect further enhanced upon addition of its CBFβ partner (Conceição et al., 2013; Simões et al., unpublished data).

A significant decrease in luciferase activity was observed when the 4 putative RUNX binding sites specific for C3 (-1429/-1424; -984/-980; -847/-843; -695/-688), and later on the 2 sites specific for C5 (-387/-380; -299/-295) were removed, indicating the presence of functional responsive elements for RUNX3/CBFβ in the regions -1531/-656 and -656/-294 regions. Although expression of RUNX3 slightly decreased luciferase activity in C8, co-expression with CBFβ failed to significantly change it, indicating that the binding site located in C7 (-64/-60) is probably not functional. Future studies should aim at identifying the functional RUNX sites (e.g. through site-directed mutagenesis of putative DNA binding elements) and at confirming RUNX3/CBFβ interaction (e.g. through electrophoretic mobility shift assay and chromatin immunoprecipitation assay).

Besides the well-documented role in neurogenesis and in gastric cancer progression, RUNX3 has been shown to be involved in mechanisms of bone and cartilage formation, inhibiting RUNX1 expression and potentiating RUNX2 during the process of endochondral ossification (Soung et al., 2007). Recently, expression of RUNX3 gene was also shown to be up-regulated during tissue mineralization in the notochord of Atlantic salmon (Wang et al., 2014). Regulation of the transcriptional activity of BMP2, a known regulator of bone metabolism, by RUNX3 provides new evidences for the role of RUNX3 in bone and cartilage metabolism.

2.1.4.4 MEF2C/SOX9b negatively regulate BMP2 gene transcription

The presence of several *cis*-regulatory elements related to MEF2 and SOX9 was also predicted in gilthead seabream BMP2 promoter (Fig. 2.1.4A). To evaluate their functionality, promoter constructs were co-transfected in HEK-293 cells with vectors expressing zebrafish MEF2C (a mixture of zebrafish MEF2Ca and MEF2Cb were used, since both forms produced similar results) or SOX9b. Because MEF2C is known to physically interact with SOX family members (Agarwal et al., 2011), MEF2C and SOX9b were also co-transfected in some experiments (Fig. 2.1.4B). A mild repression (up to 2 fold) of BMP2 promoter activity was observed upon co-transfection of C3 promoter construct with MEF2C or SOX9b

expression vectors, indicating that both factors are negative regulators of BMP2 gene transcription.

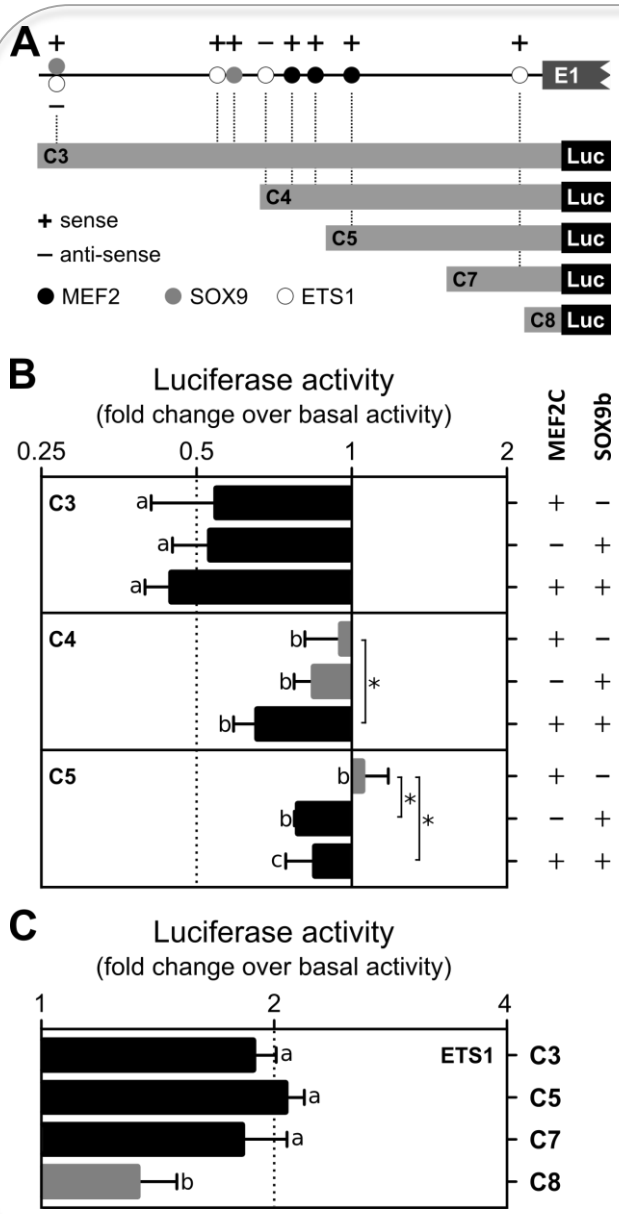


Figure 2.1.4. Effect of MEF2C, SOX9b and ETS1 on the transcriptional activity of gilthead seabream BMP2 gene promoter determined from luciferase activity of promoter constructs. **A**, Schematic representation of BMP2 promoter constructs with putative MEF2, SOX9 and ETS1 binding sites indicated as *circles*. *E1*, exon 1. **B**, Luciferase activity (Fluc/Rluc) of promoter constructs co-transfected with expression vectors carrying zebrafish MEF2C (1:1 mixture of MEF2Ca and MEF2Cb), SOX9b or ETS1 in HEK-293. Values are presented as fold induction over the basal activity of each construct. *Gray bars* represent values that are not significantly different from the basal activity. *Asterisks* and *letters* indicate values significantly different for the same construction and between different constructs, respectively (one-way ANOVA followed by Tukey's post-test; $P<0.05$).

Co-transfection of both factors did not significantly change luciferase activity but deletion of the region containing the two predicted binding sites for SOX9 in C4 abolished not only the negative regulation by SOX9b, but also the negative regulation by MEF2C, even though no binding site for MEF2C was removed. Although we cannot exclude that in silico analysis failed to predict MEF2 binding site(s) in this region, we propose that MEF2C regulation of BMP2 gene transcription is SOX9b-dependent, in a way similar to what has been reported for COL10A1 gene (Dy et al., 2012). Inhibition in C3 upon individual expression of MEF2C or SOX9b could be related to the presence of endogenous SOX9,

which expression has been detected at basal levels in HEK-293 cells (Blache et al., 2004). The negative regulation by MEF2C/SOX9b was attenuated (1.2 fold) upon deletion of the region containing two predicted MEF2C responsive elements (C5). The decrease on luciferase activity observed in C4 and C5 upon expression of SOX9b or co-expression of MEF2C and SOX9b could be related to the presence of a SOX9 binding site(s) not predicted through *in* silico analysis.

Both MEF2C and SOX9 factors have been implicated in the regulation of bone and cartilage formation (Mackie et al., 2008; Dy et al., 2012). There are also evidences of MEF2C and SOX9 being regulators of BMP signaling pathway mediators (Dalcq et al., 2012; J Wu et al., 2010) and vice versa, the expression of both transcription factors was shown to be regulated by BMP2 (Kawakami et al., 2006; Zheng et al., 2013). In agreement with data presented here, Liao and co-workers recently demonstrated that BMP2 expression is lowered in SOX9-enhanced chondrogenesis in mouse cells (Liao et al., 2014). Dy and co-authors have recently reported a cooperation between SOX9 and MEF2C during cartilage formation (Dy et al., 2012), in a way similar to the cooperation reported here. In this study, the stimulation of COL10A1 gene transcription by MEF2C was SOX9-dependent, and MEF2C-enhancing capacity was lost upon SOX9 inactivation, even after MEF2C overexpression (Dy et al., 2012)

2.1.4.5 ETS1 enhances BMP2 gene transcription

Four binding sites for ETS1 were predicted in gilthead seabream BMP2 promoter and their functionality was tested through co-transfection in HEK-293 cells of promoter deletion constructs and vector expressing zebrafish ETS1 (Fig. 2.1.4C). A 2 fold increase of luciferase activity was observed in C3 upon expression of ETS1, indicating that it is a positive albeit weakly regulator of BMP2 gene transcription. Deletion of the region containing 3 of the responsive elements in C5 and C7 constructs (-1502/-1496; -984/-980; -847/-843) did not affect ETS1 transactivation of BMP2 promoter, suggesting that those sites are most likely not functional. On the contrary, deletion of the region containing the responsive element located in the C8 construct (-64/-60), decreased luciferase activity to basal levels. Although this should be confirmed (e.g. through site-directed mutagenesis of the specific DNA binding elements), we propose that the -64/-60 binding site is functional and accounts for the totality of ETS1 activity. Members of ETS family of transcription factors are expressed at the onset of bone formation (Raouf and Seth, 2000) and have been associated with mechanisms

regulating osteogenic and chondrogenic processes in vertebrates (Rosa et al., 2014; Gao et al., 2005). Although ETS1 has been shown to cooperate with other regulatory proteins to modulate transcription of bone and cartilage related genes – e.g. with RUNX2 to regulate osteopontin gene transcription (Miyake et al., 1998), and with retinoic acid receptor to regulate RA-induced expression of PTHrP (Karperien et al., 1997) - no cooperation was observed in our experimental system. Our data suggest that ETS1 is able to activate seabream BMP2 transcription and provides additional evidence towards the osteogenic/chondrogenic role of ETS1.

2.1.5 Conclusions

We have collected within the scope of this work valuable data towards a better understanding of the transcriptional regulation of BMP2 gene. The high conservation of BMP2 gene structure among vertebrates – in particular the presence of a 5' non-coding exon – and the prediction of similar binding sites for RUNX/MEF2/SOX9/ETS1 transcription factors in gilthead seabream and human BMP2 genes suggest that these results, collected in a fish system, are probably valid in other vertebrate system, in particular in human. Functional analysis of promoter-luciferase constructs suggests that 5' flanking region contains several responsive elements for selected transcription factors and therefore corresponds to a functional BMP2 gene promoter, while intron I might contribute to silence promoter activity. The functionality of several binding sites for bone and cartilage related factors predicted in silico was confirmed *in vitro*, highlighting the relevance of performing in silico analysis prior to functional assays. RUNX3 is a negative regulator of BMP2 gene transcription and its activity is enhanced by the co-factor CBFβ. Similarly, SOX9b and MEF2C, in a SOX9-dependent manner, are also negative regulators of BMP2 gene transcription, while ETS1 stimulates BMP2 transcription, although weakly. Current knowledge on the transcriptional regulation of seabream BMP2 gene promoter has been summarized in Fig. 2.1.5.

Although the data reported here will require further studies, it provides new evidences on the regulation of BMP2 transcriptional activity by bone- and cartilage- related transcription factors.

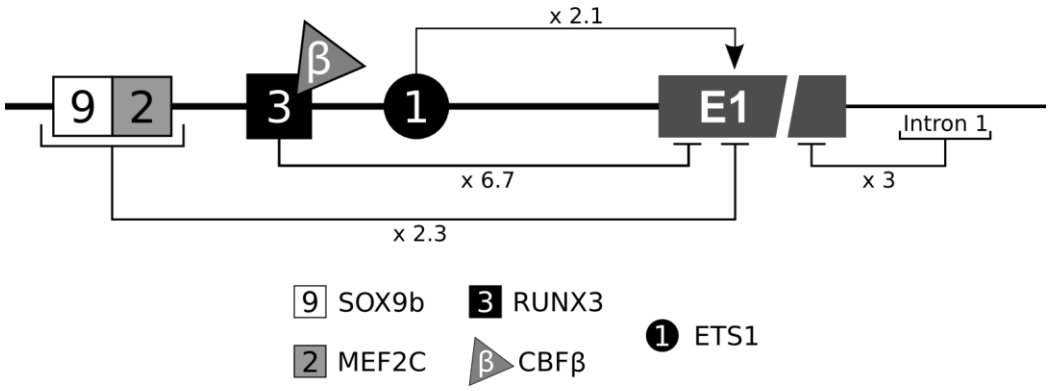


Figure 2.1.5. Schematic representation of proposed transactivation of gilthead seabream BMP2 gene by RUNX3, CBFβ, MEF2C, SOX9b and ETS1. *E1*, exon 1. *Arrows* and *intersected lines* indicate activation and repression, respectively. Fold changes in luciferase activity are indicated above the line respective to each transcription factor or pairs of transcription factors.

2.1.6 Acknowledgments

CM was supported by a doctoral grant (SFRH/BD/39964/2007) from the Portuguese Foundation for Science and Technology (FCT). The authors are grateful to Brigite Simões for the cloning of zebrafish RUNX3 and CBFβ, to Andreia Adrião for the cloning of zebrafish MEF2Ca and MEF2Cb and to Marlene Trindade for the subcloning of zebrafish SOX9b into pCMX vector. SOX9b was a kind gift from Prof. Robert Kelsh.

2.2. Mir20a regulates in vitro mineralization and BMP signaling pathway by targeting BMP2 transcript in fish

Daniel M. Tiago
Cátia L. Marques
Vânia P. Roberto
M. Leonor Cancela
Vincent Laizé

2.2.1 Abstract

MicroRNAs (miRNAs) are important regulators of vertebrate development but their role during skeletogenesis remains largely unknown. In this regard, we investigated the mineralogenic activity of miR20a, a miRNA associated with osteogenesis, in fish bone-derived cells. Expression of miR20a was up-regulated during differentiation and its overexpression inhibited mineralization, suggesting a role in fish tissue calcification. In this regard, a conserved miR20a binding site was identified in bone morphogenetic protein 2 (BMP2) 3'UTR and its functionality was evidenced through luciferase assays, and further confirmed by western blot and qPCR. Type II BMP receptor (BMPR2) is also targeted by miR20a in mammalian systems and evidence was collected for the presence of a binding site in fish sequences. We propose that miR20a is a regulator of BMP pathway through specific action on BMP2 and possibly BMPR2. Overexpression of miR20a was also shown to up-regulate matrix Gla protein (MGP) transcript, a physiological inhibitor of calcification previously found to form a complex with BMP2. We propose that MGP may play a role in the anti-mineralogenic effect promoted by miR20a by decreasing availability of BMP2. This study gives new insights into miRNA-mediated regulation of BMP2, and sheds light into the potential role of miR20a as a regulator of skeletogenesis.

2.2.2 Introduction

Skeletogenesis is a complex process (Blair et al., 2002; Karsenty and Wagner, 2002) and many key players and cellular mechanisms still remain to be identified. In this regard, the

post-transcriptional regulation of skeletal genes has been largely under studied and data on the skeletogenic and osteogenic role of microRNAs (miRNAs) is scarce.

MiRNAs are small non coding RNAs of approximately 22 nucleotides that binds to target mRNAs preventing their translation or promoting their degradation (Bartel et al., 2004). Through their post-transcriptional activity, miRNAs have been shown to regulate a broad range of biological processes (Guarnieri and DiLeone, 2008), including skeletogenesis as evidenced by defective bone and cartilage formation resulting from conditional inactivation of DICER (the enzyme processing pre-miRNA into mature miRNA) in mouse osteoprogenitor cells (Gaur et al., 2010) and in chondrocytes (Kobayashi et al., 2008). MiRNAs were also shown to specifically affect in vitro differentiation of chondrocytes (Martinez-Sanchez et al., 2012), osteoblasts (Martinez-Sanchez et al., 2012; Kapinas et al., 2009; Luzi et al., 2008; Gaur et al., 2010) and osteoclasts (Sugatani and Hruska, 2007).

Among those miRNAs, miR20a was recently identified as capable of promoting bone cell differentiation by targeting antagonists of the bone morphogenetic protein (BMP) signaling pathway in human mesenchymal stem cells (J Zhang et al., 2011). BMP pathway participates in osteoblast differentiation and plays a major role in the development of skeletal tissues (a mechanism that was shown to be conserved from fish to mammals; Yamaguchi et al., 2000; Rafael et al., 2006). Interestingly, a recent study reported the targeting of intermediates of the BMP signaling by miR20a in biological systems not related to bone or cartilage (Brock et al., 2009; Brock et al., 2012). Available results indicate that miR20a may regulate BMP signaling pathway through direct and indirect mechanisms and indicates that mechanisms for miR20a action on bone formation are far from being understood.

Because they share significant similarities with mammals in organ/tissue development, bony fish represent a suitable alternative to mammals to investigate mechanisms associated with vertebrate development (McGonnell and Fowkes, 2006), in particular skeletogenesis (Ingham, 2009). The conservation of miRNA-related mechanisms throughout vertebrate evolution (Giraldez et al., 2005; Schier and Giraldez, 2006; Takacs and Giraldez, 2010) also indicates the suitability of bony fish *in vivo* and *in vitro* models to investigate the role of miRNA during skeletogenesis/osteogenesis.

In this work, the ABSa15 cell line – developed from calcified branchial arches of the marine teleost gilthead seabream (*Sparus aurata*, Linnaeus, 1758) and capable of in vitro mineralization (Marques et al., 2007) – was used to investigate the post-transcriptional regulation of two key players of the BMP signaling pathway by miR20a. Data collected

provided evidence for the role of miR20a in the regulation of skeleton development, thus demonstrating the suitability of fish systems to study mechanisms of post-transcription.

2.2.3 Materials and methods

2.2.3.1 Cell culture and extracellular matrix mineralization

ABSa15 is a cell line previously developed from calcified branchial arches of the marine teleost gilthead seabream (*Sparus aurata*, Linnaeus, 1758) that is capable of *in vitro* mineralization (Marques et al., 2007), and was recently deposited in the European Collection of Cell Cultures (Ref. 13112201; see also Supplementary Fig. S1). ABSa15 were cultured at 33 ºC in a humidified 10% CO2 atmosphere in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), as described previously for VSa13 and VSa16 cells (Pombinho et al., 2004). Human Embryonic Kidney 293 (HEK 293) cells were cultured at 37 ºC in a humidified 5% CO2 atmosphere in DMEM supplemented with 10% FBS. For mineralization experiments, ABSa15 cells were seeded in 24-well plates at 2×10^4 cells/well and allowed to proliferate for 1 week. Then, extracellular matrix (ECM) mineralization was induced in confluent cultures by supplementing medium with 50 µg/ml of L-ascorbic acid, 10mM ß-glycerophosphate and 4mM $CaCl_2$. At appropriate times, mineral deposition was revealed through von Kossa staining and quantified by densitometry analysis (Pombinho et al., 2004). Culture medium was renewed twice a week.

2.2.3.2 RNA extraction and gene expression analysis

Total RNA was extracted from cell cultures as described by Chomczynski and Sacchi (Chomczynski, 1987) and quantified by UV spectrophotometry (NanoDrop ND-1000). Quantitative real-time PCR (qPCR) analysis of miRNAs and mRNAs was performed using the StepOnePlus system (Applied Biosystems, Invitrogen). For qPCR analysis of mRNA expression, total RNA (1 µg) was treated with RQ1 RNase-free DNase (Promega), then reverse-transcribed using MMLV-RT (Invitrogen) and oligo-d(T)-adapter primer (Supplementary Table 1). PCR amplifications were performed using 10 ng of cDNA, gene-specific primers (Supplementary Table 1) and SsoFast EvaGreen Supermix (Bio-Rad) according to manufacturer instructions. For qPCR analysis of miRNA expression, total RNA (1 µg) was polyadenylated and reverse-transcribed using NCode miRNA First-Strand cDNA

Synthesis kit (Invitrogen) according to manufacturer instructions. PCR amplifications were achieved using miRNA-specific primers (Supplementary Table 1) and NCode SYBR miRNA qRT-PCR kit (Invitrogen). Relative mRNA and miRNA expression was calculated using the ΔΔCt method (Pfaffl, 2001) and normalized using expression of three housekeeping genes (ribosomal protein L27a (RPL27a), 18S, and β-actin) for mRNAs, and U6 small nuclear RNA (U6) for miRNAs.

2.2.3.3 Vector construction

For luciferase assays, the 3' untranslated region (UTR) of gilthead seabream BMP2 transcript was inserted into XbaI site of pGL3 control vector (Promega) downstream of firefly luciferase (F-Luc) coding sequence. 3'UTR was amplified from Marathon cDNA libraries (Clontech) using gene-specific primers (Supplementary Table S1) and Klen Taq Polymerase mix (Clontech). Mutations in polyadenylation signal and miR20a binding sites were achieved using 50 ng of pGL3–3'UTR constructs, specific primers containing point mutations (designed according to manufacturer instructions; Supplementary Table S1) and the QuickChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies). PCR reaction was treated with DpnI restriction endonuclease to cut methylated template DNA and used to transform XL10-Gold cells (Agilent Technologies) according to manufacturer instructions. For miR20a overexpression, oligonucleotides containing forward and reverse sequences of zebrafish pre-miR20a (Supplementary Table S1) were annealed then inserted into pcDNA6.2-GW/EmGFP-miR vector downstream of GFP coding sequence using the BLOCK-iT Pol II miR RNAi Expression Vector kit (Invitrogen). The BMP responsive luciferase reporter vector (BRE-Luc) was kindly provided by Dr. Peter ten Dijke (Korchynskyi and ten Dijke, 2002). Luciferase

2.2.3.4 Luciferase assays

HEK 293 cells were seeded in 24-well plates at 8×10^4 cells/well, further cultured for 14–16 h and transfected with 5 ng of the pGL3–3'UTR construct and 12.5 ng of pRL-TK vector (Promega) carrying renilla luciferase gene (RLuc) using 1.5 µl of X-treme GENE HP transfection reagent (Roche). When appropriate, 5 ng of pcDNA6.2-miR20a vector was co-transfected in HEK 293 cells. ABSa15 cells were seeded in 12-well plates at 8×10^4 cells/well, further cultured for 14–16 h and transfected with 600 ng of BRE- Luc vector and

600 ng of pRL-SV40 vector using 1.5 µl of FuGene HD (Roche). After 48 h, transfected cells were lysed and luciferase activities were measured using Dual-Luciferase Reporter Assay system (Promega). Relative luciferase activity was determined from the ratio FLuc/RLuc.

2.2.3.5 Establishment of cell clones overexpressing miR20a

ABSa15 cells were seeded in 6-well plates at 2×10^5 cells/well, further cultured for 14–16 h and transfected with 2.4 µg of pcDNA6.2-miR20a vector using 3 µl of FuGeneHD (Roche). After 24 h, cells were sub-cultured into a 10-cm culture dish containing DMEM supplemented with 2 µg/ml of blasticidin (determined as described in the manual of BLOCK-iT Pol II miR RNAi Expression Vector kit). After approximately 30 days in selective medium (renewed once a week), cell colonies expressing GFP were identified using Olympus IX-81 fluorescence microscope and sequentially sub-cultured into 24-well, 6-well and 10-cm culture dishes.

2.2.3.6 Establishment of cell clones overexpressing miR20a

ABSa15 cells were seeded in 6-well plates at 2×10^5 cells/well, further cultured for 14-16 h and transfected with 2.4 µg of pcDNA6.2-miR20a vector using 3 µl of FuGeneHD (Roche). After 24 h, cells were sub-cultured into a 10-cm culture dish containing DMEM supplemented with 2 µg/ml of blasticidin (determined as described in the manual of BLOCK-iT Pol II miR RNAi Expression Vector kit). After approximately 30 days in selective medium (renewed once a week), cell colonies expressing GFP were identified using Olympus IX-81 fluorescence microscope and sequentially sub-cultured into 24-well, 6-well and 10-cm culture dishes.

2.2.3.7 Protein extraction and western-blot analysis

Proteins were extracted from cell cultures using lysis buffer containing Tris (50 mM), sodium chloride (150 mM), NP-40 (1% m/v), glycerol (10% v/v), magnesium chloride (10 mM), sodium orthovanadate (10 mM) and protease inhibitor cocktail (cOmplete, Roche). Protein concentrations were determined using the Bradford protein assay (Bio-Rad). Proteins were fractioned using 4–12% acrylamide NuPAGE Novex Bis-Tris gels (Invitrogen) and transferred onto PVDF membranes (Millipore) using the XCell SureLock Blot module

(Invitrogen). The following antibodies were used for western blot: anti-zebrafish BMP2b rabbit IgG conjugate (AnaSpec; 1:500 dilution), anti avian ß-actin mouse IgG conjugate (Santa Cruz Biotechnology; 1:500 dilution), anti-rabbit IgG peroxidase conjugate (Sigma–Aldrich; 1:30.000 dilution) and anti-mouse IgG peroxidase conjugate (Sigma–Aldrich; 1:30.000 dilution). Chemiluminescent signals were detected using the Western Lightning ECL kit (PerkinElmer) and Hyperfilm ECL (Amersham, GE Healthcare) then quantified through densitometry analysis.

2.2.4 Results and Discussion

2.2.4.1 Expression of miR20a is up-regulated during in vitro mineralization in fish

Levels of miR20a expression and ECM mineralization were determined by qPCR and von Kossa staining, respectively, in confluent cultures of gilthead seabream ABSa15 cells exposed to mineralogenic cocktail for 4 weeks or left untreated (Fig. 2.2.1).

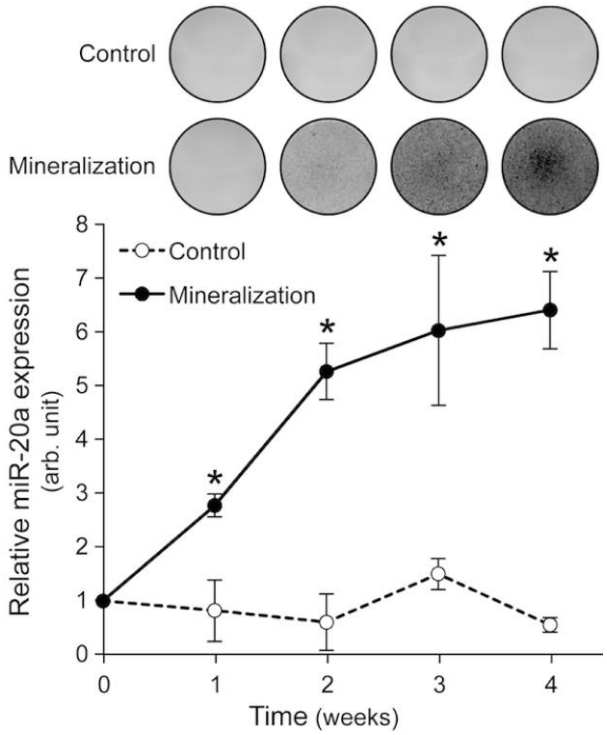


Figure 2.2.1. Relative expression of miR20a in mineralizing gilthead seabream ABSa15 cells. Extracellular matrix mineralization was induced in confluent cultures by supplementing medium with 50 µg/ml of L-ascorbic acid, 10mM b-glycerophosphate and 4mM $CaCl_2$. Control cultures were left untreated. Expression of miR20a was determined by qPCR and normalized using U6 small RNA expression. Representative pictures of von Kossa-stained cultures are presented above qPCR data. Asterisks (*) indicate values statistically different from respective control at specific time of mineralization ($n \geq 3$; Student's t-test, $p < 0.01$).

While expression of miR20a remained basal and constant in control non-mineralizing cultures, it was strongly up-regulated (up to 6-fold) during the first two weeks cultured in

mineralogenic medium, a period corresponding to the onset of in vitro mineralization, then remained constant as ECM mineralization progressed in the following 2 weeks, suggesting a role for miR20a in mechanisms of cell differentiation towards a phenotype of ECM mineralization.

2.2.4.2 Overexpression of miR20a decreases ECM mineralization and the activity of BMP canonical pathway

To further study this role, clones of ABSa15 cells overexpressing miR20a were developed through the stable transfection of pcDNA6.2-GW/EmGFP-miR20a construction. Three clones, homogeneously expressing GFP and therefore miR20a (data not shown), were isolated and overexpression of miR20a was confirmed by qPCR (6.3, 11.9 and 10.9-fold increase in clones 1, 2 and 3, respectively; Fig. 2.2.2A).

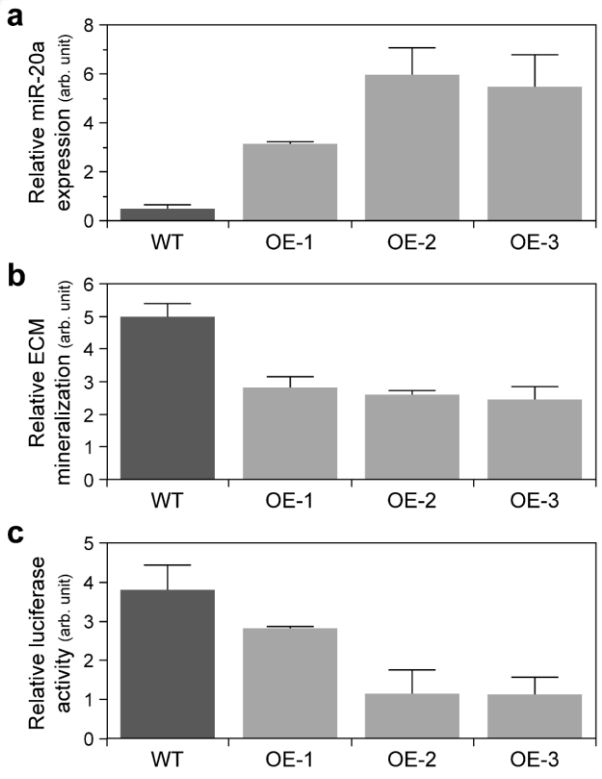


Figure 2.2.2. Effect of miR20a overexpression in gilthead seabream ABSa15 cells. Relative expression of miR20a **(a)** ECM mineralization **(b)** and reporter gene analysis of the canonical BMP signaling pathway **(c)** in wild-type cells (WT) and clones overexpressing miR20a (OE-1, -2 and -3). Expression of miR20a was determined in confluent cultures by qPCR and normalized using U6 small RNA expression. Mineral deposition was revealed after 2 weeks by von Kossa staining and evaluated by densitometry analysis. Reporter gene analysis was performed in cells transfected with BRE-Luc vector containing BMP-responsive elements upstream of luciferase gene. Relative luciferase activity was calculated as the ratio of firefly and renilla luciferase activities (FLuc/RLuc). All values in OE clones were statistically different from values in WT cells ($n \geq 3$; one-way ANOVA, $p < 0.05$).

Overexpressing (OE) clones were exposed to mineralogenic cocktail for 4 weeks and mineral deposition was evaluated once a week. Onset of ECM mineralization occurred after 2 weeks of exposure; at that time mineral deposition was significantly reduced by 44, 48 and 51% in clones 1, 2 and 3, respectively (Fig. 2.2.2B). Comparative analysis of these results

with those presented in Fig. 2.2.1 suggests that miR20a could have a specific role in early cell differentiation, i.e., from 0 to 2 weeks of mineralogenic treatment, progressively inhibiting this process and allowing mineralogenic mechanisms to occur. To support this hypothesis, ECM mineralization in cells overexpressing miR20a was delayed but not impaired. In fact, at later stages of ECM mineralization (3 and 4 weeks), mineral deposition remained lower in OE clones than in wild-type cells, but differences were not as accentuated (results not shown), indicating a partial recovery.

It has been recently reported that bone morphogenetic protein (BMP) pathway is targeted by miR20a (Brock et al., 2009); BMP pathway is central to osteogenesis, promoting osteoblast differentiation and ECM mineralization in fish (Laizé et al., 2009; Tiago et al., 2008; Tiago et al., 2011; Rafael et al., 2006) and mammals (Rosen, 2009; Laizé et al., 2009), and its repression would certainly impair ECM mineralization in ABSa15 cell line. To test this hypothesis, wild-type ABSa15 cells and clones OE-1, -2 and -3, were transfected with the BRE-Luc vector, a construct recently developed to investigate the activation of BMP pathway and where BMP responsive elements (BRE) control the expression of firefly luciferase gene (Korchynskyi and ten Dijke, 2002). Luciferase activity was significantly reduced in OE clones (Fig. 2.2.2C), indicating that BMP canonical pathway was affected upon overexpression of miR20a.

Expression of miR20a was also silenced in ABSa15 cells using a construct where a siRNA against miR20a was cloned into pcDNA6.2-GW/EmGFP vector. Two cell clones displaying reduced miR20a expression were isolated and treated for mineralization. Surprisingly, no significant changes in ECM mineralization were observed in these clones versus wild-type ABSa15 cells (data not shown), suggesting that compensatory mechanisms may exist.

2.2.4.3 BMP2 transcripts contain an evolutionary conserved binding site for miR20a

To further investigate the anti-mineralogenic action of miR20a through BMP pathway, the 3'UTR of gilthead seabream BMP2 transcript was analyzed in silico for the presence of miRNA binding sites. A search in GenBank sequence database using on-site blast facilities identified 2 transcript variants different in the length of their 3'UTR (GenBank accession numbers AY500244 and JF261172). A canonical polyadenylation signal (AAUAAA) was identified in both transcripts 16–24 nt upstream of poly(A) tail (Fig. 2.2.3A). Approximately

50% of mammalian protein-coding genes have more than one polyadenylation signal and can code for transcripts that differ in their 3'UTR (Yan and Marr, 2005). Since 3'UTRs contain binding sites for proteins that regulate mRNA stability (Huntzinger and Izaurralde, 2011) and for miRNAs that regulate mRNA translation (Legendre et al., 2006), alternative polyadenylation has been associated with post-transcriptional regulation of genes. Thus the 3'UTR region of the long variant of gilthead seabream BMP2 transcript was searched for miRNA binding sites using PITA algorithm (genie.weizmann.ac.il) and respective on-site miRNA database. Since PITA database only contained mammalian miRNA sequences, the conservation of predicted miRNAs (from mammals to zebrafish) was assessed using miRBase (mirbase.org). A binding site for miR20a was predicted with a ΔΔG score of -10.48 J/mole (sites with a score below -10 J/mole are likely to be functional if miRNA is endogenously expressed (Kertesz et al., 2007). BMP2 transcript was further analyzed using TargetScanFish release 6.2 (targetscan.org/fish_62), an online tool recently developed to search miRNAs binding sites in zebrafish sequences, and the presence of a miR20a binding site was confirmed in BMP2 3'UTR.

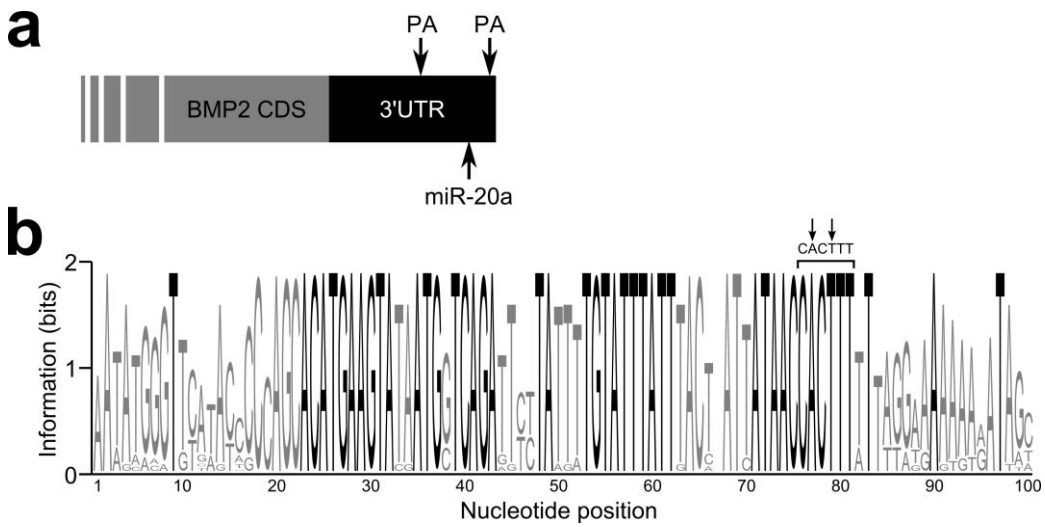


Figure 2.2.3. Prediction of miR20a binding sites in the 3' untranslated region (UTR) of BMP2 transcripts using PITA algorithm and TargetScanFish release 6.2. **(a)** schematic representation of the 3'UTR of gilthead seabream BMP2 transcript where polyadenylation signals (PA) and miR20a binding site are indicated. **(b)** Sequence logos of the 3'UTR of mammalian, sauropsidian, amphibian and fish BMP2 transcripts; miR20a seed region (CACTTT) is indicated on top of the logo and arrows indicate nucleotides mutated for functional analysis of miR20a binding; overall height of each letter corresponds to level of nucleotide conservation among species at that position; black letters indicate 100% conservation.

Although TargetScanFish analysis also predicted the presence of binding sites for other members of the miR17 family, which share similar seed regions, low binding energies were calculated by PITA for these binding sites, suggesting that they are less likely to bind miRNA binding site than miR20a.

The conservation of miR20a binding site in BMP2 transcripts throughout vertebrate evolution, a critical feature in miRNA binding predictions (Massirer and Pasquinelli, 2006; Kuhn et al., 2008), was investigated using BMP2-related sequences available in GenBank database. Nineteen 3'UTR sequences of mammalian (10), sauropsidian (2), amphibian (3) and bony fish (4) BMP2 transcripts were collected, aligned using ClustalW (align.genome.jp; Supplementary Fig. S2), then displayed as sequence logos using Weblogo (weblogo.berkeley.edu). A remarkable conservation of putative miR20a binding site, in particular the seed region, was observed (Fig. 2.2.3B), further evidencing the probable post-transcriptional regulation of BMP2 by miR20a. A second seed region for miRNAs of the miR17 family (including miR20a) was identified in the 3'UTR of mammalian, birds and amphibians BMP2 transcripts; it was however absent in fish BMP2 transcripts. Although its $\Delta\Delta G$ score was low (-6.76 J/mole in human sequence), which may indicate a false positive, future studies should aim at determining whether any miRNA of the miR17/92 cluster, in particular miR20a, bind to this tetrapod-specific site and whether post-transcriptional regulation of BMP2 transcripts has evolved throughout vertebrate evolution towards a tighter control by miR17 family.

Following the report by Brock et al. (Brock et al., 2009) evidencing the presence of a binding site for miR20a in the 3'UTR of human BMPR2 transcript, GenBank database was searched for vertebrate BMPR2-related sequences. Fourteen sequences were collected (mammals (8), sauropsids (3), and bony fish (3)) and aligned using ClustalW. A remarkable conservation of miR20a binding site was observed, in particular the seed region (Supplementary Fig. S3). Although both PITA and TargetScanFish returned low scores for the binding of miR20a to the miRNA site identified in zebrafish BMPR2 transcript, the remarkable conservation of the seed region suggests that this site could be functional in fish. Although this remains to be demonstrated, we propose that BMP signaling pathway and downstream processes are regulated by miR20a through its action on BMP2 but also on BMPR2 transcripts.

2.2.4.4 Seabream BMP2 transcript is post-transcriptionally regulated by miR20a

In silico prediction of miRNAs binding sites in BMP2 3'UTR sequences clearly indicated miR20a as a strong candidate for the post-transcriptional regulation BMP2. To validate this hypothesis, firefly luciferase activity was measured in extracts of HEK 293 cells transfected with reporter vector carrying 3'UTR region downstream of luciferase gene and, when appropriate, with expression vector carrying miR20a downstream of CMV promoter (Fig. 2.2.4).

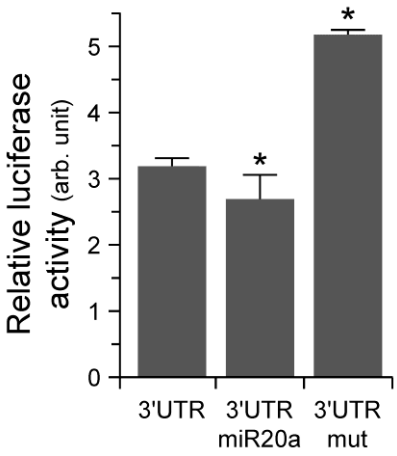


Figure 2.2.4. Interaction between miR20a and the 3'UTR of gilthead seabream BMP2 (SauBMP2) transcript. HEK 293 cells were transfected with pGL3 vector carrying the 3'UTR of SauBMP2 transcript (3'UTR) or the 3'UTR mutated for miR20a binding sites (3'UTR mut) downstream of luciferase gene. HEK 293 cells were also co-transfected with 3'UTR constructs and pcDNA6.2 carrying miR20a (3'UTR miR20a) downstream of CMV promoter. Relative luciferase activity was calculated as the ratio of firefly and renilla luciferase activities. Asterisks (*) indicate values statistically different from respective 3'UTR value ($n \geq 3$; one-way ANOVA, $p < 0.05$).

Polyadenylation signal upstream of miR20a binding site was mutated to avoid premature termination of the fusion transcript. Since miR20a expression was recently reported in HEK 293 cells (Tian et al., 2012), possible suppressive effects of endogenous miR20a in control conditions were considered in these experiments. While luciferase activity of BMP2 construction was slightly reduced (19%) upon miR20a overexpression, it was up-regulated (62%) upon mutation of the binding site, suggesting a pre-existing repression by endogenous miR20a and therefore a post-transcriptional regulation of BMP2 through miR20a binding site. Although a slight decrease of luciferase activity was observed upon overexpression of miR20a, we could not exclude that other members of the miR17 family, which share the same binding sites with miR20a and are expressed at similar levels in HEK 293 cells (Tian et al., 2012), may also regulate seabream BMP2 transcript. However, as stated previously, according to PITA analysis binding of other members of miR17 family to BMP2 transcript is less probable due to their association to lower binding energies. Nevertheless, to better understand the specific effect of miR20a, protein and transcript levels of gilthead

seabream BMP2 were determined in OE clones. In all OE clones, overexpression of miR20a resulted in a slight reduction of BMP2 production at 0 weeks (from 18% to 46%; Fig. 2.2.5A) and in a strong reduction after 2 weeks (from 80% to 100%; Fig. 2.2.5A).

A strong reduction (from 60% to 80%) of BMP2 transcript levels (long transcript) was also observed after 2 weeks in all OE clones (Fig. 2.2.5B). These data further demonstrated the regulation of BMP2 by miR20a and also suggested that miR20a action on BMP2 is probably related to mRNA degradation. Due to the lack of a suitable antibody to detect fish BMPR2, the action of miR20a on this protein in fish is still unknown. Furthermore, qPCR analysis of BMPR2 transcripts in WT ABSa15 cells and OE clones did not reveal any significant changes. Data collected in mammalian systems pointed towards the inhibition of protein translation (Brock et al., 2009), but whether this mechanism applies also in fish remains to be confirmed.

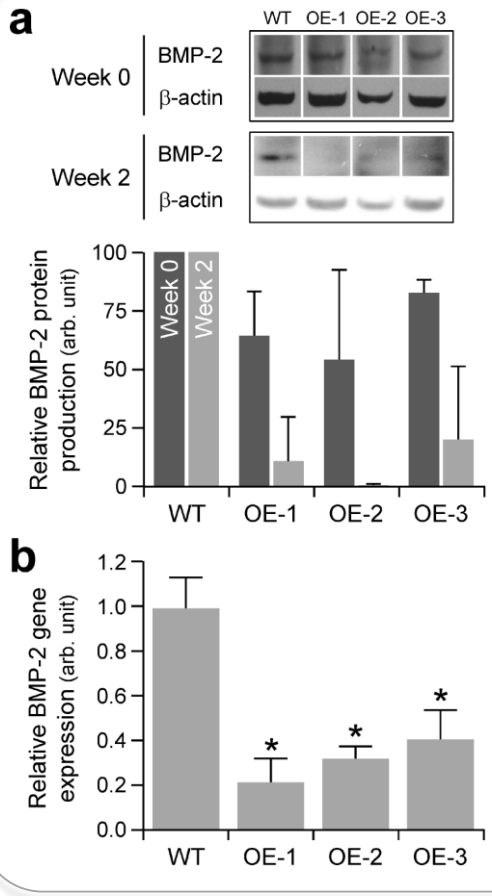


Figure 2.2.5. Levels of BMP2 protein production **(a)** and gene expression **(b)** in wild-type ABSa15 cells (WT) and clones overexpressing miR20a (OE-1, -2 and -3). Production of BMP2 protein was determined in cell cultures at time 0 and after 2 weeks of mineralization by densitometry analysis of western blot signals and normalized using ß-actin signals. Expression of BMP2 gene was determined by qPCR and normalized using RPL27a housekeeping gene expression (n.b. similar expression data were collected using 18S and ß-actin housekeeping genes; data not shown). Asterisks (*) indicate values statistically different from WT ($n \geq 3$; one-way ANOVA, $p < 0.05$).

The role of miR20a in bone formation is far from being understood. On one hand, the nature of known targets of miR20a, either identified through this study (BMP2) or identified in a previous study (BMPR2; Brock et al., 2009) suggest that miR20a can repress bone cell differentiation and ECM mineralization (Rosen, 2009). On the other hand, miR20a repression

of MAPK (Beveridge et al., 2009), a pathway that was shown to inhibit bone cell differentiation in mammals and fish (Laizé et al., 2009; Tiago et al., 2008; Tiago et al., 2011), suggests that miR20a could also promote bone cell differentiation in fish. Accordingly, miR20a was recently shown to induce osteogenic differentiation in human mesenchymal stem cells (hMSC) through repression of antagonists of BMP pathway BAMBI, CRIM1 and PPAR-γ (R Zhang et al., 2011). In contrast, the role of miR20a in bone formation was further investigated *in vivo* in a study using mouse knockout models for the miR17/92, a cluster of miRNAs in which miR20a is included (Ventura et al., 2008). While the development of homozygotic miR17/92 knockouts is severely compromised due to lethal cardiac and lung defects, heterozygous models showed significantly reduced trabecular and cortical bone formation, and impaired osteoblast differentiation (Zhou et al., 2014). Data collected within the scope of this work is in favor of miR20a inhibiting bone cell differentiation/mineralization through the repression of BMP pathway. This discrepancy could be related to distinct regulatory mechanisms in different cell systems: ABSa15 is a skeletal cell line established from calcified cartilage of branchial arches of a teleost fish; it displays gene expression patterns resembling those of chondrocyte-like cell types, including: (i) mild up-regulation of TNAP (tissue non-specific bone-related alkaline phosphatase), COL1A1 (type I collagen a1), SPARC (secreted protein acidic cysteine-rich; also known as osteonectin) and SOX9a (SRY (sex determining region Y)-box 9a) earlier in differentiation and down-regulation later during mineralization; (ii) strong up-regulation of MGP (matrix Gla protein) and SPP1 (secreted phosphoprotein 1; also known as osteopontin) from non-differentiated to mineralized cells; and (iii) absence of osteocalcin expression in all stages (Supplementary Fig. S4). Supporting the dual effect of miR20a, a recent study showed that miR17, a miRNA that belongs to miR17/92 cluster and shares the same ‘‘seed’’ and predicted targets with miR20a, could either inhibit or promote osteogenic differentiation in human periodontal ligament tissue stem cells depending on whether these where collected from healthy donors or patients suffering from inflammatory process, respectively (Liu et al., 2011). Interestingly, this opposite effect was associated with the differential expression of SMAD ubiquitin regulatory factor one (SMURF1), a regulator of BMP pathway and a direct target of miR17.

In another study, Brock and colleagues evidenced in human and mouse the post-transcriptional regulation of the cell surface receptor BMPR2 by miR20a and proposed that the up-regulation of miRNA expression may be a key feature in the development of

pulmonary arterial hypertension (PAH), through the action of BMP signaling pathway on endothelial and smooth muscle cell differentiation and matrix formation (Brock et al., 2012; Brock et al., 2009). These authors showed that antagonizing miR20a using an antagomiR restored BMPR2 mRNA/protein levels and a functional BMP signaling in a mouse model of hypoxia-induced PAH (Brock et al., 2012). Consequently disease development was reduced and pulmonary arterial haemodynamics were improved in antagomiR20a treated animals (Brock et al., 2012). Smad5, an intermediate of BMP signaling pathway, was also proposed to be targeted by miR20a but this hypothesis was never confirmed either by luciferase reporter assays or western blot analysis (Brock et al., 2012). Interestingly, miR17/92 cluster, in particular miR20a, has been shown to repress type II transforming growth factor ß receptor (TGFBR2; Li et al., 2012), which is involved in osteogenesis (and chondrogenesis) through the action of TGFß signaling on osteoblast recruitment and proliferation, and matrix formation (G Chen et al., 2012). Since BMPR2 and TGFBR2 belong to the same cell surface receptor family and activate similar transduction pathways, it will be interesting to address in future studies the role of their post-transcriptional regulation by miR20a during osteogenesis (and in a more general manner during skeletogenesis) and whether a dysregulation of this mechanism could lead to bone/skeletal diseases.

2.2.4.4 Overexpression of miR20a up-regulates the expression of the matrix Gla protein, a calcification inhibitor

In order to better understand the underlying mechanisms of miR20a inhibitory role on ECM mineralization of ABSa15 cells, expression of several markers of bone cell differentiation/mineralization was investigated in wild-type cells and in clones over-expressing miR20a. Initially, the expression of each bone-related marker was investigated during ECM mineralization of ABSa15 cells, and as mentioned before, this analysis suggested a possible association of ABSa15 cells to a chondrocytic lineage (Supplementary Fig. S4). After 2 weeks of treatment, when overexpression of miR20a strongly inhibited mineral deposition, expression levels of TNAP, COL1A1 and MGP were significantly up-regulated in all three OE clones, ranging between 3.7–6.4-fold, 3.1–5.5-fold and 2.1–5.1-fold, respectively (Fig. 2.2.6), while other bone-related genes did not reach significant differences or remained undetected (data not shown).

The strong up-regulation of MGP, a well known inhibitor of arterial calcification in mammals (Schurgers et al., 2008) which has been also associated with ECM mineralization

in fish (Gavaia et al., 2006; Pinto et al., 2003; Simes et al., 2003; Pombinho et al., 2004), could explain miR20a inhibitory effects in ABSa15 cell mineralization. In mammalian systems, MGP was demonstrated to bind BMP2 and prevent its association with BMPR2, which is necessary for the activation of BMP pathway and consequent stimulation of bone formation (Zebboudj et al., 2002; Wallin et al., 2000). Furthermore, in calcifying vascular cells this mechanism involved a feedback control regulation, where MGP expression levels appear to be negatively correlated with BMP2 availability (Zebboudj et al., 2003). Therefore, an up-regulation of MGP in ABSa15 OE clones is likely to increase its binding to BMP2 and thus contribute to block BMP pathway, enhancing the effect of post-transcriptional regulation of BMP2 by miR20a.

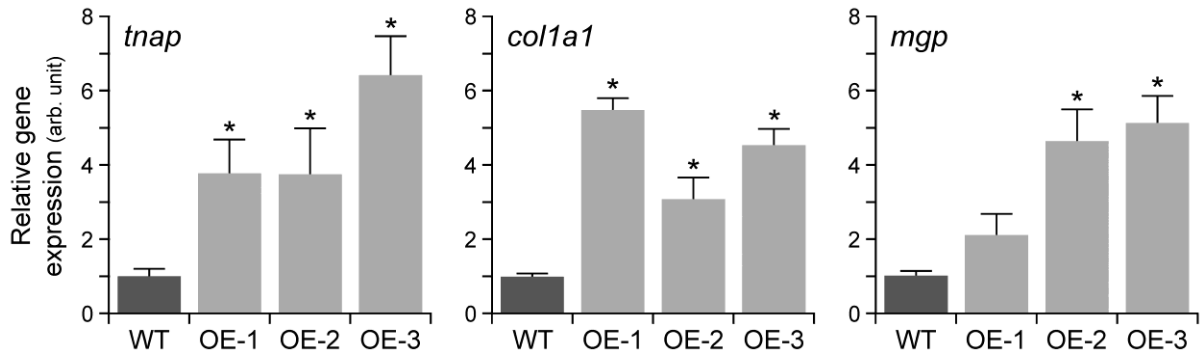


Figure 2.2.6. Levels of alkaline phosphatase (TNAP), type I collagen a1 (COL1A1) and matrix Gla protein (MGP) gene expression in mineralizing wild-type ABSa15 cells (WT) and clones overexpressing miR20a (OE-1, -2 and -3). Gene expression was determined after 2 weeks of mineralization by qPCR and normalized with RPL27a housekeeping gene expression (n.b. similar expression data were collected using 18S and ß-actin housekeeping genes; data not shown). Asterisks (*) indicate values statistically different from WT ($n \geq 3$; one-way ANOVA, $p < 0.05$).

Regarding TNAP and COL1A1, these are promoters of ECM formation and mineralization (Murshed et al., 2005) and it is therefore difficult to explain the anti-mineralogenic effect of miR20a through their up-regulation. Since data available on literature regarding regulation of TNAP and COL1A1 by miR17/92 is still contradictory (Zhou et al., 2014), this effect should be addressed in future studies.

2.2.5 Conclusions

We present here novel data (i) describing the up-regulation of miR20a during ECM mineralization of a fish mineralogenic cell line (ABSa15), (ii) evidencing the anti-mineralogenic effect of miR20a in this cell line, (iii) showing inhibition of the BMP pathway

by miR20a, (iv) identifying binding sites for miR20a in the 3'UTR of gilthead seabream BMP2 and zebrafish BMPR2 transcripts that were conserved in vertebrate and (v) demonstrating the post-transcriptional regulation of BMP2 by miR20a (binding site in BMPR2 transcript may also be active but this remains to be demonstrated). We propose that low levels of expression of miR20a in undifferentiated cells may account for a higher activity of BMP signaling and consequent osteogenic differentiation. Then, in the course of ECM mineralization, miR20a becomes more expressed to inhibit this process through BMP2 (and possibly also BMPR2) regulation (Fig. 2.2.7).

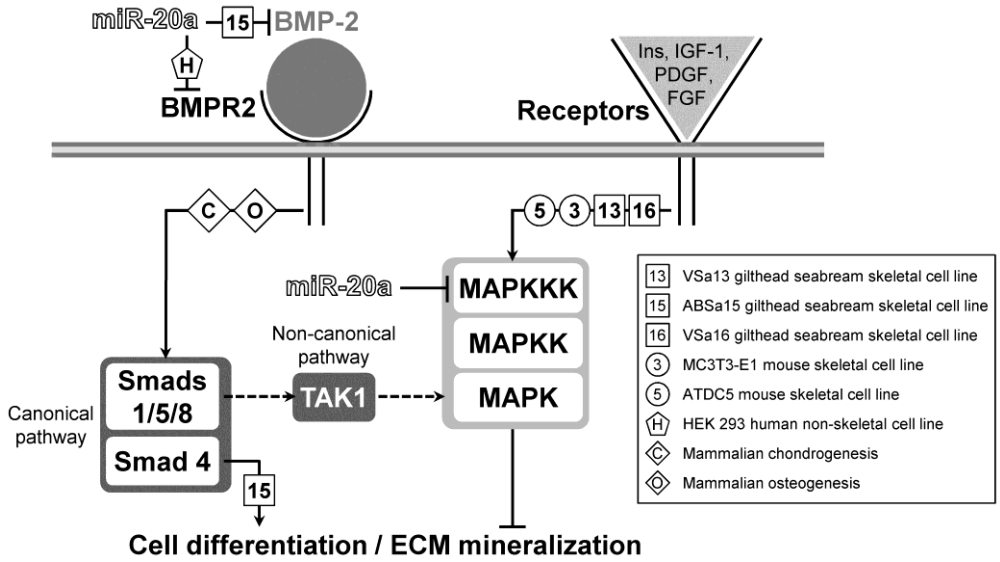


Figure 2.2.7. Putative mechanisms of action for miR20a anti-mineralogenic effect. Arrows and intersected lines indicate activation and repression, respectively. Solid arrows indicate pathways likely to be activated during differentiation of ABSa15 cells. Dashed arrows indicate pathways most probably not activated in ABSa15 cells. Smads 1, 5, 8 and 4 are intermediates of the BMP canonical pathway. TAK1 (TGFß-activated kinase 1) is an intermediate of the BMP non-canonical pathway. MAPK, MAPKK and MAPKKK are intermediates of the mitogen-activated protein kinase (MAPK) pathway.

Alternative mechanisms of action, such as activation of MAPK pathway either directly by miR20a or indirectly through non-canonical BMP pathway, cannot be excluded but remain to be demonstrated in ABSa15 cells. Furthermore, effect on MGP suggests that this protein is likely to play a role in the inhibitory mechanism observed. Results obtained from previous studies combined with data hereby demonstrated, suggest that miR20a preferentially targets BMP pathway to promote (hMSC) or inhibit (ABSa15 cells) osteogenic differentiation.

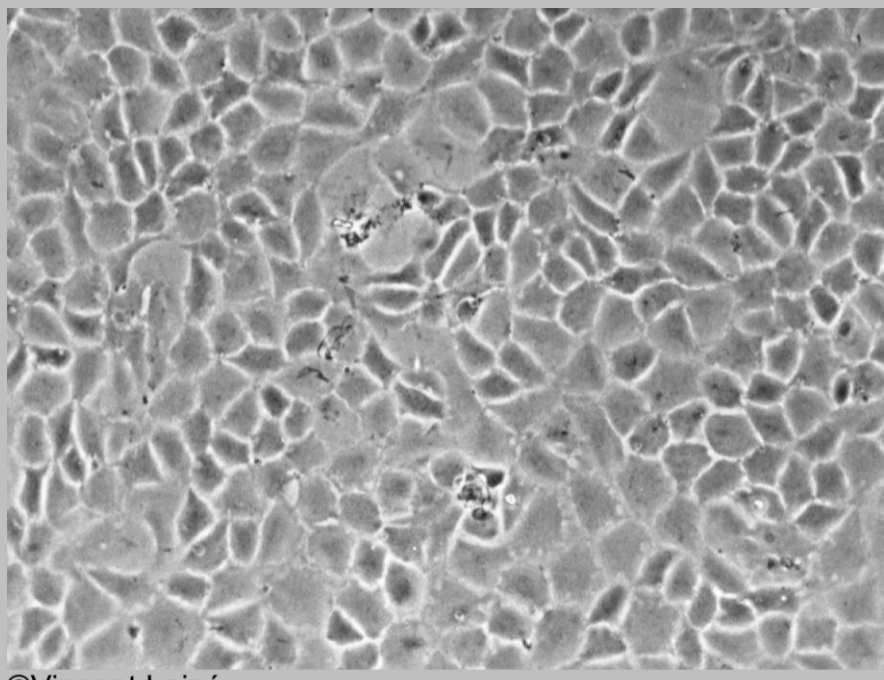
2.2.6 Acknowledgements

Authors are grateful to Dr. Peter ten Dijke (Leiden University Medical Center, Leiden, The Netherlands) for kindly providing the BMP-responsive luciferase reporter vector (BRE-Luc). This work was supported by grants from the Calouste Gulbenkian Foundation (program "Na Fronteira das Ciências da Vida"; to DMT) and the Centre of Marine Sciences (to DMT, VL and MLC). DMT, VPR and CLM were the recipients of post-doctoral (SFRH/BPD/45034/2008) and doctoral (SFRH/BD/41392/2007 and SFRH/BD/39964/2007) fellowships, respectively, from the Portuguese Foundation for Science and Technology (FCT). This work was also partially funded by European ASSEMBLE project (FP7-227799).

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.abb.2013.12.009

Chapter 3

CHARACTERIZATION OF BMP2-MGP INTERACTION



©Vincent Laizé

Preamble

MGP was initially described as a contaminant of BMP2 purification from demineralized bone matrix. During the last decades several studies proposed that both proteins may interact and that this interaction may affect the osteoinductive capacity of BMP2 and therefore represent another level on BMP2 signaling modulation. Our initial idea was to further investigate BMP2-MGP interaction in a fish system and to identify residues/domains responsible for the binding of the two proteins, using a luciferase reporter construct (BRE-Luc), previously validated in mammalian systems. While we could demonstrate the functionality of the BRE-Luc construction in gilthead seabream skeletal cell line and its suitability to study the transduction of BMP2 signal, we failed to evidence a negative effect of MGP on BMP2 signaling activation. Data presented here are not conclusive of BMP2-MGP interaction and will not be published.

CHAPTER 3. CHARACTERIZATION OF BMP2-MGP INTERACTION

3.1 Identification of the domains involved in bone morphogenetic protein (BMP) 2 and matrix Gla protein (MGP) interaction

Cátia L. Marques

M. Leonor Cancela

Vincent Laizé

3.1.1 Abstract

Bone-related diseases are considered a major health problem, affecting millions of people worldwide. BMPs are secreted proteins that act as bone inducers and have proven to be essential for tissue mineralization. Among the BMP family members, BMP2 has demonstrated remarkable osteogenic properties by inducing osteoblasts differentiation. However, mechanisms underlying BMP2 regulation are far from being understood. MGP is a small matrix protein initially purified from bone as a contaminant of BMP2. The two proteins were found to be physically associated and the calcification inhibitory role of MGP is thought to be related with its interaction with BMP2. Although evidences point to the involvement of MGP Gla residues in the binding, no conclusive results have been reported. In this work we have used the BRE-Luc system as a tool to characterize residues involved in BMP2-MGP binding. In our fish system this approach, although functional, was shown to be unsuitable to study protein-protein interactions since the effect observed with the negative controls was similar to the effect observed when using MGP, indicating an unspecific response. In future works we propose to use an alternative and direct method, i.e. protein-fragment complementation assay, to characterize residues involved in BMP2-MGP interaction.

3.1.2 Introduction

Bone morphogenetic proteins (BMPs) form the largest group of proteins in TGFβ superfamily of cytokines and growth factors (Katagiri and Tsukamoto, 2013) and to date more than 20 members have been described and characterized (Modica and Wolfrum, 2013). BMP2 is most probably the best studied member of this family due to its pleiotropic action and involvement in the regulation of various physiological processes, from embryonic patterning and organogenesis to bone metabolism (reviewed in Hogan, 1996). As for all BMP proteins, BMP2 is synthetized as a large precursor composed of a signal, a pro and a mature peptides (Shimasaki et al., 2004). After signal peptide removal, the pro and mature domains undergo dimerization and are sequentially processed into the dimeric biological form by members of the subtilisin-like convertase family, which recognize the optimal RXR/KR or the minimal RXXR sequences present just before the mature peptide (Constam and Robertson, 1999; de Caestecker, 2004). Once processed into the active mature form, BMP2 is secreted outside the cell into the extracellular compartment (Ducy and Karsenty, 2000). In order to exert its function the dimeric form of BMP2 binds to type I BMP receptors (BMPRI) present at the cell surface and recruit type II BMP receptors (BMPRII) that will phosphorylate the serine and threonine residues on type I receptors. Activated type I receptors are then responsible for the transduction of BMP signaling through mothers against decapentaplegic homologs (SMADs) intermediates to the nucleus, where regulation of target genes will occur (Balemans and Van Hul, 2002; Marcellini et al., 2012).

Due to its wide range of functions and critical role in animal physiology, BMP2 protein is regulated at various levels and several modulators of BMP2 activity have been described (Yanagita, 2005), including the matrix Gla protein (MGP; Zebboudj et al., 2002). MGP is a small extracellular matrix protein, member of the vitamin K-dependent family, also known as Gla (γ-carboxyglutamate) proteins (Cancela et al., 2014). Members of this family undergo a post-translational modification, where glutamate (Glu) residues are converted to Gla residues by the γ-glutamyl-carboxylase (Schurgers et al., 2013). Price and Williamson (1985) reported the co-purification of MGP and BMP2 from bone demineralized extracts, suggesting a physical association between the two proteins. MGP is a known calcification inhibitor and, although mechanisms of action are not fully understood at the molecular level, it has been proposed that MGP may prevent soft tissue calcification by binding to BMP2 and inhibiting BMP2-related induction of osteogenic differentiation and mineralization (Boström et al., 2001; Schurgers et al., 2008). BMP2-MGP interaction has been demonstrated by co-

immunoprecipitation studies (Zebboudj et al., 2002) and although γ-carboxylated residues in MGP have been proposed as important mediators in MGP-BMP2 interaction (Sweatt et al., 2003), conclusive evidences are still missing, mostly due to the lack of a suitable system.

The aim of this study was to characterize BMP2-MGP interaction by identifying contributing residues/domains in both proteins. BMP responsive element reporter assay (BRE-Luc; Korchynskyi and ten Dijke, 2002) was used to evaluate the effect of BMP2-MGP interaction on BMP2-specific activation of BMP signaling using mineralogenic fish-derived cells. In this system, previously used in mammalian cells to demonstrate and characterize BMP4-MGP interaction (Yao et al., 2008), BMP2 protein should, after expression and secretion, bind the corresponding receptors and induce luciferase activity through the BMP signaling cascade (Fig. 3.1A). When added to the system, MGP should bind BMP2 and consequently repress the BMP2-induced luciferase activity (Fig. 3.1B), allowing binding capacity assessment. Residues and/or domains involved in BMP2-MGP binding will be also identified by introducing mutations in the specific sequences coding for particular residues on the MGP protein (e.g. Gla residues from the Gla domain and Serine residues from the phosphorylation domain).

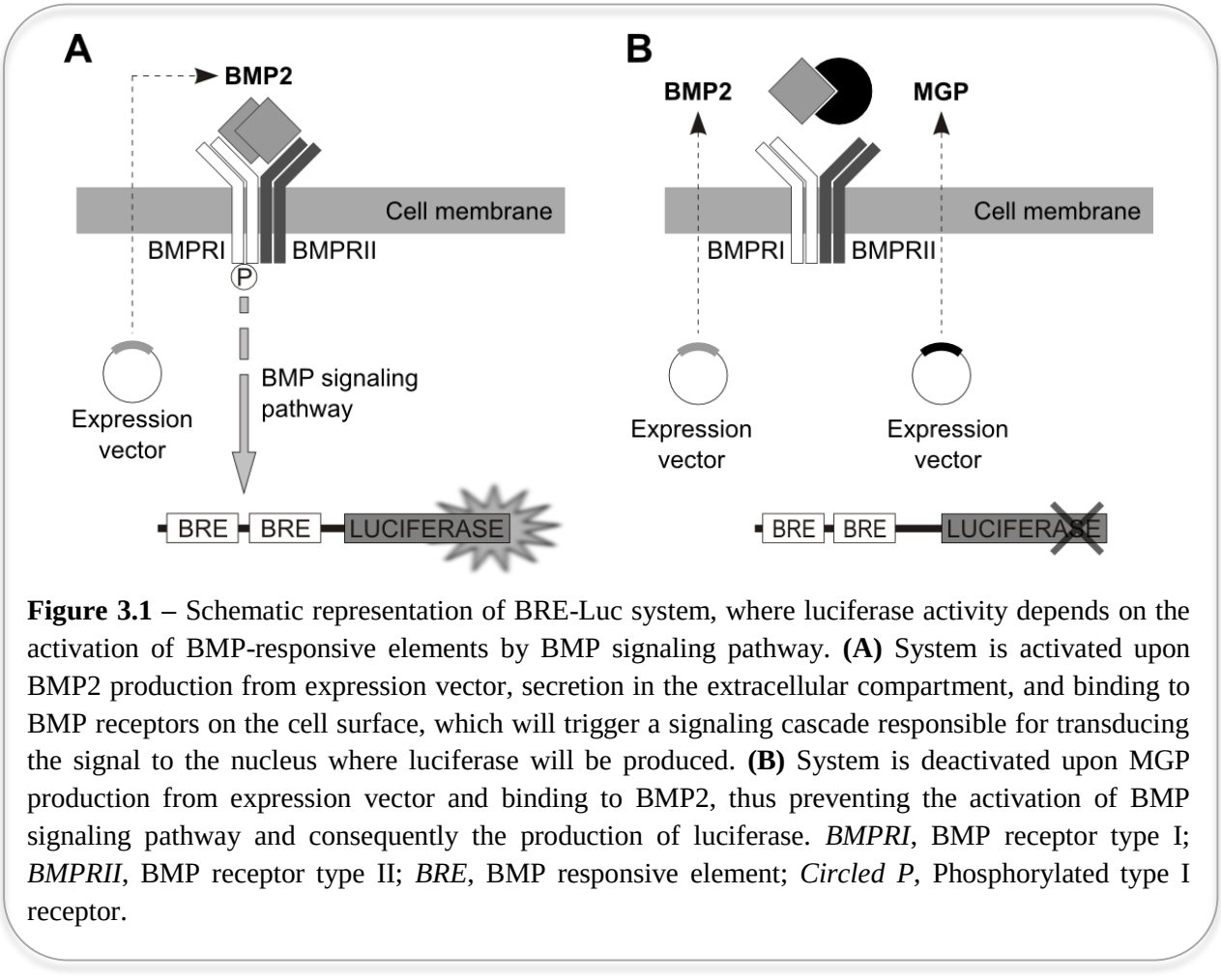


Figure 3.1 – Schematic representation of BRE-Luc system, where luciferase activity depends on the activation of BMP-responsive elements by BMP signaling pathway. **(A)** System is activated upon BMP2 production from expression vector, secretion in the extracellular compartment, and binding to BMP receptors on the cell surface, which will trigger a signaling cascade responsible for transducing the signal to the nucleus where luciferase will be produced. **(B)** System is deactivated upon MGP production from expression vector and binding to BMP2, thus preventing the activation of BMP signaling pathway and consequently the production of luciferase. *BMPRI*, BMP receptor type I; *BMPRII*, BMP receptor type II; *BRE*, BMP responsive element; *Circled P*, Phosphorylated type I receptor.

3.1.3 Materials and methods

3.1.3.1 Cell culture maintenance

ABSa15 cells (ECACC catalogue no. 13112201; Tiago et al., 2014) were cultured in Dulbecco's Modified Eagle Medium (DMEM; Life technologies), supplemented with 10% of fetal bovine serum (FBS; Sigma-Aldrich), 1% of penicillin-streptomycin (Life technologies), 1% of L-glutamine (Life technologies) and 0.2% of fungizone (Life technologies) and maintained at 33ºC in a 10% CO_2 humidified atmosphere.

3.1.3.2 Transfection assays

The day before the transfection ABSa15 cells were seeded in 24-well plate, at a density of 4×10^4 cells per well, and cultured for an additional 16 h in the same conditions. Cultures at 60-80% of confluence were transfected with BRE-Luc vector (BMP-responsive element driving the expression of the firefly luciferase; Korchynskyi and ten Dijke, 2002) using 1 µl of X-tremeGENE HP DNA transfection reagent (Roche). When appropriate, expression vectors (pcDNA3 backbone) containing the coding sequence of gilthead seabream BMP2, MGP, insulin-like growth factor 1 (IGF1) and osteocalcin 1 (OC1) under the control of CMV promoter, were co-transfected with BRE-Luc vector (pcDNA3-IGF1 and pcDNA3-OC1 expression vectors were constructed in M.L. Cancela's laboratory by D. Tiago and D. Molha, respectively). pRL-SV40 vector (Promega), expressing *Renilla* luciferase (R-luc) under the control of CMV promoter, was used in all the transfections (200 ng) to normalize the firefly luciferase (F-luc) activity and empty pcDNA3 vector was used to keep the amount of DNA constant during cell transfections. After 48 h, cells were lysed and luciferase activities were measured in a BioTek Synergy 4 plate reader using Dual-Luciferase Reporter Assay system (Promega). Relative luciferase activity was determined from the ratio F-Luc/R-Luc and results presented as the fold change over BRE-Luc vector (for BMP2) or BRE-Luc+BMP2 (for MGP, IGF1 and BGP).

3.1.4 Results and Discussion

In order to validate the BRE-Luc reporter system in fish (it has been developed for mammalian systems), gilthead seabream ABSa15 cells were co-transfected with BRE-Luc

vector and pcDNA3-BMP2 vector expressing gilthead seabream BMP2 (Fig. 3.2A). Seabream BMP2 was able to induce luciferase activity demonstrating that it could bind to BMPR and activate BMP signaling pathway but also indicating a conservation of both the BMP signaling pathway and the BMP responsive elements throughout evolution. It also validated the use of our fish system to study BMP2-MGP interaction. To test the responsiveness of the system and determine the optimal amount of BRE-Luc and pcDNA3-BMP2 vectors to use in the experiments, different quantities of the two vectors were transfected (Fig. 3.2A).

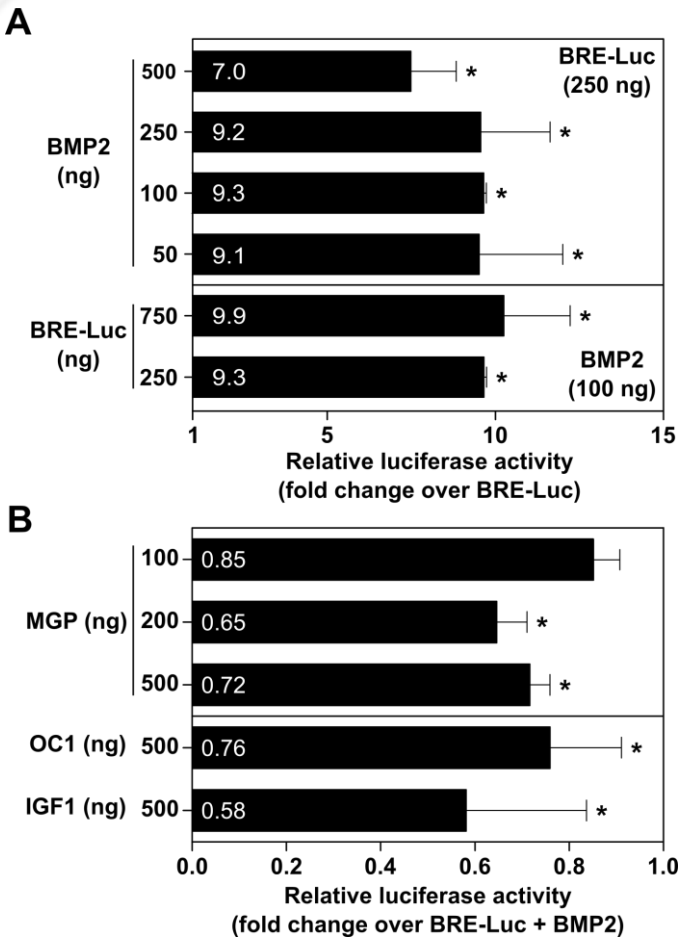


Figure 3.2 – Relative luciferase activity upon co-transfection of BRE-Luc reporter vector with BMP2 (**A**) or BMP2/MGP (**B**). ABSa15 cells were first co-transfected with different amounts of BRE-Luc and pcDNA3-BMP2 vectors to validate BRE-Luc system in a fish environment and determine optimal vector concentrations. BMP2-MGP interaction and inhibition of BMP2 signaling pathway by MGP was evaluated through co-transfection of pcDNA3-BMP2 and pcDNA3-MGP expression vectors. Vectors expressing insulin-like growth factor1 (pcDNA3-IGF1) and osteocalcin 1 (pcDNA3-OC1) were used as negative controls. *Numbers* inside the bars indicate fold changes over BRE-Luc vector (panel A) or over BRE-Luc + pcDNA3-BMP2 vectors (panel B). *Asterisk* indicate values significantly different from control (one-way ANOVA followed by Dunnett's post-test; $P<0.05$).

A 9-fold induction of luciferase activity was observed, independently of the amounts of pcDNA3-BMP2 and BRE-Luc vectors used.

Since it is always preferable to transfect cells with the lowest possible amounts of DNA and because lowest variations in luciferase values were achieved with these amounts, 100 ng of pcDNA3-BMP2 vector and 250 ng of BRE-Luc vector were used thereafter.

The capacity of MGP to inhibit BMP2 activity was evaluated through co-transfection of pcDNA3-BMP2 and BRE-Luc vectors with different amounts of pcDNA3-MGP expression vector (Fig. 3.2B). A slight decrease of luciferase activity was observed upon co-transfection with pcDNA3-MGP expressing vector, possibly indicating an inhibition by MGP of BMP2-induced luciferase activity. However, this inhibition was somehow weak in light of the results reported by Yao and co-workers (Yao et al., 2008) and this could be a consequence of an insufficient γ-carboxylation of MGP in ABSa15 cells (although γ-glutamyl carboxylase is produced endogenously in these cells we are not sure that the levels are sufficient to induce a correct gamma carboxylation when MGP is overexpressed) or a lack of responsiveness by our cell system (forced protein production of BMP2 and MGP under the control of CMV constitutive promoter can overload the cellular machinery and affect cell fitness and consequent luciferase expression). Following this reasoning, it is also possible that decrease in luciferase activity was not MGP specific but resulted from altered cell fitness upon delivery of massive DNA content (approximately 1 μg) into the ABSa15 cells. To test this hypothesis, ABSa15 cells were co-transfected with 500 ng of pcDNA3 vectors expressing IGF1 or OC1 (Fig. 3.2B), two proteins not reported to interact with BMP2. Both proteins slightly decreased luciferase activity to levels similar to those observed upon MGP co-transfection, confirming that MGP had probably no direct effect on induction of BMP signaling pathway in our system. Given the published data on BMP2-MGP interaction (Zebboudj et al., 2002) and on the validation of Bre-Luc approach to study protein-protein interaction (Yao et al., 2008), we have to question the functionality of our fish system. While BMP2 expression from pcDNA3 vector and production by cell machinery has been demonstrated by the strong increase of luciferase activity (therefore confirming the functionality of CMV promoter in fish cells), MGP expression and production has not been confirmed. Since the same cloning strategy and vector were used to prepare BMP2 and MGP constructs we assume that MGP protein was being overexpressed in ABSa15 cells, however this should be confirmed (e.g. by western blot or immunocytochemistry). Posttranslational modification of BMP2 (N-glycosylation) and MGP (phosphorylation/γ-carboxylation) are critical to protein function. While non-glycosylated BMP2 is still functional, although to a lower extent (Ruppert et al., 1996; P C Bessa et al., 2008), unphosphorylated or uncarboxylated MGP loses its functionality and fails to prevent soft tissue calcification as demonstrated by (Murshed et al., 2004). Although this should be confirmed (e.g. by evaluating the carboxylation of MGP overexpressed in ABSa15 cells, using Gla-specific

antibodies), we believe that over-produced MGP may be under- or uncarboxylated in ABSa15 cells and alternative cell systems with a higher capacity of γ-carboxylation (e.g. liver-derived cells) should be tested. Although it would increase the DNA load, ABSa15 cells could also be co-transfected with a vector expressing gilthead seabream γ-glutamyl carboxylase, in order to increase their carboxylation capacity. Finally, the use of purified BMP2 and MGP proteins instead of cDNA-expression vectors should also be considered, to overcome possible issues related with posttranslational modifications and DNA overload of the cells. Alternatives to the use of the BRE-Luc system are also available. For example, protein-fragment complementation assays (PCAs) have been described as powerful systems to study protein-protein interactions (Michnick et al., 2007; Remy and Michnick, 2006). In this approach, proteins of interest are fused with fragments of *Gaussia princeps* luciferase, which when brought to proximity through the association of the two interacting proteins, are able to reconstitute enzyme activity and produce a quantitative signal.

In conclusion, we were able to demonstrate BRE-Luc system functionality in ABSa15 cells and to confirm that seabream BMP2 is an effector of the BMP signaling pathway. Unfortunately, for reasons that remain to be identified, we failed to show that MGP interacts with BMP2 and is therefore a direct regulator of BMP signaling pathway. In the future, we propose to either modify the current cell system to improve posttranslational capacity, to use purified and fully functional proteins instead of overexpressing them or to use alternative methods (e.g. PCAs) to characterize BMP2-MGP interaction and identify residues contributing to this interaction.

3.1.5 Acknowledgments

We thank D. Tiago and D. Molha for the pcDNA3-IGF1 and pcDNA3-OC1 expression vectors, respectively and Dr. Peter ten Dijke (Leiden University Medical Center, Leiden, The Netherlands) for kindly providing the BMP-responsive luciferase reporter vector (BRE-Luc).

Chapter 4

Comparative Analysis of Fish BMP2, BMP4 and BMP16

©Paulo Gavaia

Preamble

BMPs have been divided into subfamilies according to their primary structure and function. Until recently, BMP2 and BMP4 were the founding and only members of BMP2/4 subfamily, however in 2009, a new member, named BMP16, was assigned to this subfamily based on its high homology with both BMP2 and BMP4. This chapter is divided in two parts that are based on manuscripts that address the evolutionary relationship of BMP2, BMP4 and BMP16 but also compare structural and expression data in an attempt to get insights into protein specificity and gene regulation. In the first manuscript, published in the Journal of Applied Ichthyology, taxonomic distribution was investigated in a particular taxonomic group (Teleostei) and spatiotemporal expression data were collected from the marine teleost Senegalese sole. In the second manuscript, submitted to Cellular and Molecular Life Sciences, molecular phylogeny was inferred from genomic data collected in the major vertebrate taxa, while structural data (from protein modeling), functional data (from BRE-Luc reporter assay), expression data (during development and in adult tissues) and regulatory data (by retinoic acid) were collected from the fish model zebrafish.

CHAPTER 4. COMPARATIVE ANALYSIS OF FISH BMP2, BMP4 AND BMP16 AT GENE AND PROTEIN LEVELS

4.1 Spatiotemporal expression and retinoic acid regulation of bone morphogenetic protein (BMP) 2, 4 and 16 in Senegalese sole

Cátia L. Marques
Ignacio Fernández
Joana Rosa
Michael N. Viegas
M. Leonor Cancela
Vincent Laizé

4.1.1 Abstract

Bone morphogenetic proteins (BMPs) are involved in various physiological processes from early life stages throughout adulthood. Specific characteristics of BMPs have been used to define different subfamilies and BMP2/4 subfamily (composed of BMP2 and BMP4) has been linked to osteogenesis and skeleton development. BMP16 was recently identified as a new member of the BMP2/4 subfamily and reported as a teleost fish-specific form. In this work, we collected a comprehensive set of ray-finned fish BMP2, BMP4 and BMP16 sequences and demonstrated, through its presence in Holostei, that BMP16 is not restricted to teleost fish genome. Comparative analysis of BMP2, BMP4 and BMP16 primary structures revealed that most of the residues required for protein stabilization, dimer formation, glycosylation and receptor binding are substantially conserved between the three proteins, suggesting that BMP16, BMP2 and BMP4 may share similar mechanisms of action. In contrast, comparative analysis of gene expression profiles during Senegalese sole development revealed differences in onset and extent of gene expression, indicating that BMP16, BMP2 and BMP4 may contribute to different developmental and physiological processes. High levels of transcripts in adult calcified tissues and the up-regulation of gene

expression by retinoic acid, a known regulator of skeletal development, suggests that BMP16 shares with BMP2 and BMP4 a role in bone metabolism and skeletal development. This study provides new insights into the taxonomic distribution and the spatiotemporal expression of BMP16 gene, and suggests that it may share structural and functional similarities with other members of the BMP2/4 protein subfamily.

4.1.2 Introduction

Bone morphogenetic proteins (BMPs) are members of the transforming growth factor beta (TGFβ) superfamily of multifunctional growth factors, initially described as bone inducers (Wozney et al., 1988). To date, more than 20 members of the BMP family have been identified and new functions, not related to bone formation, have been reported (Chen et al., 2004; Wu and Hill, 2009). BMPs are synthesized as long inactive precursors and processed, through sequential cleavage, by members of the subtilisine-like convertases (SCPs) family, at optimal (RXK/RR) and minimal (RXXR) sites (Heng et al., 2010; Akiyama et al., 2012). The resulting mature and biological active forms are capable of modulating various cellular processes through their binding to BMP receptors (BMPRs) (Canalis et al., 2003; X Wang et al., 2012). BMPs have been divided into subgroups according to their primary amino acid sequences, and the subfamily formed by BMP2 and BMP4, homologs of the fly decapentaplegic (DPP) protein, was the first to be identified and characterized (Wozney et al., 1988; Kingsley, 1994). Although both proteins were associated to mechanisms of bone and cartilage formation (Nishimura et al., 2008), BMP2 and BMP4 have also been implicated in several important embryonic processes (e.g. gastrulation, neural patterning and gut development, reviewed by Hogan, 1996). Most likely, due to their high degree of conservation throughout evolution, members of the BMP2/4 subfamily are interchangeable, e.g. the drosophila protein DPP can induce endochondral bone formation when introduced subcutaneously in mouse (Sampath et al., 1993) while the mammalian BMP4 is able to rescue the dorsal-ventral defects resulting from the lack of DPP in drosophila (Padgett et al., 1993). Also, mechanisms of regulation of BMP signaling by retinoic acid seem to be conserved (Sheng et al., 2010), evidencing the remarkable conservation, both in structure and function of members of this family. Recently a new member of this sub-family has been identified and named BMP16. It was reported to be teleost fish-specific and presumed to have appeared from a duplication event that occurred early in vertebrate lineage, and then to have been lost in tetrapods soon after their divergence

from ray-finned fish 450 million years ago (Feiner et al., 2009). BMP16 gene expression was also evaluated and mainly detected in the developing heart, gut epithelium and swim bladder.

This study aims at providing new data on the occurrence of BMP2, BMP4 and BMP16 throughout vertebrate taxonomy but also on the conservation of mature peptides in ray-finned fish. Gene expression profiles will also be investigated during development and in adult tissues of Senegalese sole *Solea senegalensis*, while regulation upon exposure to all-*trans* retinoic acid (*at*RA) will be determined using a mineralogenic cell culture established from vertebra of Senegalese sole. Similar data will be collected for BMP2 and BMP4 and compared with those of BMP16 in order to evidence expression patterns and structural differences and/or similarities.

4.1.3 Materials and methods

4.1.3.1 Sequence collection and reconstruction

BMP2, BMP4 and BMP16 sequences were retrieved from GenBank (ncbi.nlm.nih.gov) and Ensembl (ensembl.org) databases using on-site BLAST facilities, or reconstructed from expressed sequence tags (EST), genome survey sequences (GSS), whole genome shotgun (WGS) sequences and transcriptome shotgun assembly (TSA). Species-specific sequences were clustered and assembled using the ContigExpress module, from Vector NTI version 9 (Invitrogen), to generate, after manual correction, highly accurate consensus sequences. Virtual transcripts were deduced from joined consensus sequences using stringent overlap criteria.

4.1.3.2 Multiple sequence alignment and sequence logos

Distinct alignments of BMP2, BMP4 and BMP16 sequences were performed using T-Coffee facilities at tcoffee.org (Notredame et al., 2000). Manual adjustments were done to improve sequence alignments. Sequence logos were created using Weblogo facilities at weblogo.threeplusone.com (Crooks et al., 2004). Sequence logos are presented as graphical display, where the height of each letter is directly proportional to its conservation, with the more conserved residues being represented as larger characters.

4.1.3.3 Larval rearing and sampling

Senegalese sole eggs were obtained from natural spawning of farmed broodstock from the Aquaculture Research Station of the National Institute for the Sea and Atmosphere (EPPO-IPMA, Olhão, Portugal) and transferred to the CCMAR fish facilities. Environmental parameters were maintained as follows: water temperature 19.9 ± 1.2°C; salinity 36.9 ± 1.2 g L^{-1}; dissolved oxygen > 6 mg L^{-1}; photoperiod 12:12 h light:dark cycles. Hatched larvae were reared in 100-L tanks at a density of 100 larvae L^{-1}. After settlement (19 days post-hatching, dph), larvae were transferred to 3-L flat bottom plastic trays at a density of 40 larvae L^{-1}. Larvae were progressively fed, three times a day, with live preys: rotifers (*Brachionus rotundiformis*) from 2 to 10 dph, *Artemia* nauplii (AF strain; Salt Lake) from 6 to 10 dph and *Artemia* metanauplii (EG strain; Salt Lake) enriched with Easy DHA SelcoTM (INVE Aquaculture) from 9 to 17 dph. The same *Artemia* metanauplii, but immediately frozen after enrichment, were provided to the larvae from 18 dph until the end of the experiment (30 dph).

Fish eggs and larvae were sampled from 2 hours post-fertilization (hpf) to 30 dph. A total of 60 eggs and up to 20 larvae, depending on their size, were collected for RNA extraction. Several tissues (spleen, eye, brain, testis, ovaries, muscle, bone, skin, liver, heart, kidney, branchial arches and intestine) were sampled and pooled from 3 adult fish. All fish were anesthetized with a lethal dose of Tricaine methanesulfonate (MS-222, Sigma-Aldrich) and washed with sterile distilled water before sampling. Specimens and tissues collected for gene expression analysis were placed in 10 volumes of TRI-Reagent (Ambion) and stored at -80 ºC until processed.

4.1.3.4 Cell culture maintenance and exposure to retinoic acid

SS1C cell culture was developed from vertebra of Senegalese sole as described in Marques et al., 2007 and maintained in Leibovitz's L-15 medium supplemented with 10% fetal bovine serum (Sigma-Aldrich), 1% gilthead seabream *Sparus aurata* serum (Rosa et al., 2010) and 1% antibiotics/antimycotics. Cell cultures were incubated at 22ºC in a humidified atmosphere and sub-divided every 3-4 days. Unless otherwise stated, cell culture reagents and plasticware were from Invitrogen and Sarstedt, respectively. Sub-confluent cultures of SS1C cells were exposed for 24 h to 1 µM all-*trans* retinoic acid (Sigma-Aldrich) or 0.01%

of DMSO (vehicle), washed 3 times in ice-cold phosphate-buffered saline, scrapped and stored in TRI-Reagent.

4.1.3.5 RNA extraction and quantitative real-time PCR

Total RNA was extracted from samples stored in TRI-Reagent following manufacturer instructions and purified using the High Pure RNA Isolation kit (Roche Applied Science). RNA integrity was confirmed using Experion Automated Electrophoresis system (Bio-Rad) and quantity was determined using NanoDrop spectrophotometer (Thermo Scientific). Total RNA (1 µg) was reverse-transcribed for 1 h at 37ºC, using M-MLV reverse transcriptase (Invitrogen), oligo-d(T) primer and RNase OUT (Invitrogen). All quantitative real time PCR (qPCR) reactions were performed in triplicates using SsoFast EVAgreen Supermix (Bio-Rad), 0.25 µM of isoform-specific primers designed according to sequences available in GenBank (Table 4.1.1) and 1:10 dilution of reverse transcribed RNA, in the StepOnePlus Real-Time PCR system (Applied Biosystems).

Table 4.1.1 GenBank accession numbers of Senegalese sole BMP2, BMP4, BMP16 transcripts and 18S ribosomal RNA and qPCR primers used to assess gene expression levels

Gene	GenBank accession no.	Primer sequence (5'-3')	
BMP2	KF147830	Forward	GGGCCGGAGGAAATACAGCGAATCC
		Reverse	CTTGCTTGCTCGGCGTCGGTCT
BMP4	KF147828	Forward	CCAAGTCCTGCTGGGAGAGAGCAAC
		Reverse	GCAGTTCATGGCTCTGACGGGC
BMP16	KF147829	Forward	TTCCTCCTTGTCGCAGGCACC
		Reverse	TACCGAGGCACTGGCACTCCG
18S	AM882675	Forward	GAATTGACGGAAGGGCACCACCAG
		Reverse	ACTAAGAACGGCCATGCACCACCAC

PCR amplification was as follows: an initial denaturation step of 1 min at 95ºC and 40 cycles of amplification (5 s at 95ºC and 10 s at 65ºC). A calibrator sample (cDNA pooled from all samples) was included in each qPCR plate (Derveaux et al., 2010). Efficiency of amplification was between 96-104% for all primer sets. Levels of gene expression were calculated using the ΔΔCt comparative method (Pfaffl, 2001) and normalized using 18S ribosomal RNA levels. Gene expression at 24 hpf and in testis were set to 1 and used as reference samples for relative expression during development and in adult tissues, respectively.

4.1.4 Results and discussion

4.1.4.1 Bone morphogenetic protein 16 is not teleost specific

Sequences showing significant similarity to BMP2, BMP4 and BMP16 genes or transcripts were retrieved from public sequence databases, i.e. NCBI and Ensembl, using on-site blast facilities and multiple query sequences. In some cases, transcripts and genes were reconstructed from EST and TSA sequences, and from GSS and WGS sequences, respectively.

Three copies of a BMP2/4 ortholog, present in invertebrate genomes as a single copy, were found in lampreys (Agnatha; Fig. 4.1.1).

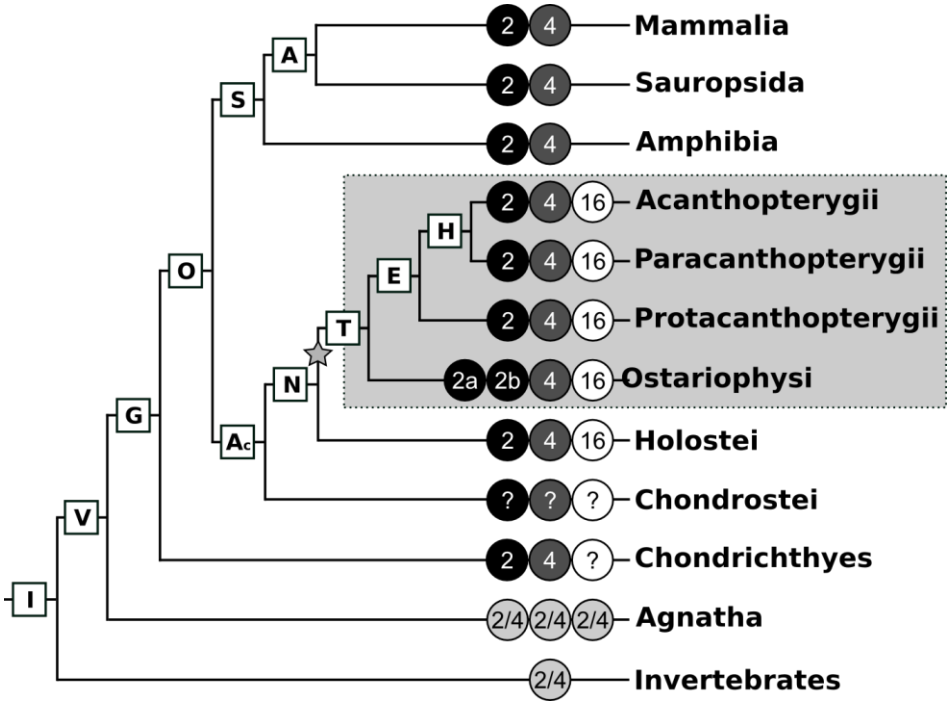


Figure 4.1.1 Occurrence of bone morphogenetic proteins 2, 4 and 16 throughout vertebrate taxonomy. Presence/absence of BMP2, BMP4 and BMP16 (*circled 2, 4* and *16*, respectively) were inferred from sequence data collected from NCBI and Ensembl sequence databases. *Circled 2a* and *2b*, Ostariophysi-specific BMP2 homologs; *Circled 2/4*, cyclostome- and invertebrate-specific BMP2/BMP4 homologs; *Circled ?*, unknown occurrence; *I*, invertebrate; *V*, vertebrata; *G*, gnathostomata; *O*, osteichthyes; *S*, sarcopterygii; *A*, amniota; *Ac*, actinopterygii; *N*, neopterygii; *T*, teleostei; *E*, euteleostei; *H*, holacanthopterygii. *Gray star* indicates the whole genome duplication event that occurred in the teleost fish lineage (*gray box*).

It has been proposed by McCauley and Bronner-Fraser (2004) that the three lamprey BMP2/4 forms (*a*, *b* and *c*) are the result of lineage specific duplication events that occurred

after the divergence of cyclostomes from gnathostomes. This fact was further confirmed through a molecular phylogenetic analysis that revealed all three lamprey genes to be equally dissimilar to gnathostomes BMP2 and BMP4, with lamprey BMP2/4a being the ortholog of gnathostome BMP2 and BMP4 (McCauley and Bronner-Fraser, 2004). The long period of divergence between agnathans and gnathostomes may difficult the understanding of BMP2 and BMP4 relationships. However, vertebrate BMP2 and BMP4 genes would be the result of an ancient duplication event that occurred at the basis of the vertebrate lineage and thus product of a single invertebrate ancestor BMP2/4 gene (Panopoulou et al., 1998; McCauley and Bronner-Fraser, 2004). Taxonomic data, inferred from genomic search, revealed the presence of a BMP16 ortholog in the genome of the spotted gar *Lepisosteus oculatus* (Fig. 4.1.1). Since holosteans did not experience the fish-specific whole genome duplication (3R) event (Sato and Nishida, 2010), we can exclude this possibility to explain the origin of BMP16. Based on this new data, we propose the following alternative scenarios: 1) BMP16 is not found in chondrosteans (e.g. sturgeon) nor chondrichthyans (e.g. sharks) and its appearance is related to a gene duplication event that occurred in the Neopterygii lineage after divergence from the Chondrostei or 2) BMP16 is present in both chondrosteans and chondrichthyans – still not identified – and its appearance is related to a gene duplication event that occurred early in the vertebrate lineage, as ultimately suggested by Feiner et al., (2009). Future works should therefore aim at determining whether BMP16 occurs in Chondrostei and Chondrichthyes. Although our search did not identify any sequences related to BMP16 in these taxonomic groups, we cannot conclude on their absence due to the lack of complete genomic information for species in these groups.

While only one BMP4 was identified in most vertebrates (not in cyclostomes), two BMP2 were evidenced in a restricted group of teleost fish, the ostariophysians. The occurrence of two BMP2 isoforms in this taxonomic group is probably related to the whole genome duplication event that occurred in the teleost fish lineage, 450 million years ago (Rafael et al., 2006), as proposed for many other genes (Gates et al., 1999; Taylor et al., 2001; Taylor et al., 2003; Sato and Nishida, 2010). Since only one BMP2 form was identified in other teleost fish (i.e. the Euteleostei), we propose that an eventual functional redundancy may have led to the subsequent loss of one of the BMP2 forms in this lineage while both were maintained in the Ostariophysi for reasons that remain to be uncovered.

4.1.4.2 Primary structure of BMP16 is very similar to those of BMP2 and BMP4

Comparative analysis of BMP2, BMP4 and BMP16 sequence logos, which have been constructed from the alignments of 18, 15 and 13 fish mature peptides, respectively (Fig. 4.1.2 and Supplementary Fig. S1 and S2), revealed a high conservation of protein primary structure for each form, with BMP4 being more conserved (77%) than BMP2 (67%) and BMP16 (66%). The high conservation of members of the BMP2/4 subfamily (represented at that time by BMP2 and BMP4), especially in their C-terminal region, has been already described by Kingsley (1994). Our data showed that BMP16, the most recent member of this subfamily, is also highly conserved. The average size of each mature peptide is also very similar, with BMP4 (117 residues) being slightly smaller than the other two members (120 residues). The levels of conservation of protein primary structure were not only high within each form, but also remarkably similar between all forms, with residues important for protein 3D structure, function and receptor binding being totally preserved in BMP2, BMP4 and BMP16 (Fig. 4.1.2).

BMPs are distinguished from other members of the TGFβ superfamily by the presence, in most cases, of seven conserved cysteine residues, instead of nine. All 7 cysteines were 100% conserved in the three members of the BMP2/4 family. Six of these cysteines form three intramolecular disulfide bridges organized in a cystine knot (Alvarez et al., 2009), which provides structural stability to the protein. The remaining cysteine is involved in an intermolecular disulfide bridge implicated in homo- and hetero-dimerization of the BMPs (Lin et al., 2006). Because of the absolute conservation of these cysteine residues, we propose that BMP16 exhibits a protein structure similar to those of BMP2 and BMP4, and requires the formation of dimers (homo and/or hetero) for proper function. BMPs are secreted as long inactive precursors and need to be processed into mature active ligands through cleavage of the propeptide at the optimal SCP cleavage site, depicted in Fig. 4.1.2 as S1. Suboptimal SCP cleavage sites were also predicted in all precursor sequences (data not shown); in other members of the TGFβ superfamily, suboptimal SCP cleavage sites have been shown to influence the amount of protein produced (Cui et al., 2001), thus functioning as a regulatory mechanism. Altogether, our data suggest that BMP16 is probably processed and regulated as BMP2 and BMP4 (Cui et al., 1998; Heng et al., 2010).

Another posttranslational modification present in proteins is glycosylation. Our in silico analysis of BMP2, 4 and 16 revealed the presence of several putative sites of N-linked

glycosylation; three of them are located on asparagine residues of the propeptide (data not shown) and only one site (100% conserved in all isoforms) is detected in the mature protein (Fig. 4.1.2). Israel and co-workers (1992) reported the secretion by CHO cells of glycosylated and non-glycosylated forms of human recombinant BMP2. Even though the posttranslational modification of proteins through N-glycosylation is an important mechanism regulating protein production and function, the non-glycosylated form of BMP2, despite less functional, still shows some biological activity *in vitro* and *in vivo* (Ruppert et al., 1996; Bessa et al., 2008).

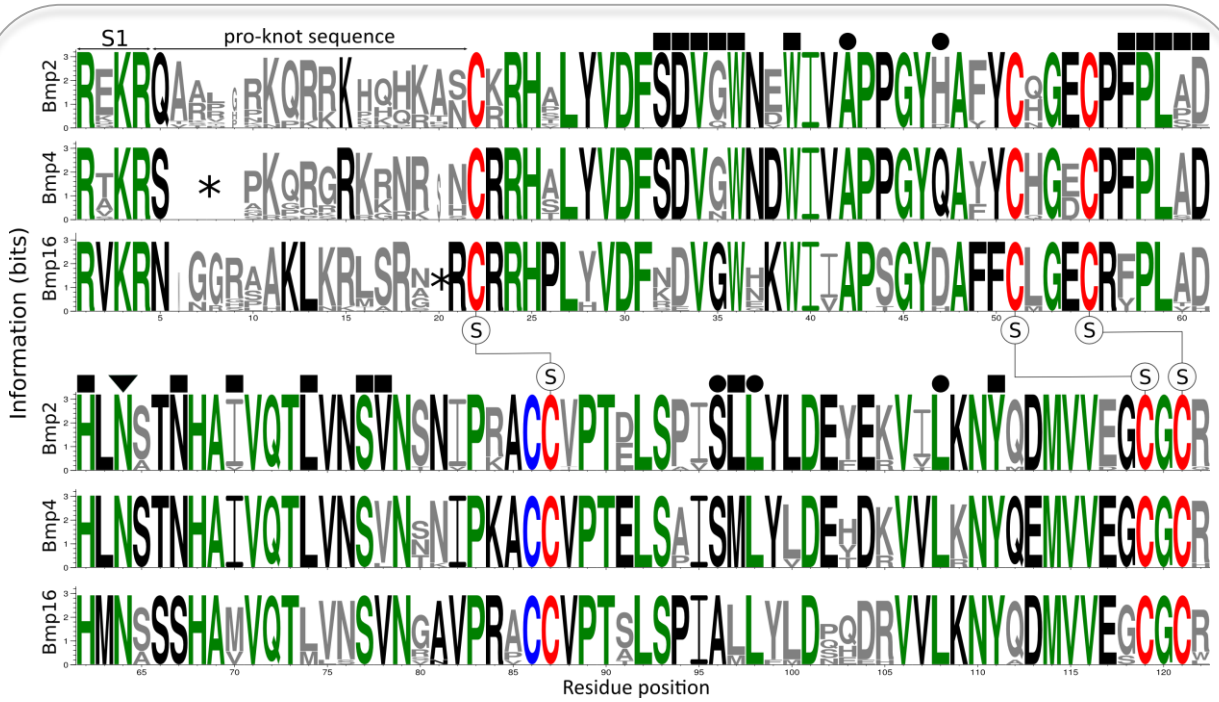


Figure 4.1.2. Conservation of fish bone morphogenetic proteins 2, 4 and 16. Logos of BMP2, BMP4 and BMP16 amino acid sequences were constructed from isoform-specific alignments of mature peptides. The height of the letters is directly proportional to its frequency. Letter color code is as follow: *Black*, residues 100% conserved in each isoform; *Green*, residues 100% conserved in all isoforms; *Red*, cysteines 100% conserved forming intramolecular disulfide bounds (cystine knot); *Blue*, cysteines 100% conserved forming intermolecular disulfide bounds responsible for dimerization. *Black triangle* indicates putative *N*-linked glycosylation site. *Black squares* indicate residues forming the wrist epitope and involved in the binding of BMP2 to type I BMP receptor (BMPR1). *Black circles* indicate residues forming the knuckle epitope and involved in binding to type II BMP receptor (BMPR2). *S1* indicates the conserved optimal SCP cleavage site. *Asterisks* indicate sites where gaps were manually inserted.

Once secreted and in order to transduce the signal to the nucleus and regulate the expression of target genes, BMPs bind to a complex of serine/threonine kinase receptors on the surface of the cells (von Bubnoff and Cho, 2001). Our protein logos showed a

considerable conservation of residues responsible for binding to BMP receptors in the BMP16 protein comparing to BMP2 and BMP4 (Fig. 4.1.2). TGFβ members always have a preferred receptor on the complex, which is called the high affinity receptor (Kirsch, Nickel, et al., 2000). In the case of BMPs and contrary to other TGFβ members, this preferential receptor is the type I (BMPR1). Indeed, only after binding of the wrist epitope of BMP monomers to the BMPR1 the lower affinity receptor (BMPR2) can interact with the knuckle epitope of each molecule of the dimer (Allendorph et al., 2006). After the interaction between ligand and receptor is completed, a cascade of phosphorylation events will activate the Smad intermediates responsible for signal translocation to the nucleus (Ducy and Karsenty, 2000). The substantial homology between BMP2, BMP4 and BMP16 in terms of residues involved in receptor binding, together with a limited promiscuity characteristic of BMP receptor systems (Kirsch, Sebald, et al., 2000), suggests that BMP16 may function as a ligand of the TGFβ signaling pathway, although functional assays will be required to confirm this possibility.

The less conserved area of BMP mature peptides is the region called pro-knot (Fig. 4.1.2), localized upstream of the first conserved cysteine, composed mainly of basic amino acids, and reported as essential for the high affinity binding of this sequence to heparin (Ruppert et al., 1996). It has also been shown that the pro-knot sequence can modulate cleavage efficiency, interact with the extracellular matrix and play a critical role in the establishment and maintenance of a morphogenetic gradient during development (Ruppert et al., 1996; Ducy and Karsenty, 2000). The lower percentage of conserved residues on the pro-knot sequence may be a consequence of not being covalently attached to the rest of the complex (Scheufler et al., 1999), becoming the most flexible part of the protein structure and being submitted to lower conservation pressure throughout evolution.

4.1.4.3 Patterns of BMP2, BMP4 and BMP16 gene expression are spatially and temporally distinct

After evidencing the remarkable conservation of BMP2, 4 and 16 protein primary structures, we analyzed the expression pattern of BMP2, BMP4 and BMP16 by qPCR during Senegalese sole larval development and in adult tissues.

All three genes were differentially expressed along Senegalese sole development, with BMP2 and BMP4 being highly expressed in early (13.5 hours post-fertilization; hpf) and late

(36 hpf) embryogenesis, respectively. Conversely, BMP16 expression was increased at a later stage, at the end of the lecitotrophic phase (5 dph) and onwards (Fig. 4.1.3).

No polyadenylated mRNA transcripts were found for BMP2 nor BMP4 on the first stages of development and the onset of expression was, for both, identified at 13.5 hpf, a period that corresponds to gastrulation in Senegalese sole. The high expression of BMP2 and BMP4 during early development, suggests that both genes may play important roles at those stages, as already documented for other vertebrate species: in gilthead seabream (*Sparus aurata*) larvae, an up-regulation of BMP2 expression was observed at gastrulation, an important stage for cell fate and embryonic patterning (Rafael et al., 2006); in zebrafish (*Danio rerio*), BMP2b was shown to be expressed in the ventral part of gastrulating embryo (Martínez-Barbera et al., 1997); and similarly, BMP2 was expressed at early gastrulation in mouse (*Mus musculus*) embryos (Ying and Zhao, 2001); in *Xenopus*, BMP4 acts as a ventralizing agent during mesoderm and neural plate formation (Wilson and Hemmati-Brivanlou, 1995) while in chicken (*Gallus gallus*), BMP4 is involved in somite patterning (Pourquié et al., 1996). A role for BMP2 and BMP4 in dorsoventral patterning was also evidenced in invertebrates (Hwang et al., 2003), suggesting a conservation of the BMP2/4 family members functions throughout evolution.

BMP16 expression was detected as soon as 2 hpf, much earlier than BMP2 and BMP4 expression, indicating a possible maternal inheritance of the transcript. The levels of BMP16 expression increased at 18.5 hpf, in accordance with data reported by Feiner et al., (2009), but remained considerably low until 5 dph. The high expression of BMP16 in later stages may indicate a more significant role during late development, in particular during metamorphosis, which onset is around 7 dph in Senegalese sole.

Tissue distribution of BMP2, 4 and 16 transcripts revealed that all isoforms are expressed in calcified tissues, although to different extents (Fig. 4.1.4), supporting the well described role of BMP2 and BMP4 in bone formation (see for example Nishimura et al., 2012) and suggesting that BMP16 may somehow share an osteogenic role. High levels of BMP16 expression in branchial arches also suggest a role in cartilage formation and/or mineralization as reported for BMP2 and BMP4. In addition, and because of the important role of branchial arches in the osmoregulation and respiration (BMP signaling is involved in osmosensing processes in fish; Kültz, 2012), a possible involvement of BMP16 in these processes should not be discarded. High levels of BMP2 expression were observed in the eye and skin samples, which contain a mixture of soft and calcified tissues. Those high levels

may be related to the presence of sclerotic cartilage in the eye, in agreement with previous reports for other BMP family ligands (Wordinger and Clark, 2007), and of scales in the skin.

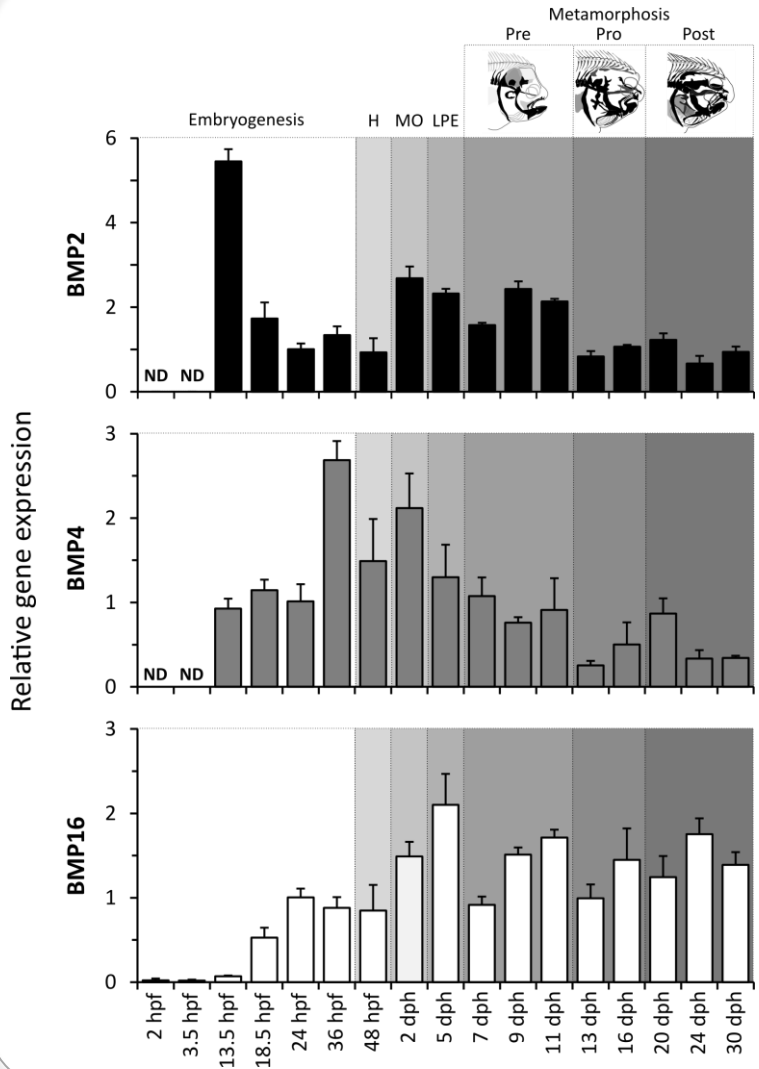


Figure 4.1.3. Relative gene expression of bone morphogenetic proteins 2, 4 and 16 throughout Senegalese sole development. Transcript levels were determined by qPCR from 3 technical replicates and normalized using 18S ribosomal RNA as housekeeping gene. Levels at 24 hpf were used as reference and set to 1. *ND*, not detected; *hpf*, hours post-fertilization; *dph*, days post-hatching; different relevant developmental processes or phases of Senegalese sole are indicated at the top of the figure: H, hatching; MO, mouth opening; LPE, lecitotrophic phase end. A schematic representation of head metamorphosis and eye migration is presented on the top of the figure, where the quantity of cartilage present is represented by the increasing intensity on the gray scale, while black areas represent bone tissue.

In addition to bone formation and dorsoventral patterning, BMPs have been associated with other processes such as reproduction. Our expression data, showing high levels of BMP4 gene expression in ovary, is in agreement with those results and also with data showing BMP4 expression in ovarian tissue of several vertebrate species (Li and Ge, 2011; Shimasaki et al., 2004). Moreover, and as previously reported in gilthead seabream, high levels of BMP2 expression were detected in Senegalese sole liver, suggesting an involvement in hepatocyte trans-differentiation (Rafael et al., 2006). Besides the high expression in branchial arches (Fig. 4.1.4), sole BMP16 is also significantly expressed in several non-calcified tissues such as intestine, heart and brain, suggesting an important role in these

tissues both during development and in adulthood, as previously described in other species and for other BMPs (Hogan, 1996; Feiner et al., 2009).

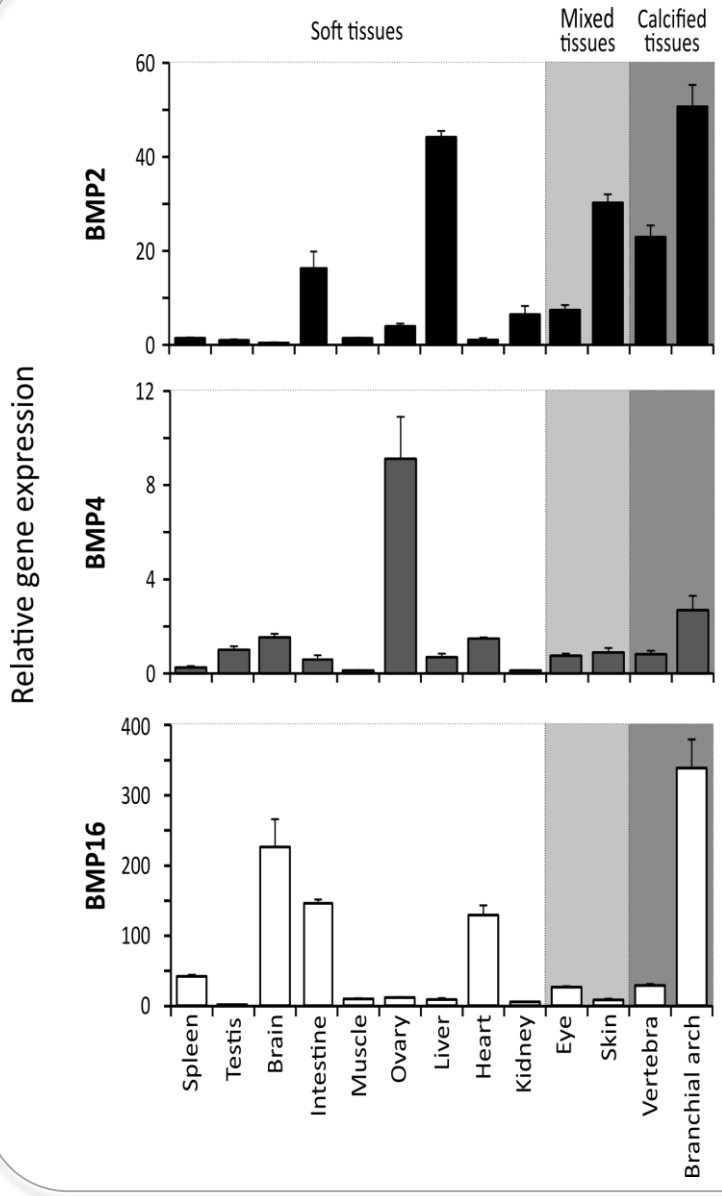


Figure 4.1.4. Relative gene expression of bone morphogenetic proteins 2, 4 and 16 in adult Senegalese sole tissues. Transcript levels were determined by qPCR from 3 technical replicates, normalized using 18S ribosomal RNA as housekeeping gene. Levels in testis were used as reference and set to 1. Eye and skin samples are a mix of soft and calcified tissues and were classified as mixed tissues.

4.1.4.4 Expression of BMP2, BMP4 and BMP16 genes is positively regulated by retinoic acid

Considering the high expression of BMP2, BMP4 and BMP16 in calcified tissues, we investigated the effect of retinoic acid (RA) on gene transcriptional regulation using a Senegalese sole bone-derived cell culture, capable of *in vitro* mineralization (SS1C cells). RA is a morphogen involved in several developmental processes, regulating target gene transcription through binding to nuclear receptors (RAR and RXR; Sheng et al., 2010). RA is

also a known repressor of the BMP signaling pathway by reducing Smad1 stability (Sheng et al., 2010). After a 24 h exposure to 1 µM of all-*trans* RA, the expression of all three genes was up-regulated (Fig. 4.1.5).

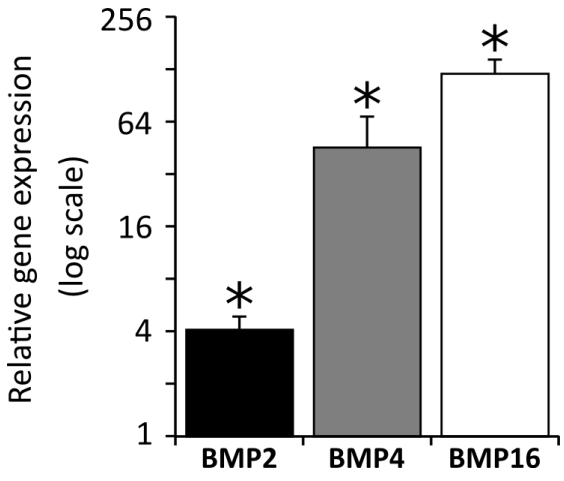


Figure 4.1.5. Relative gene expression of bone morphogenetic proteins 2, 4 and 16 in Senegalese sole SS1C cell cultures upon exposure to 1 µM all-*trans* retinoic acid for 24 h. Transcript levels were determined by qPCR from 3 technical replicates, normalized using 18S ribosomal RNA as housekeeping gene and presented as fold changes over control (cells treated with DMSO, vehicle for retinoic acid). * indicate values significantly different from each other (one-way ANOVA followed by Tukey's post-test; $P<0.05$)

Regulation of members of the BMP2/4 family by RA was previously described for other fish species, e.g. BMP4 expression was up-regulated in *Dicentrarchus labrax* larvae fed a high vitamin A dietary content (Villeneuve et al., 2006) as well as in other organisms (Hoffman et al., 2006; Sheng et al., 2010). This up-regulation may be caused by 1) a mechanism of feedback response, where the cells try to compensate the RA-dependent inhibition of BMP pathway by increasing expression of BMPs or 2) a direct transcriptional regulation through binding of RA-RAR or -RXR complexes to RA responsive elements present in the promoter regions of target genes. Our *in vitro* data indicates that BMP16, the newest member of BMP2/4 subfamily, is also regulated by RA evidencing a possible conservation of gene regulation throughout evolution and a possible role in bone metabolism.

4.1.5 Conclusions

Our data showed that BMP16 is not restricted to teleost fish, as previously described, but it is more widely distributed as evidenced by its presence in the Holosteans. Despite the overall conservation of BMP2, 4 and 16 sequences, that may suggest similar protein structure, processing mechanisms and signaling pathways, the three genes showed different developmental and tissues specificities. Retinoic acid was able to up-regulate BMP2, BMP4 and BMP16 expression, although to different extent. Altogether, our results contribute to the

better understanding of vertebrate distribution and characterization of BMP16, the newest member of the BMP2/4 family.

4.1.6 Acknowledgments

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. It was also partially financed by the European Community (EC) through ASSEMBLE (FP7/227799) research project. CM and JR were supported by doctoral grants (SFRH/BD/39964/2007 and SFRH/BD/47433/2008, respectively) from the FCT. IF was supported by a post-doctoral grant (SFRH/BPD/82049/2011) from the FCT. The authors are grateful to P. Gavaia (EDGE-CCMAR) for Senegalese sole BMP partial sequences and to Pedro Pousão-Ferreira (EPPO-IPMA) for Senegalese sole eggs. SS1C cell culture was developed within the scope of the collaborative action AI-E-54/11 (SOLVITAK), between Portugal and Spain and funded by the Conselho de Reitores das Universidades Portuguesas (CRUP).

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1111/jai.12539

4.2 Comparative analysis of zebrafish bone morphogenetic proteins 2, 4 and 16: molecular and evolutionary perspectives

Cátia L. Marques
Ignacio Fernández
Michael N. Viegas
Cymon J. Cox
Paulo Martel
Joana Rosa
M. Leonor Cancela
Vincent Laizé

4.2.1 Abstract

BMP2, BMP4 and BMP16 form a subfamily of bone morphogenetic proteins acting as pleiotropic growth factors during development and as a bone inducers during osteogenesis. BMP16 is the most recent member of this subfamily and basic data regarding protein structure and function, and spatio-temporal gene expression is still scarce. In this work, insights on BMP16 were provided through the comparative analysis of structural and functional data for zebrafish BMP2a, BMP2b, BMP4 and BMP16 genes and proteins, determined from tri-dimensional models, patterns of gene expression during development and in adult tissues, regulation by retinoic acid and capacity to activate BMP signaling pathway. Structures of Bmp2a, Bmp2b, Bmp4 and Bmp16 were found to be remarkably similar, with residues involved in receptor binding being highly conserved All proteins could activate the BMP signaling pathway, suggesting that they share a common function. On the contrary, stage- and tissue-specific expression of *bmp2*, *bmp4* and *bmp16* suggested that gene transcription may be regulated by different factors but also that they are involved in distinct physiological processes, although with the same function. Retinoic acid, a morphogen known to interact with BMP signaling during bone formation, was shown to down-regulate the expression of *bmp2*, *bmp4* and *bmp16*, although to different extents. Taxonomic and

phylogenetic analyses indicated that *bmp16* diverged before *bmp2* and *bmp4*, is not restricted to teleost fish lineage as previously reported, and that it probably arose from a whole genomic duplication event that occurred early in vertebrate evolution and disappeared in various tetrapod lineages through independent events.

Note

Names/acronyms of genes/proteins of species with different nomenclature conventions are used throughout this study. To reduce heterogeneity no convention will be used and acronyms will be uppercased. However, convention will be maintained for zebrafish which is the main specie studied here.

4.2.2 Introduction

Whole genome duplications (WGDs) are key features in species evolution that allow organisms to develop new characteristics (Magadum et al., 2013). WGDs are often related to bursts in organism diversity and complexity (Magadum et al., 2013; Santini et al., 2009) and many families of genes are known to have evolved through genome duplication (Magadum et al., 2013; Panopoulou et al., 1998). Despite some controversy, it is commonly accepted that three whole genome duplication occurred during vertebrate evolution: the first and second duplication events occurred early in the vertebrate lineage, approximately 500 million years ago, while the third event only affected teleost fish genome (Santini et al., 2009; Sharman and Holland, 1998; Escriva et al., 2002; Furlong et al., 2007; Kuraku et al., 2009). After duplication, the paralogous gene (copy of the original gene) can co-exist with the original copy and complement its function (subfunctionalization) or diverge and develop a new function (neofunctionalization) (Dehal and Boore, 2005; Wolfe, 2001). However, WGDs are typically followed by massive gene loss, and in most cases only a single copy of the duplicated genes will be maintained (Rabier et al., 2014). Moreover, gene loss among distantly related lineages often results in hard-to-interpret molecular phylogenies as in the case of bone morphogenetic proteins (BMPs) 2, 4 and the recently identified BMP16 (Kuraku, 2010).

Products of the BMP2 and BMP4 genes belong to the transforming growth factor β (TGFβ) superfamily of multifunctional growth factors, and are involved in several key mechanisms of vertebrate development (Chen et al., 2004; Wu and Hill, 2009), such as bone

formation (Nishimura et al., 2012). The primary structures of BMP2 and BMP4 proteins have been remarkably conserved throughout evolution and human mature BMP2 and BMP4 share 90% identity and are 75% identical to their *Drosophila* homolog decapentaplegic (DPP). The conservation of protein function has also been demonstrated through the interchangeability of *Drosophila* and mammalian proteins: DPP can induce endochondral bone formation when introduced subcutaneously in mouse (Sampath et al., 1993), while mammalian BMP4 protein is able to rescue the dorsal-ventral defects resulting from the lack of DPP in *Drosophila* (Padgett et al., 1993). BMP2 and BMP4 genes are expressed in a wide variety of tissues, and their expression in early stages of development was shown to be crucial for organism viability, as demonstrated by the early lethality of mice deficient for BMP2 or BMP4 gene (Zhang and Bradley, 1996).

Based on sequence similarities, Feiner et al. (Feiner et al., 2009) identified in 2009 a new member of the BMP2/4 subfamily, which was later named BMP16. While it was initially presented as a teleost fish specific protein, its presence has recently been reported in the genome of non-teleost fish species suggesting that its origin is not related to the third, teleost-specific, WGD event (Marques et al., 2014). Expression of *bmp16* gene was analyzed by *in situ* hybridization during early zebrafish development and detected mainly in the developing heart, gut epithelium and swim bladder (Feiner et al., 2009). In adult Senegalese sole tissues, BMP16 transcript was detected through quantitative real-time PCR in branchial arches, brain, intestine and heart (Marques et al., 2014). Not much more is known about BMP16, in particular its capacity to activate BMP signaling pathway as BMP2 and BMP4 (Goldman et al., 2008; Upton et al., 2008; Hung et al., 2012) or its role during vertebrate development.

The aim of this work is to provide new insights into the origin of BMP16 and characterize the evolutionary relationship of the members of BMP2/4/16 subfamily but also to collect data on protein function and gene regulation through the comparative analysis of protein structure homology models, capacity to activate BMP signaling pathway, spatiotemporal gene expression profiles and regulation by retinoic acid.

4.2.3 Materials and methods

4.2.3.1 Ethics Statement

Animal handling and experiments are legally accredited by the Portuguese Direcção Geral de Veterinária (DGV) and all the experimental procedures involving animals were performed according the EU (Directive 86/609/CEE) and National (Portaria nº 1005/92 de 23 de Outubro; Portaria nº 466/95 de 17 de Maio; Portaria nº 1131/97 de 7 de Novembro) legislation for animal experimentation and welfare.

4.2.3.2 Gene sequence collection and reconstruction

Annotated sequences for BMP2, BMP4 and BMP16 were retrieved from GenBank (ncbi.nlm.nih.gov) and Ensembl (ensembl.org) databases using on-site BLAST facilities. In some cases, sequences were reconstructed from expressed sequence tags (EST), genome survey sequences (GSS), whole genome shotgun (WGS) sequences, and transcriptome shotgun assembly (TSA) available through GenBank sequence databases. Species-specific sequences were clustered and assembled using the ContigExpress module of Vector NTI software (Invitrogen). Gene structures were predicted using the Spidey mRNA-to-genomic alignment tool (ncbi.nlm.nih.gov). Genomic organization of BMP16 gene flanking regions was determined using genomic data available in Ensembl database.

4.2.3.3 Multiple sequence alignment and phylogenetic reconstruction

BMP2, BMP4 and BMP16 gene sequences were aligned using TranslatorX V1.1 (Abascal et al., 2010). The nucleotide alignment was manually adjusted using SeaView V3.2 (Galtier et al., 1996) where parts of sequences were arbitrarily aligned (2 accessions, namely, sea squirt and fruit fly). Unambiguously aligned characters were defined using Gblocks V0.91b (Castresana, 2000) and the following options in the SeaView interface: “allow gaps in final positions” (-b5=h), "do not allow many contiguous non conserved positions" (-b3=4), and "allow smaller final blocks" (-b4=5). A single block of positions was eliminated where the alignment with outgroup sequences (lancelet, fruit fly, sea squirt) was arbitrary. The final alignment consisted of 58 taxa and 612 aligned nucleotides, which translated to 204 amino acids. Phylogenetic analyses, using maximum likelihood, were conducted using RAxML

V7.8.4-MP (Stamatakis, 2006) and Bayesian phylogenetic inferences were conducted using P4 V0.89.r234 (Foster, 2004). Appropriate models were determined using ModelGenerator V0.85 (Keane et al., 2006): for nucleotides this was a general time-reversible substitution model (GTR) with a gamma-distribution of among-site rate variation (4 discrete categories) (+Γ) and a proportion of invariant sites (+I), and for amino acids the LG (Le and Gascuel, 2008) empirical substitution matrix with +Γ and estimated stationary amino acid frequencies (+Fest). Non-stationary composition model analyses were performed in P4 with the addition of extra composition vectors (CV) to the best model (see **Online resource 2** for details on individual analyses).

4.2.3.4 3D model building of zebrafish Bmp2a, Bmp2b, Bmp4 and Bmp16

Mature peptides of zebrafish BMP2, NMP4 and BMP16 and human BMP2 were aligned using the ClustalW Omega server (ebi.ac.uk/Tools/msa/clustalo). The pairwise identities ranged from 64.7% (BMP16) to 83.3% (BMP2b). At this level of identity, reliable homology models can be built from one structure determined experimentally (Eswar et al., 2006). Structure 1REW available in the Protein Data Bank (complex between human BMP2 and the ectodmain of its type IA receptor (Keller et al., 2004)) was selected to serve as template. Water molecules and receptor chains were removed from the structure file and the remaining BMP2 dimer used as template for homology building using the MODELLER v9.12 software package (Sali and Blundell, 1993). Since the first 11 residues of human BMP2 mature peptide are missing in PDB structure, matching zebrafish segments were removed from the alignment. Structures were modelled as dimers with an intermolecular SS bond, the active form of BMPs. For each protein 50 models were generated and the one with the best DOPE score (Shen and Sali, 2006) was selected. The absolute quality of the four selected models was evaluated using the Z-DOPE score, and also checked using the QMEAN and MOLPROBITY servers. The final models were displayed and analyzed using the PyMOL Molecular Graphics System version 1.5.0.4 Schrödinger, LLC.

4.2.3.5 Vector construction

Coding sequences of zebrafish *bmp2a*, *bmp2b*, *bmp4*, and *bmp16* (accession numbers NM_131359, NM_131360, NM_131342 and NM_001171776, respectively) were amplified by PCR using the proofreading Advantage cDNA polymerase (Clontech), reverse-transcribed mRNA extracted from ZFB1 cells (Vijayakumar et al., 2013) and gene-specific primers designed according to available sequences (Table 1) and directionally inserted into pcDNA3.1 expression vector (Invitrogen), under the control of pCMV promoter. DNA integrity was confirmed through sequencing (Note: cloned sequences contained single nucleotide polymorphisms that did not alter protein sequence; they have been deposited into GenBank database with the following accession numbers: *bmp2a*, KM820423; *bmp2b*, KM820424; *bmp4*, KM820425 and *bmp16*, KM820426).

4.2.3.6 Luciferase assays

ABSa15 cells (ECACC catalogue no. 13112201) (Tiago et al., 2014) were seeded at 4 $\times$ 10^4 cells/well in 24-well plates and cultured in DMEM medium for 16 h at 33ºC under 10% CO_2. Sub-confluent cultures were transfected using 1.5 µl of X-tremeGENE HP DNA transfection reagent (Roche) with vectors expressing (1) zebrafish *bmp2a*, *bmp2b*, *bmp4* or *bmp16* under the control of CMV promoter (pcDNA3 vector backbone; 200 ng), (2) firefly luciferase under the control of BMP-responsive elements (BRE-Luc vector; 250 ng (Korchynskyi and ten Dijke, 2002)) and/or (3) renilla luciferase under the control of SV40 promoter (pRL-SV40 vector; 200 ng; Promega). After 48 h, cells were lysed and luciferase activities were measured in a BioTek Synergy 4 plate reader using Dual-Luciferase Reporter Assay system (Promega). Relative luciferase activity was determined from the ratio F-Luc/R-Luc and is presented as the fold change over pGL3 basic vector.

4.2.3.7 Larval rearing and sampling

Zebrafish eggs were obtained from natural spawning of in-house broodstock maintained in a ZebTec housing system (Tecniplast). Water parameters were maintained as follows: pH 7.6 $\pm$ 0.2; conductivity 700 mS; dissolved oxygen 7.8 mg L^{-1}; photoperiod 14:10 hours light:dark. Fertilized eggs were maintained until hatching in 1-L water tanks at a

density of 200 eggs L^{-1} with 0.5 ppm of methylene blue to avoid fungi development. Hatched larvae were raised until 30 days post-fertilization (dpf) in 1-L water tanks at a density of 100 larvae L^{-1}, with 90% of the water renewed every two days. Larvae from 5 to 10 dpf were fed twice a day with *Artemia* nauplii (AF strain INVE, 5-10 nauplii mL^{-1}) and from 8 to 30 dpf with *Artemia* metanauplii (EG strain INVE, 10 metanauplii mL^{-1}). Juveniles, adults and broodstock were fed twice a day with commercial dry food and once a day with *Artemia* metanauplii.

Fish embryos and larvae were sampled at 1 (4 cells), 3 (1k cell), 16 (14 somites), 24, 32, 48, 72 and 96 hours post fertilization (hpf), and 5, 7, 9, 12, 15, 20, 25 and 30 dpf. The amount of material sampled at each developmental stage was adapted to specimen size and ranged between 100 eggs and 5 early juveniles (30 dpf). Adult zebrafish tissues were collected and pooled from 3 males and 2 females. All fish were anesthetized with a lethal dose of tricaine methanesulfonate (MS-222, Sigma-Aldrich) and washed with sterile distilled water before sampling. Specimens and tissues collected for gene expression analysis were placed in 10 volumes of TRI-Reagent (Ambion) and stored at -80 ºC until processed.

4.2.3.8 Cell exposure to retinoic acid

Sub-confluent cultures of ZFB1 cells (Vijayakumar et al., 2013) were exposed for 24 h to 1 µM all-*trans* retinoic acid (*at*RA; Sigma-Aldrich) or 0.01% of dimethyl sulfoxide (DMSO; vehicle), washed 3 times in ice-cold phosphate-buffered saline, scrapped out and stored in TRI-Reagent.

4.2.3.9 RNA extraction and quantitative real-time PCR

Total RNA was extracted from samples stored in TRI-Reagent following manufacturer instructions and purified using the High Pure RNA Isolation kit (Roche). RNA integrity was confirmed using Experion Automated Electrophoresis system (Bio-Rad) and quantity was determined using NanoDrop spectrophotometer (Thermo Scientific). Total RNA (500 ng) was reverse-transcribed for 1 h at 37ºC using M-MLV reverse transcriptase (Invitrogen), oligo-d(T) primer and RNase OUT (Invitrogen). All quantitative real-time PCR (qPCR) reactions were performed in triplicates using SsoFast EvaGreen Supermix (Bio-Rad), 0.25 µM of isoform-specific primers (Table 4.2.1) and 1:10 dilution of reverse-transcribed RNA, in a StepOnePlus Real-Time PCR system (Applied Biosystems).

Table 4.2.1. PCR primers used in this study to assess gene expression and subclone coding sequences in expression vectors. Primers were designed according to zebrafish sequences BMP2a (accession no. NM_131359), BMP2b (accession no. NM_131360), BMP4 (accession no. NM_131342), BMP16 (accession no. NM_001171776)

Name	Primer sequence (5'-3')*	Application
DreBMP2a_1Fw	CTGAGCCCGTCTGATCTCCTTCGTC	qPCR
DreBMP2a_1Rv	GCTGCTGGGAGTGGGTCTGTGCTGGAG	
DreBMP2b_1Fw	GAGGAACTTAGGAGACGACGGGAACGC	
DreBMP2b_1Rv	TCTCGGGAATGAGTCCAACGGCAC	
DreBMP4_1Fw	CGCCGTCGTACCACAGTATCTGCTC	
DreBMP4_1Rv	ATAGTCGAAGCTGACGTGCTGCGC	
DreBMP16_1Fw	CGTCATCGACAACTCAAAGGGACCAA	
DreBMP16_1Rv	GCGAAGAAGTGCCCTGCAATCAGTTA	
DrebACTIN2_Fw	GCAGAAGGAGATCACATCCCTGGC	
DrebACTIN2_Rv	CATTGCCGTCACCTTCACCGTTC	
DreRPS_Fw	AACACGAACATTGATGGAAGACG	
DreRPS_Rv	ATTAGCAAGGACCTGGCTGTATTT	
DreBMP2a_2Fw_HindIII	CGAAGCTTATCATCATGGTCTCGTCCACCGCC	Cloning
DreBMP2a_2Rv_XhoI	CCCTCGAGGTGGCGTCAGCGGCACCCGCATCC	
DreBMP2b_2Fw_KpnI	CCGGAGGGTACCTGATCATGGTCGCCGTGGTCC	
DreBMP2b_2Rv_XhoI	CCCTCGAGAGATTGTTCTCATCGGCACCC	
DreBMP4_2Fw_HindIII	CGAAGCTTGACATCATGATTCCTGGTAATCGAATG	
DreBMP4_2Rv_XhoI	CCCTCGAGCTCCGTTTAGCGGCAGCCACACC	
DreBMP16_2Fw_HindIII	CGAAGCTTTCCAACATGTTCCCTGCTAGCCTA	
DreBMP16_2Rv_XhoI	CCCTCGAGATCTGGCTATCGACAGCCACATCC	

*Underlined sequences indicate recognition site for endonucleases cited in primer name

PCR amplification was as follows: an initial denaturation step of 1 min at 95ºC and 40 cycles of amplification (5 s at 95ºC and 10 s at 65ºC). Efficiency of amplification was above 95% for all primer sets. Levels of gene expression were calculated using the ΔΔCt comparative method (Pfaffl, 2001) and normalized using housekeeping genes, which suitability was evaluated using Normfinder and BestKeeper algorithms (Andersen et al., 2004; Pfaffl et al., 2004). *β-actin 2* and *rps18* were selected to normalize gene expression in developmental stages and cells, and tissues, respectively.

4.2.4 Results

4.2.4.1 Taxonomic distribution of vertebrate BMP2, BMP4 and BMP16: BMP16 is not fish-specific

Sequences with a high similarity to BMP2, BMP4 and BMP16 were retrieved from GenBank and Ensembl databases using on-site BLAST tools and multiple sequences as queries. An overview of the taxonomic distribution of BMP2, BMP4 and BMP16 is presented in Fig. 4.2.1 A single BMP2/4 isoform is present in several invertebrate genomes and three copies of the same gene are found in the jawless fish superclass Agnatha (e.g. lampreys). BMP2 and BMP4 are present as two independent genes in all Gnathostomes (clade G in Fig. 4.2.1).

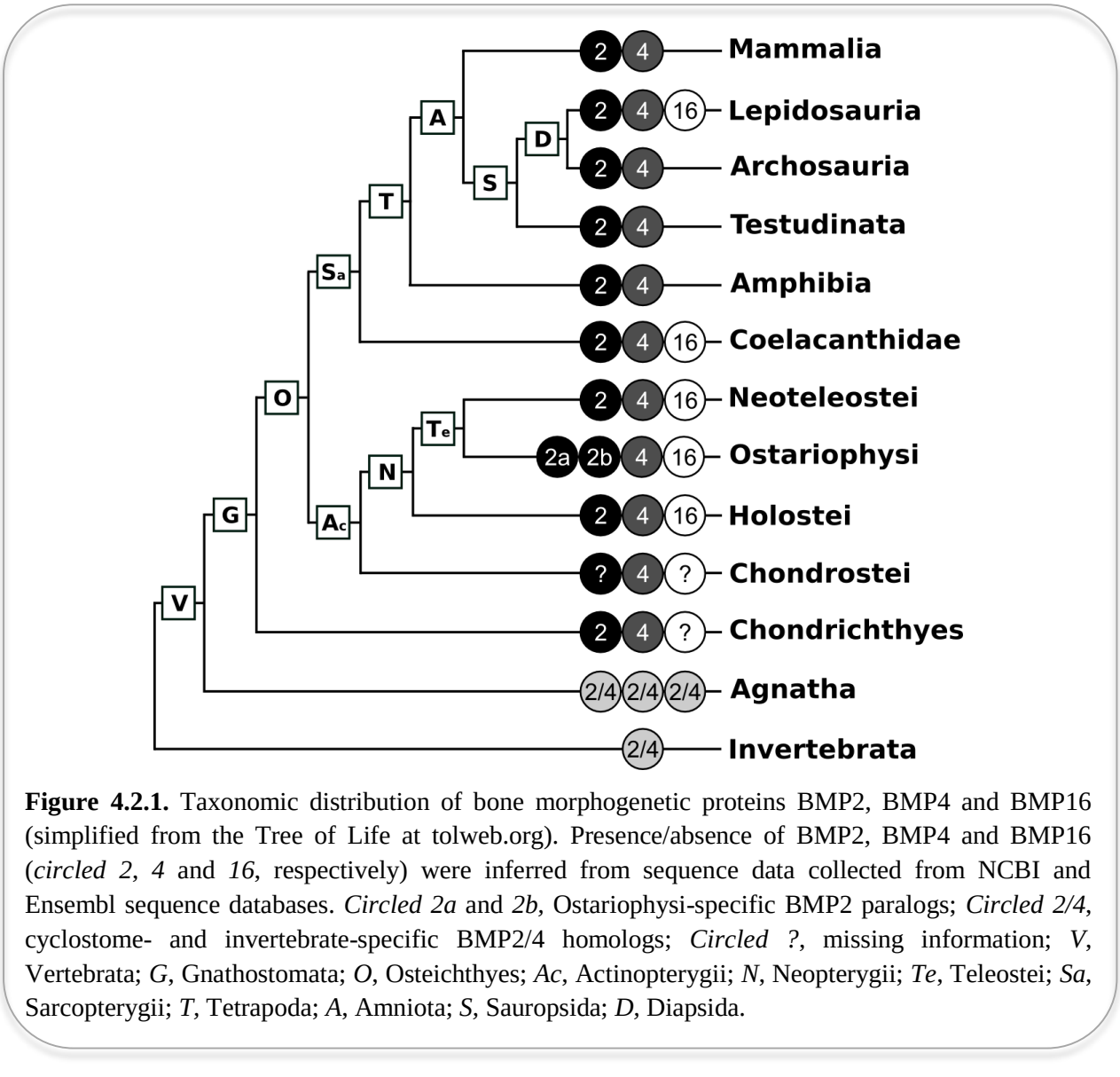


Figure 4.2.1. Taxonomic distribution of bone morphogenetic proteins BMP2, BMP4 and BMP16 (simplified from the Tree of Life at tolweb.org). Presence/absence of BMP2, BMP4 and BMP16 (*circled 2*, *4* and *16*, respectively) were inferred from sequence data collected from NCBI and Ensembl sequence databases. *Circled 2a* and *2b*, Ostariophysi-specific BMP2 paralogs; *Circled 2/4*, cyclostome- and invertebrate-specific BMP2/4 homologs; *Circled ?*, missing information; *V*, Vertebrata; *G*, Gnathostomata; *O*, Osteichthyes; *Ac*, Actinopterygii; *N*, Neopterygii; *Te*, Teleostei; *Sa*, Sarcopterygii; *T*, Tetrapoda; *A*, Amniota; *S*, Sauropsida; *D*, Diapsida.

No BMP2 sequence was found in Chondrostei (e.g. sturgeon and bichir) and although we cannot exclude the possibility of gene loss in the lineage, it is probable that its absence in sequence databases is a consequence of the scarce genomic data available for those species. In contrast, two BMP2 genes (BMP2a and BMP2b) are present in Ostariophysi genomes (e.g. zebrafish). The presence of BMP16 gene is restricted to the genome of few Gnathostome taxa. BMP16 gene has been identified in all Neopterygii (clade N in Fig. 4.2.1) but not in Chondrichthyes and Chondrostei, which are both early-branching fish lineages of gnathostomes and again have limited genomic data available. The BMP16 gene was also found in Coelacanthidae (coelacanth) and in Lepidosauria (e.g. lizards and snakes) but not in other Sarcopterygii (clade Sa in Fig. 4.2.1), including Amphibia, Testudinata, Archosauria, and Mammalia. In the latter organisms genomic data is abundant suggesting that the absence of BMP16 is not due to a lack of data.

To better understand the mechanisms underlying the loss of the BMP16 gene, genomic regions flanking BMP16 locus were analyzed from a subset of species representing the major vertebrate taxonomic groups (Fig. 4.2.2). As expected, gene composition and synteny surrounding the BMP16 locus are more conserved in closely related species (e.g. the Japanese medaka, the spotted green pufferfish, the three-spined stickleback and the southern platyfish) than in evolutionarily distant species (e.g. zebrafish, coelacanth and green anole). However in some cases, for instance in three-spined stickleback and southern platyfish, gene substitution or translocation was observed. In tetrapods, where the BMP16 gene is missing, two different scenarios were observed: (1) genes surrounding the BMP16 locus (i.e. GEMIN7 and PPP1R37 (core genes) but also RELB, CLPTM1, RTN2) are present, suggesting that the BMP16 gene was selectively removed from their genome (e.g. western clawed frog, painted turtle, mouse and human) and (2) genes surrounding the BMP16 locus were also absent, suggesting that the entire chromosome region was lost (e.g. chicken and zebra finch).

An analysis of the regions that would typically have contained the BMP16 gene (scenario 1) did not reveal the presence of undetected genes, pseudogenes or remnants of the BMP16 gene, favoring the hypothesis of an active removal of BMP16 gene in these species. Our data demonstrate the presence of the BMP16 gene in ray-finned, lobe-finned fish and also in tetrapods, although it has been independently lost in several tetrapods during evolution. Moreover, our data show that the BMP16 gene is not specific to the teleost fish lineage as previously claimed (Feiner et al., 2009).

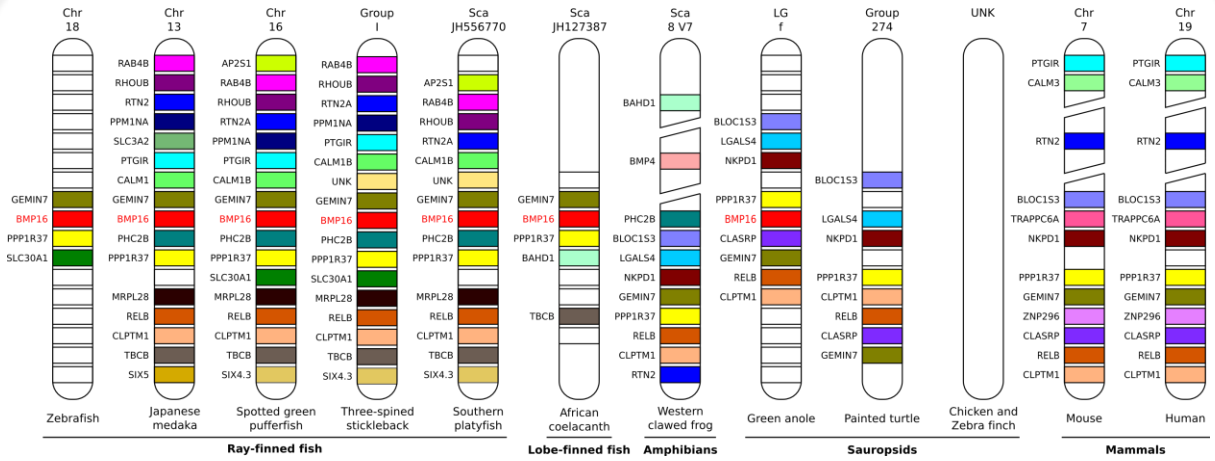


Figure 4.2.2. Schematic representation of the genomic region flanking vertebrate BMP16 gene using data from Ensembl project. Genes present in the vicinity of BMP16 locus are indicated in colored boxes irrespectively of their orientation. Gene names are indicated on the left side of each scheme. Unnamed white boxes indicate genes present in the vicinity of BMP16 locus in only one species. *Chr*, chromosome; *LG*, linkage group; *Sca*, scaffold; *UNK*, unknown. Vertebrate species are: zebrafish *Danio rerio*, Japanese medaka *Oryzias latipes*, spotted green pufferfish *Tetraodon nigroviridis*, three-spined stickleback *Gasterosteus aculeatus*, Southern platyfish *Xiphophorus maculatus*, African coelacanth *Latimeria chalumnae*, Western clawed frog *Xenopus tropicalis*, green anole *Anolis carolinensis*, painted turtle *Chrysemys picta bellii*, chicken *Gallus gallus*, zebra finch *Taeniopygia guttata*, mouse *Mus musculus*, human *Homo sapiens*. Gene names are: *Gemin7*, gem associated protein 7; *BMP16*, bone morphogenetic protein 16; *PPP1R37*, protein phosphatase 1 regulatory subunit 37; *SLC30A1*, solute carrier family 30 member 1; *RAB4B*, RAB4B member RAS oncogene family; *RHOUB*, Ras homolog gene family member Ub; *RTN2*, reticulon 2; *PPM1NA*, protein phosphatase Mg^{2+}/Mn^{2+} dependent 1Na; *PTGIR*, prostaglandin I2 receptor; *CALM1*, calmodulin 1; *PHC2B*, polyhomeotic homolog 2b; *MRPL28*, mitochondrial ribosomal protein L28; *RELB*, avian reticuloendotheliosis viral oncogene related B; *CLPTM1*, cleft lip and palate associated transmembrane protein 1; *TBCB*, tubulin folding cofactor B; *SIX5*, sine oculis homeobox 5; *AP2S1*, adaptor-related protein complex 2 sigma 1 subunit; *BAHD1*, bromo adjacent homology domain containing 1; *BMP4*, bone morphogenetic protein 4; *BLOC1S3*, biogenesis of lysosomal organelles complex 1 subunit 3; *LGALS4*, lectin galactoside-binding soluble 4; *NKPD1*, NTPase KAP family P-loop domain containing 1; *CLASRP*, CLK4-associating serine/arginine rich protein; *TRAPPC6A*, trafficking protein particle complex 6A; *ZNF296*, zinc finger protein 296.

4.2.4.2 Evolution of the BMP2, BMP4, and BMP16 genes: BMP16 diverged before BMP2 and BMP4

The molecular phylogeny of BMP2, BMP4 and BMP16 genes was inferred from a subset of 58 complete coding sequences (Supplementary figure 4.2.1), representing the main vertebrate taxa (i.e. jawless fish, cartilaginous fish, ray-finned and lobe-finned fish, amphibians, sauropsids and mammals). The optimal maximum likelihood (ML) tree and the bootstrap proportions of the BMP nucleotide data under the GTR+Γ+F_{est} model are presented in Supplementary figures 4.2.2A and 4.2.2B. Bayesian Markov chain Monte Carlo (MCMC)

analyses revealed that the data were non-stationary that and 2 composition vectors using the node-discrete compositional heterogeneity (NDCH) model implemented in P4 were necessary to model the among lineage composition (Supplementary figures 4.2.2C, 4.2.2D and 4.2.2E). ML bootstrap and Bayesian MCMC analyses of the translated amino acid sequences of the BMP genes resulted in trees which were neither analysis well-resolved or well-supported due to a lack of substitutional information at the protein level (data not shown). A 50% majority rule consensus tree of trees sampled from the posterior distribution of the non-stationary composition P4 MCMC (GTR+Γ+I+CV2; Supplementary figure 4.2.2E) analysis of the nucleotide sequence data (analytical details are provided in Suplementary figure 4.2.2) is presented in Fig. 4.2.3, with posterior probabilities and ML bootstrap support values indicated at nodes.

In this tree, BMP16 is seen to diverge first before the split of BMP2, BMP4 and BMP2/4 of the Agnatha. Branch lengths tend to be slightly longer in the BMP16 clade compared to BMP2 and BMP4 clades, indicating a higher substitution rate and molecular divergence of this isoform. The BMP2/4 isoforms (2/4a, 2/4b or 2/4c) from lamprey are more closely related to the BMP2 and BMP4 members than they are to BMP16, indicating that the three isoforms of BMP2/4 present in lamprey genomes are the result of lineage specific duplication and would have occurred after the split of BMP16 from BMP2 and BMP4. In all the Gnathostomes only one homolog for each of the three genes (BMP2, BMP4 and BMP16) was found, except for a particular group of teleost fish (Ostaryophisi), which includes zebrafish. It is probable that the two isoforms of BMP2 (BMP2a and BMP2b) present in Ostariophysi resulted from the third, teleost-specific, WGD that occurred approximately 350 million years ago (Meyer and Van de Peer, 2005), and suggesting that the second isoform was subsequently lost in the lineage leading to modern teleosts.

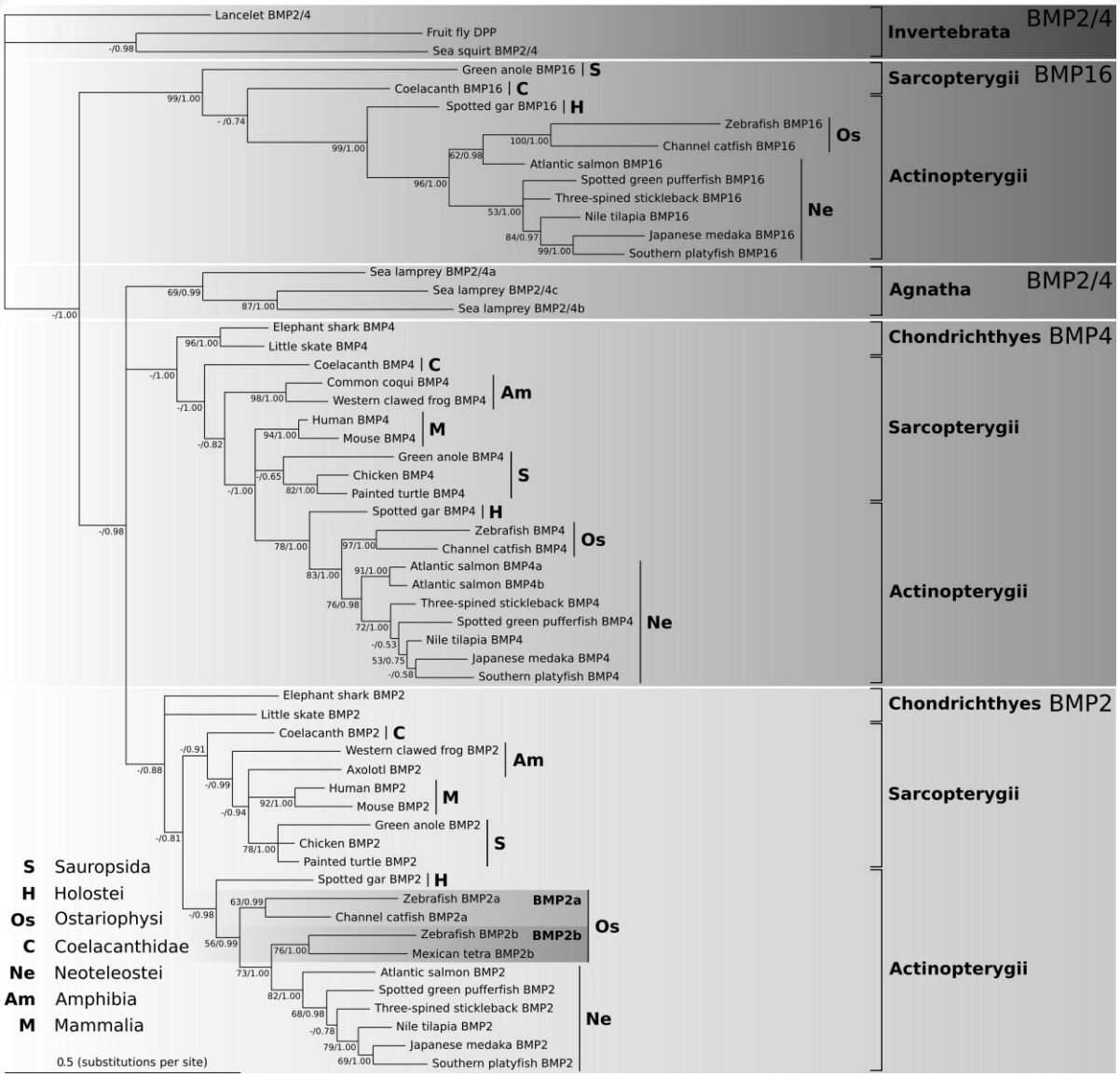


Figure 4.2.3. Phylogenetic relaionship among vertebrate BMP2, BMP4 and BMP16 homologs. The tree is a 50% majority-rule consensus tree of trees obtained from the posterior distribution of a composition heterogeneous P4 Bayesian MCMC analysis (Fig. E Online resource 2): model GTR+Γ+I+CV2, marginal likelihood $-\ln L_h$ = 17045.6169, posterior predictive simulations of $\chi 2$ statistic of composition homogeneity p-value = 0.4613. Numbers on the branches represent bootstrap values (**Online resource 2b)** and posterior probabilities of the maximum likelihood and Bayesian analysis, respectively. The tree is rooted in the outgroup taxa Invertebrata.

4.2.4.3 Zebrafish BMP2a, BMP2b, BMP4 and BMP16 gene and protein structures are remarkably similar

The structure of zebrafish *bmp2a*, *bmp2b*, *bmp4* and *bmp16* was determined from genomic information available in GenBank database (accession numbers: *bmp2a*,

NC_007128; *bmp2b*, NC_007131; *bmp4*, NC_007128 and *bmp16*, CAAK05042509 and CAAK05042510) by direct comparison with mRNA sequences (Fig. 4.2.4).

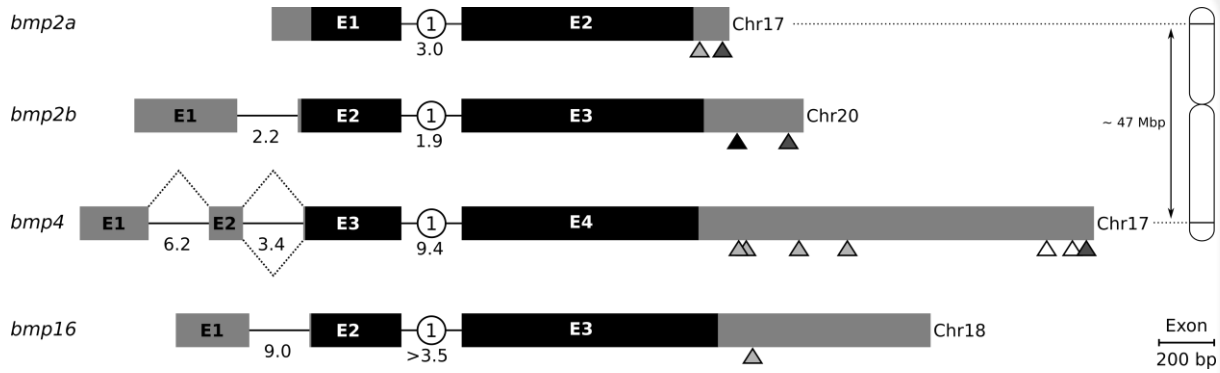


Figure 4.2.4. Structure of zebrafish *bmp2a*, *bmp2b*, *bmp4* and *bmp16* genes. Gene sequences were collected/reconstructed from GenBank using whole-genome shotgun (WGS), genome survey sequence (GSS) and/or high-throughput genome sequencing (HTGS). Exons (*E*) are displayed as *gray boxes* (non-coding exons) and *black boxes* (coding exons). Introns are displayed as *solid black lines* and respective size (kbp) is indicated below the phase of intron insertion (*white circles*). *Dashed lines* indicate alternative splicing *Triangles* indicate local of predicted polyadenylation signals in 3' untranslated region: *white*, manual prediction; *light gray*, manual and bioinformatics predictions; *dark gray*, manual prediction and EST evidence; *black*, manual and bioinformatics predictions, and EST evidence. A schematic representation of chromosome 17 (Chr17) is presented on the right side, with the approximate location of *bmp2a* and *bmp4* genes and the estimated distance between both genes.

Protein-coding sequences were found to be remarkably conserved among the four genes, exhibiting the same number of coding exons (2), the same phase of intron insertion (1), a similar length (ranging from 1161 to 1251 nt), and sharing a high sequence identity (ranging from 52.2 to 65.7% identity over total CDS length). Non-coding sequences were however quite variable both in size and in structure. While absent in *bmp2a*, a 5´-non-coding exon was observed in *bmp2b*, *bmp4* and *bmp16*. In *bmp4*, an alternative spliced transcript containing two 5'-non-coding exons was also identified (evidence supported by several ESTs). The size of 3'-untranslated regions and the number of consensus sites for polyadenylation signals (predicted with different confidence intervals) were also different among the four genes. Differences in untranslated regions could be indicative of different mechanisms of regulation at transcriptional and posttranscriptional levels.

Similarly, the primary structures of zebrafish Bmp2a, Bmp2b, Bmp4 and Bmp16 mature peptides were compared (Fig. 4.2.5A) and found to be remarkably conserved. Sequence identity ranged from 62.9 to 84.7% and sequence similarity ranged from 74.1 to 92.4%, with Bmp2a and Bmp2b being the most similar, and Bmp4 and Bmp16 being the most dissimilar. These observations suggest that Bmp2a, Bmp2b, Bmp4 and Bmp16 may

share a similar 3D structure and a similar function. Tridimensional models of the four zebrafish proteins were built based on the human BMP2 structure 1REW available in the Protein Data Bank (Fig. 4.2.5B).

a

```
                                          v v x                      v vv
DreBMP2a   ---QARNNKQRKKHKANCRRHSLYVDFSDVGWNDWIVAPPGYHAFYCQGECPFPLADHLN 57
DreBMP2b   QARRGQKPRRKHHQRSNCRRHALYVDFSDVGWNEWIVAPPGYHAFYCHGECPFPLPDHLN 60
DreBMP4    -----SPKQRGRKRNRNCRRHALYVDFSDVGWNDWIVAPPGYQAYYCHGECPFPLADHLN 55
DreBMP16   --NNRRAAKLKRLSRARCRRHPLYVDFKDVGWNKWIIAPSGYDAFFCLGECRFPLADHMN 58
HsaBMP2    QAKH---KQRKRL-KSSCKRHPLYVDFSDVGWNDWIVAPPGYHAFYCHGECPFPLADHLN 45
                         *:** *****.*****.**:** **.*::* *** *** **:*

             o         v                      o v
DreBMP2a   STNHAIVQTLVNSVNSNIPRACCVPTDLSPVSLLYLDEYERVILKNYQDMVVEGCGCR  115
DreBMP2b   STNHAIVQTLVNSVNSNIPKACCIPTELSPISLLYLDEYEKVILKNYQDMVVEGCGCR  118
DreBMP4    STNHAIVQTLVNSVNTNIPKACCVPTELSAISMLYLDETDRVVLKNYQEMVVEGCGCR  113
DreBMP16   SSSHAMVQTLVSSVNGAVPRPCCVPTALSPIALLFLDQEERVVLKNYQDMVVEGCGCR  116
HsaBMP2    STNHAIVQTLVNSVNSKIPKACCVPTELSAISMLYLDENEKVVLKNYQDMVVEGCGCR  103
           *:.**:*****.***  :*: **:** ** :::*:**: ::*:*****:*********
```

b

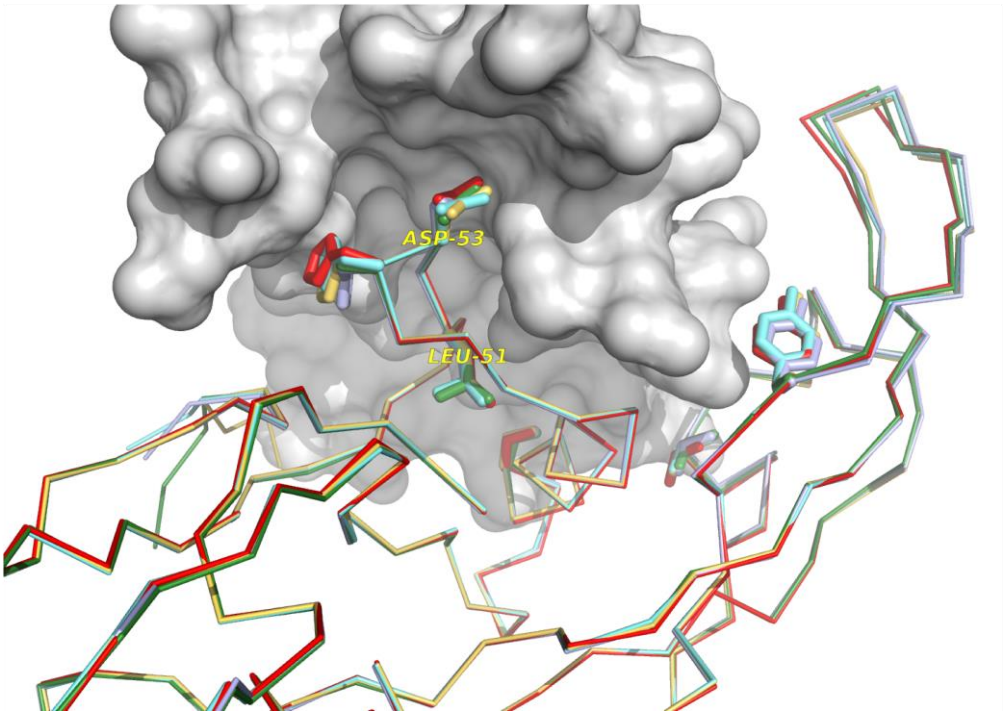


Figure 4.2.5. Structures of zebrafish Bmp2a, Bmp2b, Bmp4 and Bmp16 and human Bmp2. **(a)** Alignment of the primary structure of zebrafish and human mature peptides. Positions marked with letters are residues of human BMP2 withing 3.0 Å of the BMP receptor IA (BMPRIA). *v*, fully conserved; *o*, conservative replacement; *x*, non-conservative replacement. Amino acids involved in solvent-mediated interactions are marked in *bold*; amino acids involved in receptor binding (hot spots) are *underlined*. *, *:* and . indicate positions in the alignment with total conservation, conserved substitution and non-conserved substitution, respectively. **(b)** Superposition of zebrafish mature Bmp2a (gold), Bmp2b (cyan), Bmp4 (green) and Bmp16 (violet) protein models with human BMP2 (red). BMP residues important for interface contact are displayed as sticks, and the two hot spot residues ASP53 and LEU51 are labeled. The molecular surface of the BMPRIA receptor is displayed in white.

Given the high similarity between zebrafish proteins and human BMP2 (ranging from 64.7 to 83.3% of sequence identity and from 78.4 to 90.4% of sequence similarity, Bmp4 being the most similar and Bmp16 the most dissimilar), homology models were expected to be very reliable and this was confirmed with various assessment scores: the four models have QMEAN and Z-DOPE scores within near nativeness (see material and methods and Supplementary Table 4.2.1 for details). The four zebrafish protein models overlapped completely, evidencing their high structural conservation and further suggesting that the correspondent proteins may have the same function, e.g. signal transduction through binding to surface receptors. Indeed, seven of the ten residues of the human BMP2 localized at a 3Å distance from the BMPRIA (Fig. 4.2.5A), including the binding hotspots Leu51 and Asp53, were found to be fully conserved among zebrafish Bmp2a, Bmp2b, Bmp4 and Bmp16, suggesting that members of the BMP2/4/16 subfamily may share the capacity to activate the same receptors. Regarding the important residues for receptor binding in the human BMP2, BMP16 is the most divergent protein, suggesting that it may not bind to BMP receptor(s) with the same affinity.

4.2.4.4 Zebrafish Bmp2a, Bmp2b, Bmp4 and Bmp16 are activators of the BMP-signaling pathway

The capacity of zebrafish proteins to activate the BMP-signaling pathway was assessed through the use of the BRE-Luc system, where BMP responsive elements (BRE) drive the expression of the firefly luciferase reporter gene. Assays were initially performed in the zebrafish ZFB1 cell line (Vijayakumar et al., 2013), but the low transfectability of these cells resulted in low levels of luciferase activity - close to background levels - precluding the use of zebrafish bone-derived cells (results not shown). To maintain a certain homogeneity in our experimental system, gilthead seabream mineralogenic cell line ABSa15 – of fish origin and previously used for this purpose (Tiago et al., 2014) – was alternatively used to perform the assays. Relative luciferase activity remained low upon co-transfection of the BRE-Luc vector with the empty expression vector, but higher than background levels (determined by the promoter-less pGL3 basic), suggesting that endogenous BMPs produced by ABSa15 cells were capable of activating BMP signaling pathway. Each of the four zebrafish proteins strongly and significantly activated the BMP signaling pathway, although to different extent (Fig. 4.2.6). Bmp2a was the more effective (14.1 folds) followed by Bmp2b, Bmp4 and Bmp16 (8.1, 7.5 and 6.5 folds, respectively). Although values were not significantly different

for those last three proteins, Bmp16 exhibited the lower activation capacity. Whether this is correlated with the higher divergence of Bmp16 for the residues involved in receptor binding remains to be determined.

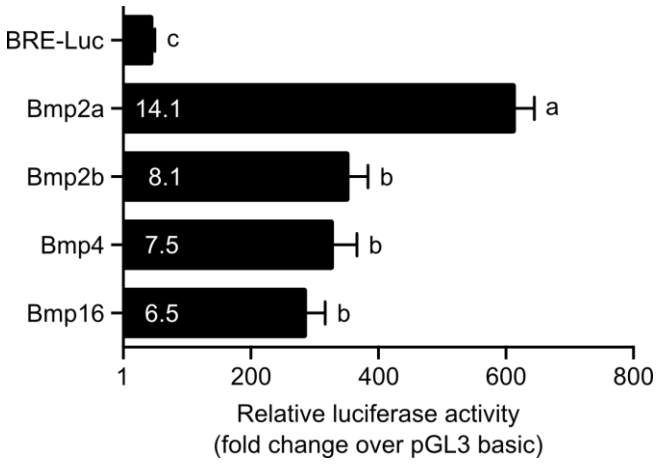


Figure 4.2.6. Activation of BMP-signaling pathway by zebrafish Bmp2a, Bmp2b, Bmp4 and Bmp16. ABSa15 cells were co-transfected with BRE-Luc reporter vector, containing BMP-responsive elements upstream the luciferase gene and vectors containing each of the zebrafish *bmp* genes. *Numbers* inside the bars indicate fold changes over BRE-Luc vector. The different *letters* indicate values significantly different from each other (one-way ANOVA followed by Tukey's post-test; $P<0.05$).

4.2.4.5 Expression of bmp2a, bmp2b, bmp4 and bmp16 during zebrafish larval development and in adult tissues

To better understand spatial-temporal expression patterns of *bmp2a*, *bmp2b*, *bmp4* and *bmp16*, transcript levels were determined by qPCR throughout larval development and in adult zebrafish tissues. Since housekeeping genes used to normalize the expression in qPCR showed some variation during early stages of zebrafish development (i.e. at 1 and 3 hpf), the raw Ct values will be considered as a measure of gene expression in those samples. Comparative analysis of Ct values (Fig. 4.2.7A) and relative gene expression (Fig. 4.2.7B) for the 4 genes revealed limited variations of transcript levels in fish older than 16 hpf. During this developmental window, a slight increase of *bmp2a* and *bmp2b* expression was observed after hatching and at late larval development; patterns of *bmp4* and *bmp16* expression were inversely related, *bmp4* being more expressed from 16 hpf to 7 dpf and *bmp16* being more expressed from 5 dpf to 30 dpf. The most striking differences however occurred during early embryonic development (1-3 hpf), where expression of *bmp* genes appeared to be sequentially switched on (Fig. 4.2.7). While *bmp16* transcript was detected since 1 hpf (4-cells stage) at levels similar to those observed in subsequent stages, suggesting that it may be maternally inherited, expression of *bmp2a* and *bmp4* remained extremely low

(high Ct values) at both stages, and significant levels were only detected at 16 hpf (*bmp2a* and *bmp4*) and 3 hpf (*bmp2b*).

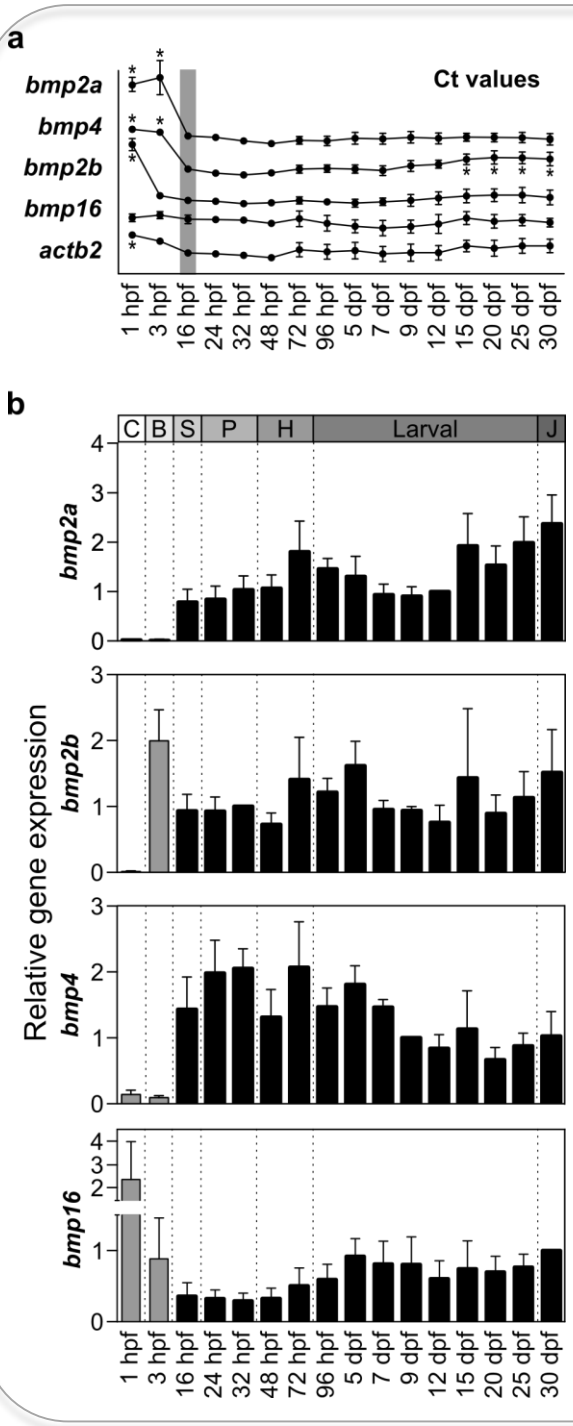


Figure 4.2.7. Expression of *bmp2a*, *bmp2b*, *bmp4* and *bmp16* throughout zebrafish larval development. **(a)** Ct values for *bmp2a*, *bmp2b*, *bmp4*, *bmp16* and *act2b* genes. Values are the mean from at least 3 technical replicates ± standard deviation. Asterisks indicate values that are significantly different from the values at 16 hpf (one-way ANOVA followed by Dunnett's post-test; $P<0.05$). **(b)** Transcript levels were determined by qPCR from at least three technical replicates and normalized using housekeeping *β-actin2* gene. Gray bars indicate initial stages of development, where the expression of housekeeping gene is not constant. Expression levels at 12 dpf (*bmp2a*), 32 hpf (*bmp2b*), 9 dpf (*bmp4*) and 30 dpf (*bmp16*) were used as references and set to 1. *hpf*, hours post-fertilization; *dpf*, days post-fertilization. Different relevant developmental processes are indicated on the top of the figure: *C*, cleavage; *B*, blastula; *S*, segmentation; *P*, pharyngula; *H*, hatching; *J*, juvenile.

Pattern of *bmp2b* expression was intermediate, its transcript being absent or poorly expressed at 1 hpf but present at 3 hpf (blastula stage) at levels similar to those observed in subsequent stages. Differences in the temporal expression of *bmp2a*, *bmp2b*, *bmp4* and

bmp16 strongly suggest that each isoform plays a different role in early embryonic development.

Comparative analysis of the distribution of *bmp2a*, *bmp2b*, *bmp4* and *bmp16* expression in adult tissues (spatial expression) reveals that all isoforms were expressed in both soft and calcified tissues. The highest levels were found in calcified tissues, scale being particularly rich in *bmp2b*, *bmp4* and *bmp16* transcripts and gills (including branchial arches) rich in *bmp2a* transcripts **(**Fig. 4.2.8**)**.

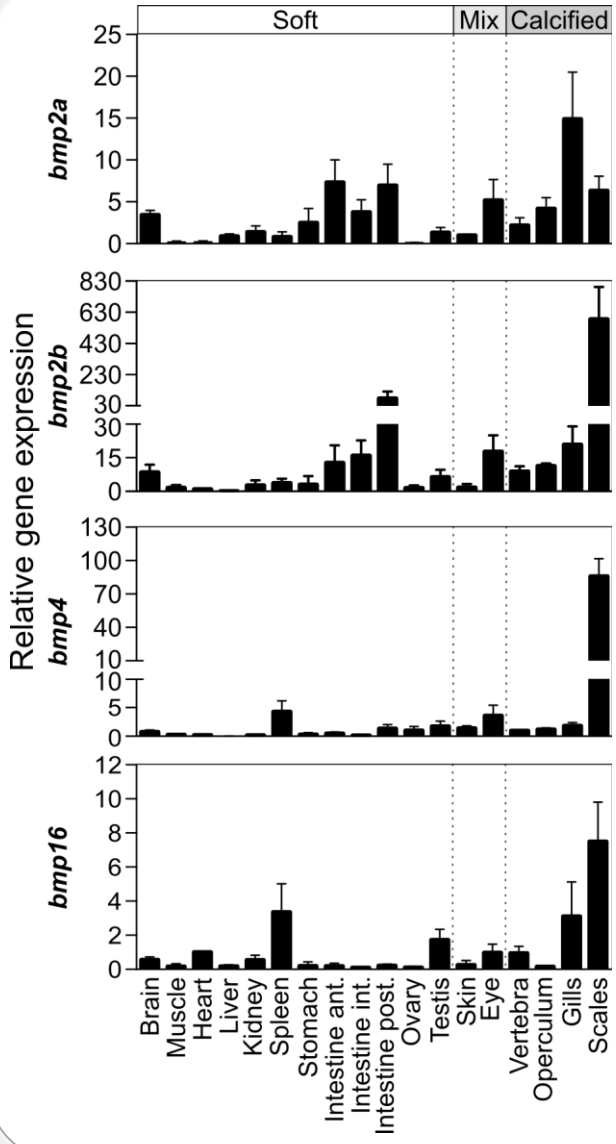


Figure 4.2.8. Relative gene expression of zebrafish *bone morphogenetic proteins 2a*, *2b*, *4* and *16* genes in adult zebrafish tissues. Transcript levels were determined by qPCR from at least three technical replicates and normalized using housekeeping *rps18* gene. Expression levels in skin (*bmp2a*), heart (*bmp2b* and *bmp16*), and vertebra (*bmp4*) were used as references and set to 1. The type of tissue is indicated on top of the figure.

High and intermediate levels of expression were also observed in specific soft tissues: intestines and brain for *bmp2a* and *bmp2b*, spleen for *bmp4* and *bmp16*, and testis for *bmp16*. Eye tissue, which is a mix of soft and calcified (e.g. sclerotic cartilage) tissues, was also positive for all transcripts, although the contribution of each type of tissue remains to be

determined. Spatial expression of *bmp2a*, *bmp2b*, *bmp4* and *bmp16* strongly suggests that each isoform plays a different role in organ and tissue homeostasis, with a central role in calcified tissues, and particularly in scales.

4.2.4.6 Retinoic acid negatively regulates the expression of zebrafish bmp2a, bmp2b, bmp4 and bmp16

Expression of all *bmp* genes was negatively regulated upon a*t*RA treatment, although to different extents (Fig. 4.2.9). The two *bmp2* genes suffered the highest inhibition, *bmp2a* being most affected (5.8 folds), while expression of other *bmp* genes was down-regulated to a lesser extent. *Bmp4* was only slightly down-regulated (1.6 folds), suggesting that although regulatory mechanisms driven by RA were conserved, they may have distinct gene specificities.

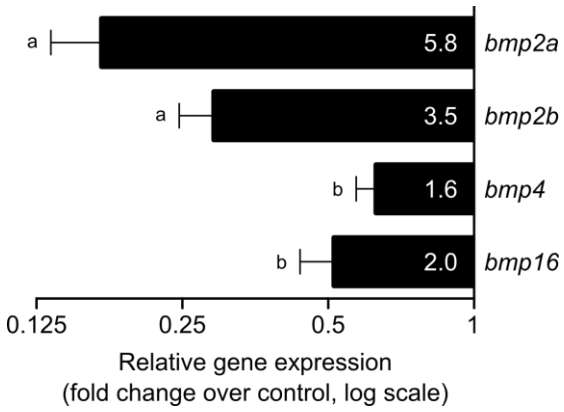


Fig. 9 Relative expression of *bmp2a*, *bmp2b*, *bmp4* and *bmp16* in zebrafish ZFB1 cells upon exposure to 1 μM of all-*trans* retinoic acid during 24 h. Transcript levels were determined by qPCR from at least three technical replicates, normalized using housekeeping *β-actin2* gene and presented as fold change over control (cells treated with DMSO, vehicle for retinoic acid). *Numbers* inside the bars indicate fold changes over control. The different *letters* indicate values significantly different from each other (one-way ANOVA followed by Tukey's post-test; $P<0.05$).

4.2.5 Discussion

Taxonomic and phylogenetic data revealed the complex evolution of BMP2/4/16 family, where members are derived from an ancestral BMP2/4/16 isoform through gene duplication events that occurred during vertebrate evolution and were subsequently independently lost in specific animal lineages. Fig. 4.2.10 illustrates this complex evolutionary relationship and according to the molecular phylogeny presented here, the origin of BMP16 would precede the appearance of BMP2 and BMP4 in Chondrichthyes and of BMP2/4 in lamprey. Given the phylogenetic relationships of taxa possessing BMP16, and taking into account the two WGD events that are known to have occurred early in vertebrate evolution (before and after the agnathans/gnathostomes split; Sharman and Holland, 1998;

Escriva et al., 2002; Panopoulou and Poustka, 2005), the most parsimonious explanation for the origin of BMP16 is that the gene is the product of the first WGD event that occurred in a chordate ancestor prior to the branching of the jawless fish. Our proposal contradicts the recent report by Feiner et al. (Feiner et al., 2009), but is congruent with the proposal that the three lamprey BMP2/4 genes are the result of lineage specific duplications after the divergence of agnathans and gnathostomes (McCauley and Bronner-Fraser, 2004). From an ancestral BMP2/4/16 isoform the first WGD event produced both the BMP16 and the BMP2/4 isoforms, the latter of which, after the second WGD event, gave rise to BMP2 and BMP4 isoforms (Fig. 4.2.10).

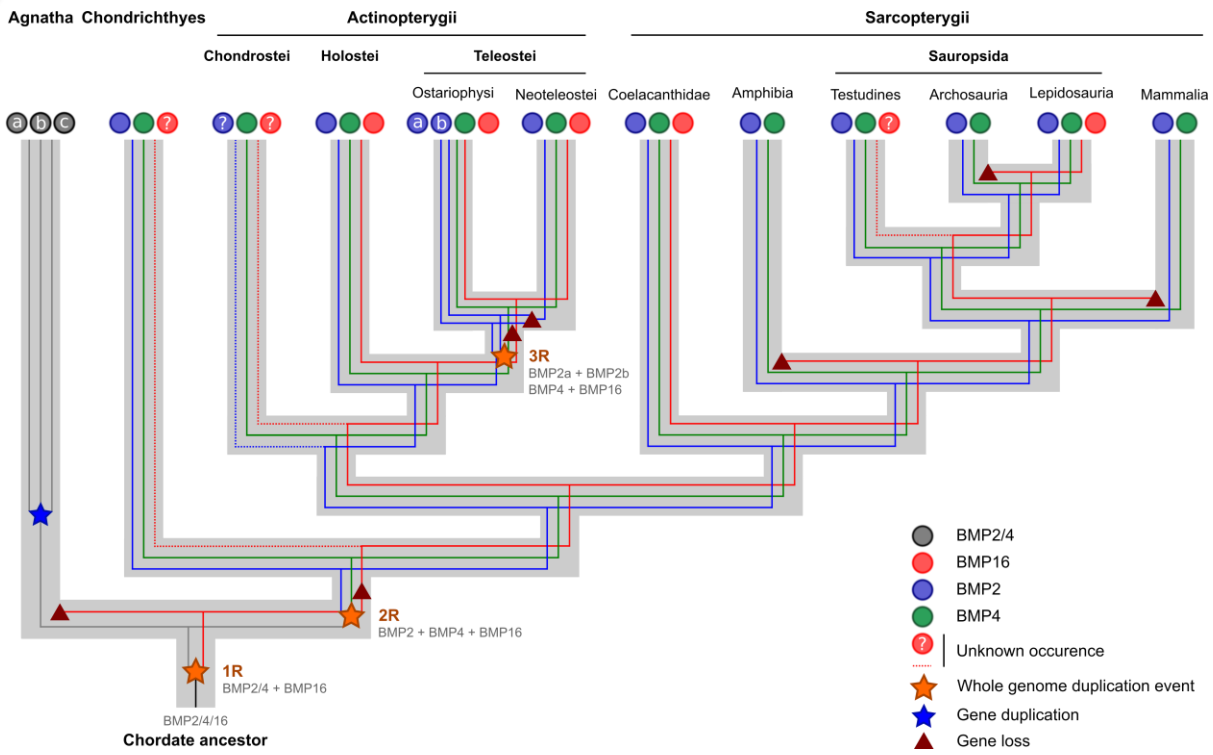


Figure 4.2.10. Schematic representation of a possible evolutionary model for members of the BMP2/4/16 subfamily. Members of this subfamily are products of the same ancestor BMP2/4/16 gene that, after the two whole genome duplication events, that affected vertebrates, originated BMP2, BMP4 and BMP16 genes. Ostariophysi-specific BMP2a isoform resulted from the teleost fish specific whole genomic duplication event and would have been lost in the lineage leading to modern teleosts (Neoteleostei). Absence of BMP16 in some vertebrate taxa is explained by lineage-specific gene loss.

The presence of BMP16 in two lepidosaurian species (the green anole and the Burmese python) and its absence in other tetrapods was unexpected. If not present in these two species, BMP16 could have been lost early in the tetrapod lineage after branching from the lobe-finned fish, but its presence in lepidosaurians suggests multiple, independent and lineage-

specific losses of the gene (Fig. 4.2.10). Similarly to what happens for BMP16, parallel lineage-specific gene losses have been reported for other genes that were present at the base of the vertebrate lineage and have been lost in selected taxa throughout evolution (Yang, 2013; Pasquier et al., 2014; Castro et al., 2012; Davit-Béal et al., 2009). Gene retention is often associated with adaptive advantages (Cuypers and Hogeweg, 2014; van Hoek and Hogeweg, 2009) and the maintenance of BMP16 in Lepidousaria genomes is most likely due to a selective advantage related to specific traits (e.g. locomotion, reproduction, feeding, adaptation to a particular environment) that the gene confers (see below for our hypothesis of the adaptive advantage promoted by BMP16). Although the absence of BMP16 in agnathans and cartilaginous fish could also be related to gene loss events, we believe that limited amount of genomic and transcriptomic information for these taxonomic groups are probably the cause for this absence.

The presence of two BMP2 isoforms (BMP2a and BMP2b) in Ostariophysi could be related to a gene duplication event that occurred in an ancestor soon after branching from Neoteleostei or to the third, fish-specific, WGD that affected Teleostei (Santini et al., 2009; Gates et al., 1999; Taylor et al., 2003; Taylor et al., 2001; Sato and Nishida, 2010). If teleost-specific WGD is at the origin of the second BMP2 isoform in Ostariophysi, its absence in Neoteleostei probably occurred through gene loss and is possibly related to an eventual functional redundancy. Future studies should aim at understanding why a second isoform was maintained in this specific taxonomic group and whether it evolved a new function.

The low conservation of the genomic region neighboring BMP16 locus in zebrafish was somehow a bit surprising given the high conservation observed among other teleost fish. Interestingly, a high number of transposable elements have been detected in zebrafish genome and we propose that divergent genomic structure around BMP16 genes could be related to a higher frequency of interchromosomal gene exchange (Howe et al., 2013; Jaillon et al., 2004).

BMPs form the largest group of growth factors in TGFβ superfamily and its division into subfamilies was based on sequence identity (Kingsley, 1994). Conservation of gene structure between members of the BMP2/4/16 subfamily is particularly evident in the coding regions and high similarity of the mature peptide has been reported previously (Marques et al., 2014; Gates et al., 1999). However, non-coding regions, which are known to have important roles in gene transcriptional and post-transcriptional regulation (Barrett et al., 2013), were found to be substantially different in zebrafish *bmp2a*, *bmp2b*, *bmp4* and *bmp16*, and may indicate differences in the regulation of gene expression. The lethal phenotypes of

BMP2$^{-/-}$ or BMP4$^{-/-}$ mice demonstrated that both genes are essential and cannot compensate for the deficient function of each other (Zhang and Bradley, 1996; Winnier et al., 1995). It suggest that spatial-temporal patterns of BMP2 and BMP4 gene expression are not overlapping and that both genes are probably submitted to distinct regulatory mechanisms and be involved in separate physiological roles. This hypothesis was confirmed by expression data presented in this study but also by the high sequence divergence of BMP2 and BMP4 promoter and untranslated regions (Fritz et al., 2006; Helvering et al., 2000; Feng et al., 1994).

On the contrary, protein sequences were remarkably conserved and exhibited an almost identical 3D structure, where most residues involved in interface contact, including the BMP receptor binding hotspots, Leu51 and Asp53 (Keller et al., 2004), were conserved. Leu51 is maintained unvariable in several members of BMPs and growth differentiation factors (GDFs), all of which interact with BMPR1A, indicating that the hydrogen bond formed between this residue and the receptor is important for ligand-receptor interaction (Keller et al., 2004). Similarly, Leu51 and Asp53 residues in BMP7 were shown to have extensive contact with noggin, a known BMP antagonist, suggesting that these residues work as a general recognition motif in BMP ligands, although they are not always the main determinants (Keller et al., 2004). These data suggest that BMP2/4/16 subfamily members may be involved in similar and/or complementary processes and act through the same signaling pathways, as proposed by Feiner and co-workers (Feiner et al., 2009). The main differences between BMP2/4/16 family members would probably not be associated to its structure and ability to trigger BMP signaling but most likely related to differential regulation and different patterns of expression.

Bmp2a, Bmp2b, Bmp4 and Bmp16 were all capable of activating the BMP signaling pathway, although to different extents. Even though we cannot exclude the possibility that the differential activation of BMP signaling pathway may result from the transfection of variable amounts of DNA into the host cells or from uneven capacity of the cells to produce and/or process the different proteins, we suggest that it is a consequence of distinct receptor affinities as already proposed for other BMPs (Upton et al., 2008). Binding assays aiming at evaluating ligand-receptor affinity will need to be performed in the future to address this question.

Although they exhibited different levels of gene expression, zebrafish *bmp2a* and *bmp2b* showed comparable expression patterns in adult tissues, while different during early development. In zebrafish and Mexican tetra (both Ostariophysi) *bmp2a* expression is

detected in a small subset of *bmp2b* expression domains fact that is consistent with a possible subfunctionalization that genes may have experienced after the duplication (Wise and Stock, 2006). Developmental expression of BMP4 is reported to be an important signal for organ morphogenesis in several vertebrates (Chin et al., 1997; Jones et al., 1991; Fainsod et al., 1994; Bellusci et al., 1996). *Bmp4* is also expressed throughout zebrafish development and similarly to *bmp2* it is described to act as a ventralizing agent, during mesoderm and neural plate formation (Wilson and Hemmati-Brivanlou, 1995), a role that is maintained by their invertebrate orthologs (Hwang et al., 2003), suggesting a functional conservation of members of the subfamily throughout evolution.

While *bmp2a*, *bmp2b* and *bmp4* were all expressed at later stages, *bmp16* expression was detected as soon as 1 hpf (4-cells), in zebrafish, and 2 hpf, in Senegalese sole (Marques et al., 2014), suggesting a possible maternal inheritance of the transcript and an important role in the early stages of embryonic development. Analysis of sites of *bmp16* expression revealed that until 5 dpf transcripts were mainly detected in the developing heart, gut epithelium and swim bladder (Feiner et al., 2009). In Senegalese sole, expression of BMP16 remained very low until 5 dpf, exhibiting an increase during the metamorphosis phase (Marques et al., 2014), which may indicate a particular role of BMP16 in this process.Tissue distribution of *bmp2a*, *bmp2b*, *bmp4* and *bmp16* indicate high expression levels for all these isoforms in calcified tissues, supporting the well documented role of *bmp2* and *bmp4* in bone metabolism (see for example (Nishimura et al., 2012)), and suggesting a contribution from *bmp16* in this process. For most of the isoforms (except *bmp2a*), highest expression levels were observed in scales, in agreement with the reported role of BMP signaling in the formation and regeneration of fish and Squamata scales (Zou and Niswander, 1996; Harris et al., 2002). High expression in scales, together with the observation that BMP16 is only present in organisms exhibiting scales on their body (i.e. ray-finned fish, lobe-finned fish, lizards and snakes), indicate that BMP16 may play an important role in scale formation and maintenance. Among other tissues with high *bmp2*, *4* and *16* expression are the branchial arches, known to be important for processes such as osmoregulation and respiration, where BMP signaling is reported to be involved (Kültz, 2012). *Bmp16* expression in this tissue may also suggest a role in cartilage formation and mineralization as already described for BMP2 and BMP4 (Alexander et al., 2011). The relatively high levels of *bmp2*, *bmp4* and *bmp16* expression observed in the eye, which contain a mixture of soft and calcified tissues, may be related to the presence of sclerotic cartilage, in agreement with previous reports for other BMP family ligands (Wordinger and Clark, 2007). In addition to bone metabolism and

dorsoventral patterning, BMPs are described as being involved in several other processes (Hogan, 1996). Important roles of BMP2 in the central and, particularly, in the enteric nervous system formation are well documented (Chalazonitis and Kessler, 2012; Sato et al., 2010; Sailer et al., 2005). The high levels of *bmp2a* and *bmp2b* in brain and intestine samples further support this interpretation. Moderate levels of expression are also found in spleen (*bmp4* and *bmp16*) and in testis (*bmp16*), and BMP4 has been described as one of the signals required for the expansion of stress erythroid progenitors, in murine spleen (Perry et al., 2007), and its expression has also reported in the gonads of several vertebrate species (Shimasaki et al., 2004; Li and Ge, 2011).

Exposure of ZFB1 cells to RA demonstrated a negative regulation of zebrafish *bmp2*a, *bmp2b*, *bmp4* and *bmp16* expression. RA is a morphogen involved in several developmental processes and in skeletal formation (Thompson et al., 2003). Interaction of RA with the BMP signaling as well as the capacity to regulate BMPs expression has been reported, however opposing effects were observed. Expression of BMP2 was stimulated in HSG-S8 cells, a human adenocarcinoma cell line (Hatakeyama et al., 1996), and we have shown that RA up-regulates the expression of BMP2, BMP4 and BMP16 in Senegalese sole cells (Marques et al., 2014). On the other hand, RA was described as down-regulating the expression of BMP7 in rats with cleft palate (Guo et al., 2008), and expression of BMP2 and BMP4 was also observed to decrease in MG63 cells after RA treatment (Virdi et al., 1998). Contradictory results for the RA regulation of BMP genes suggest a context-dependent effect of the morphogen and is most likely related to the presence/absence of co-regulators (Simandi and Nagy, 2011; Waxman and Yelon, 2011; Grimsrud et al., 1998).

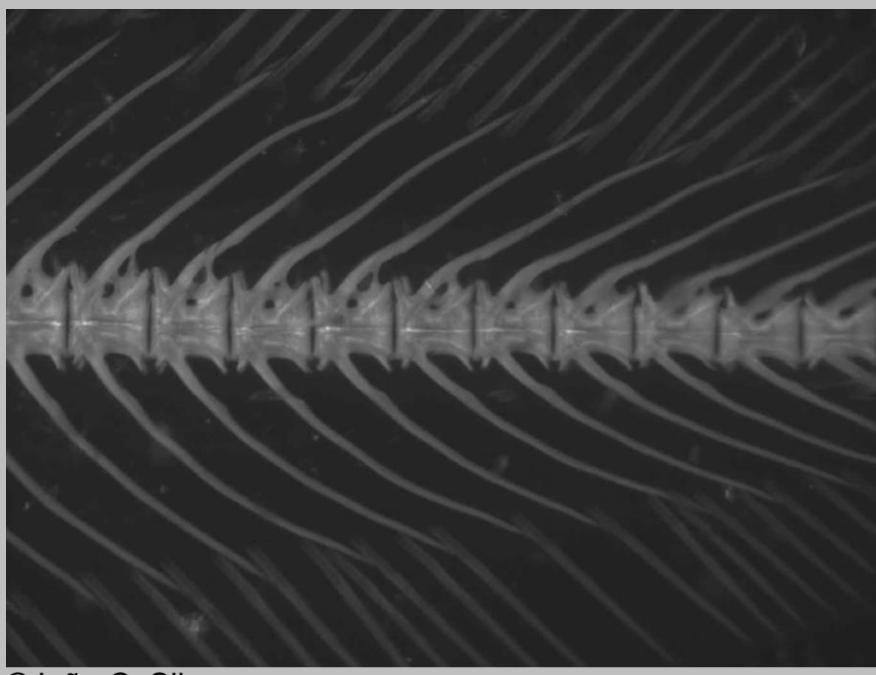
In conclusion, we have shown that BMP16 is not restricted to the teleost fish lineage but is largely absent from tetrapod genomes. Phylogenetically, BMP16 diverged in early vertebrates from an ancestral BMP2/4/16. All family members have a protein structure remarkably similar and are capable of activating the BMP signaling pathway. They would therefore perform the same function.. Differences among BMP2/4/16 family members are found in the spatial-temporal expression of the genes. They would therefore be submitted to different regulation and participate in distinct physiological processes during early embryonic development and in adult tissues, scales being the tissue expressing the highest levels of *bmp2b*, *bmp4* and *bmp16* expression. The presence of BMP16 in lepidosaurians, while it is absent in other tetrapods, remains to be elucidated although a role in scale formation and homeostasis is conceivable.

4.2.6 Acknowledgments

We thank to Dr. Peter ten Dijke (Leiden University Medical Center, Leiden, The Netherlands) that kindly provided the BMP-responsive luciferase reporter vector (BRE-Luc). 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. It was also partially financed by the European Community (EC) through ASSEMBLE (FP7/227799) research project. CM and JR were supported by doctoral grants (SFRH/BD/39964/2007 and SFRH/BD/47433/2008, respectively) from the FCT. IF was supported by a post-doctoral grant (SFRH/BPD/82049/2011) from the FCT.

Chapter 5

General Conclusions and Future Perspectives



©João C. Silva

CHAPTER 5. GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES

We have collected, within the scope of the work presented here, novel evidences on mechanisms regulating BMP2 gene expression but also data on BMP2 gene evolution and protein structure through comparative analysis with the other members of the BMP2/4/16 subfamily. This new data has been almost entirely collected using fish tools, e.g. gene, transcript and protein sequences, mineralogenic cell lines, RNA samples from developmental stages and adult tissues, transcription factor expressing vectors and microRNA. It is worth to note that tools developed for mammalian systems (e.g. BRE-Luc vector carrying mammalian BMP responsive elements, pcDNA3/pCMX vectors driving gene expression through the human CMV promoter and HEK293 cell line as an host for fish sequence-based vectors) were functional in fish systems or when used with fish tools, supporting previous data showing a conservation of cellular machinery and regulatory mechanisms, in particular those related to bone morphogenetic proteins (e.g. receptors, pathway intermediates, TF responsive elements), between mammals and fish. Because BMP2 gene and protein structures are also very much similar between species in taxa separated by hundreds of million years of evolution, we believe that insights provided here from data collected in fish systems should be valid in mammals, although this remains to be confirmed.

Regulation of gilthead seabream BMP2 gene expression was first addressed at transcriptional level through the analysis of the corresponding gene 5' flanking region, which supposedly contains the promoter and *cis*-regulatory elements. Simple and extremely well conserved throughout evolution, the structure of BMP2 gene includes a distinctive feature – a 5' non-coding exon – that may be important for transcriptional regulation. Such feature has been associated, in other genes, with the presence of enhancer/repressor of the promoter activity in the intronic sequence following the non-coding exon or with alternative promoter usage (Mikio, 2000; Conceição et al., 2008). Additional exon could also be associated with the presence, in the 5'UTR of BMP2 transcripts, of elements necessary to the stability of the messenger RNA. While no alternative BMP2 transcripts have been observed in vertebrate transcriptomes (surveyed for the past 5 years) to evidence a possible alternative promoter usage, we observed a down-regulation of BMP2 promoter activity by intron I suggesting that it may function as a silencer of BMP2 transcription. This hypothesis should be further investigated using deletion constructs and/or through site-directed mutagenesis of the

potential regulatory elements present in intron I. Binding elements of several transcription factors involved in bone and cartilage metabolism, i.e. RUNX, MEF2, SOX9 and ETS1, were identified in the promoter region of gilthead seabream BMP2 gene and some of these elements were shown to be functional. RUNX3 was identified as a negative regulator of seabream BMP2 transcription and CBFß, a co-factor of RUNX family members, enhanced this effect. MEF2C also down-regulated the transcriptional activity of BMP2 promoter in a SOX9-dependent manner. Future studies should aim at localizing precisely active binding elements for RUNX3 and MEF2C, e.g. through site-directed mutagenesis, and further characterize the cooperative effect of CBFß and SOX9, e.g. through electrophoretic mobility shift assay (EMSA) and/or chromatin immunoprecipitation (ChIP). ETS1 was identified as a mild enhancer of seabream BMP2 transcription and although RUNX factors are known to cooperate with ETS1 through the formation of heterodimers, RUNX3 failed to enhance the up-regulatory effect of ETS1. Future work should evaluate the effect of other members of the RUNX family, in particular RUNX2, but should also aim at identifying other factors that could cooperate with ETS1. In silico analysis of gilthead seabream BMP2 promoter revealed the presence of several binding elements for retinoic acid receptor (RAR) and based on *in vitro* data presented here and demonstrating an effect of retinoic acid on BMP2 gene expression, we propose to further test the possibility of retinoic acid being a transcriptional regulator of BMP2.

Regulation of gilthead seabream BMP2 gene expression was also addressed at a post-transcriptional level, through the analysis of transcript 3'UTR, which normally contains binding sites for microRNAs. A conserved binding site for miR-20a was identified and its functionality as a negative regulator of BMP2 transcript, and therefore BMP signaling pathway, was confirmed. Interestingly, mammalian type II BMP receptor was also reported to be under the control of miR-20a and future work should aim at determining whether fish ortholog of BMPR2 is also targeted by miR20a but also at evaluating whether other intermediates of BMP signaling would be regulated similarly. Whether BMP2 transcript is under the control of other miRNAs (for example heart- and brain-specific miRNAs considering the central role of BMP2 during the development of the cardiovascular system and neurogenesis) should also be evaluated to get valuable insights into mechanisms of BMP2 post-transcriptional regulation.

The interaction between BMP2 and MGP was evaluated through the use of BMP responsive elements driving the expression of luciferase reporter gene. Although we could validate the use of this approach in our fish system through the activation of BMP signaling

pathway and subsequent increase in luciferase activity by gilthead seabream BMP2, it has not been possible to collect conclusive evidences toward the effective interaction of BMP2 and MGP, possibly due to the poor capacity of our cell system to correctly γ-carboxylate overproduced MGP. To identify the probable cause of this failure, we propose to first investigate the levels of MGP carboxylation through the use of specific staining (e.g. 4-diazobenzene sulfonic acid, DBS) or Gla-specific antibodies. If MGP is shown to be under/un-carboxylated, we propose to increase the carboxylation capacity of our cell system (e.g. using an expression vector carrying the cDNA for seabream γ-glutamyl carboxylase or by supplementing cell culture medium with vitamin K, if this co-factor of reveals to be a limitation) or use an alternative cell system with a higher carboxylation capacity (e.g. liver cells). In order to reduce the overload of the system and the consequent fitness of the cells, because of the high amounts of DNA used in the transfections, the use of purified proteins would also be a suitable alternative. Some work has been initiated to produce large amounts of seabream BMP2 (with a C-terminal polyhistidine tag) in *Escherichia coli* and then purify it through a nickel affinity resin. Production of MGP is however much more complex and chemical synthesis may represent the method of choice since Gla residues can be added during the synthesis. The use of more direct methods based on reconstituted enzymes activity, such as the protein-fragment complementation assay (PCA) could represent a valid alternative to the BRE-Luc system in order to characterize the interaction of BMP2 and MGP and some work has also been initiated to fuse seabream BMP2 (bait) and MGP (prey) to each half of the *Gaussia* luciferase.

Finally, a comparative analysis of the founding members of the BMP2/4/16 subfamily was conducted at various levels. By analyzing the taxonomic distribution of BMP2, BMP4 and BMP16 in several vertebrate taxa, we evidenced for the first time the presence of BMP16 protein in tetrapods (in Squamata or scaled reptiles) and therefore could refute the previously published idea that it was teleost specific (Feiner et al., 2009). A phylogenetic analysis revealed that BMP16 diverged prior to BMP2 and BMP4 (which were shown to be more closely related) and suggested that its origin dates back to an ancient genome duplication (possibly 1R) that occurred early in the vertebrate lineage from a common BMP2/4/16 ancestral gene. Although it remains speculative and would require more genomic data, in particular from species representing early vertebrates, origin of BMP2 and BMP4 could be the result of the 2R whole genome duplication event, which presumably occurred after the split between jawless and jawed vertebrates. The absence of BMP16 in most tetrapods may indicate independent lineage-specific gene loss in several vertebrate taxa. The retention of

BMP16 gene in the genome of Squamata remains to be explained but it could be related to some adaptive advantage conferred by protein function, for example the presence of scales in all organisms that retained BMP16 gene in their genome. While this hypothesis is supported by expression data showing high levels of BMP16 transcript in zebrafish scales, more conclusive evidence e.g. sites of BMP16 gene expression in fish and lizards through *in situ* hybridization or locals of protein accumulation by immunohistochemistry, is needed. A zebrafish mutant line lacking the expression of BMP16 gene (not yet available) would be a valuable tool to investigate the possible role of this protein in scale formation and/or maintenance or, alternatively, to determine what is the role of BMP16 in fish. The striking conservation of BMP2, 4 and 16 protein structures contrasts with the large differences observed in gene expression patterns. Overall protein structure and residues involved in receptor binding have been remarkably conserved among zebrafish BMP2, BMP4 and BMP16, suggesting that all three proteins could function as effectors of BMP signaling as confirmed by luciferase reporter assays. The difference between BMP2, BMP4 and BMP16 could be related to differential gene expression during development and in adult tissues, where stage- and tissue-specific expression of each gene was observed both in Senegalese sole and zebrafish.

Mechanisms of differential regulation of gene expression remain to be determined but are apparently not related to retinoic acid since expression of BMP2, BMP4 and BMP16 gene were similarly affected upon exposure of mineralogenic fish cell lines to this morphogen. The comparative in silico analysis of zebrafish BMP2, BMP4 and BMP16 promoter regions and transcript 3'UTR will certainly allow the identification of *cis*-regulatory elements and microRNA binding sites specific to each of the three genes/transcripts and this would represent a first step toward the identification of gene-specific regulatory mechanisms.

Although much remains to be done, novel and valuable data have been collected within the scope of this work towards a better understanding of BMP2 function and regulation but also towards the evolutionary relationship between BMP2, BMP4 and BMP16. Ultimately this work contributed to further validate the use of fish as an alternative to mammalian models to investigate molecular mechanisms of bone metabolism.

GENERAL REFERENCES

Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* 38:W7–W13.

Abd-El-Barr MM, Cox JB, Antonucci MU, Bennett J, Murad GJ, Pincus DW (2011) Recombinant human bone morphogenetic protein-2 as an adjunct for spine fusion in a pediatric population. *Pediatr. Neurosurg.* 47:266–271.

Abdelwahid E, Rice D, Pelliniemi LJ, Jokinen E (2014) Overlapping and differential localization of Bmp-2, Bmp-4, Msx-2 and apoptosis in the endocardial cushion and adjacent tissues of the developing mouse heart. *Cell Tissue Res.* 305:67–78.

Abrams KL, Xu J, Nativelle-Serpentini C, Dabirshahsahebi S, Rogers MB (2004) An evolutionary and molecular analysis of Bmp2 expression. *J. Biol. Chem.* 279:15916–15928.

Agarwal P, Verzi MP, Nguyen T, Hu J, Ehlers ML, McCulley DJ, et al. (2011) The MADS box transcription factor MEF2C regulates melanocyte development and is a direct transcriptional target and partner of SOX10. *Development* 138:2555–2565.

Akiyama T, Marqués G, Wharton KA (2012) A large bioactive BMP ligand with distinct signaling properties is produced by alternative proconvertase processing. *Sci. Signal.* 5:ra28.

Alam N, St-Arnaud R, Lauzier D, Rosen V, Hamdy RC (2009) Are endogenous BMPs necessary for bone healing during distraction osteogenesis? *Clin. Orthop. Relat. Res.* 467:3190–3198.

Alexander C, Zuniga E, Blitz IL, Wada N, Le Pabic P, Javidan Y, et al. (2011) Combinatorial roles for BMPs and Endothelin 1 in patterning the dorsal-ventral axis of the craniofacial skeleton. *Development* 138:5135–5146.

Allendorph GP, Vale WW, Choe S (2006) Structure of the ternary signaling complex of a TGF-β superfamily member. *Proc. Natl. Acad. Sci. U. S. A.* 103:7643–7648.

Alvarez E, Cahoreau C, Combarnous Y (2009) Comparative structure analyses of cystine knot-containing molecules with eight aminoacyl ring including glycoprotein hormones (GPH) alpha and beta subunits and GPH-related A2 (GPA2) and B5 (GPB5) molecules. *Reprod. Biol. Endocrinol.* 7:90–96.

Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, et al. (1998) Zebrafish hox clusters and vertebrate genome evolution. *Science* 282:1711–1714.

Andersen CL, Jensen JL, Ørntoft TF (2004) Normalization of Real-Time Quantitative Reverse Transcription-PCR Data : A Model-Based Variance Estimation Approach to Identify Genes Suited for Normalization , Applied to Bladder and Colon Cancer Data Sets Normalization of Real-Time Quantitative Reverse .

Andrée B, Duprez D, Vorbusch B, Arnold HH, Brand T (1998) BMP-2 induces ectopic expression of cardiac lineage markers and interferes with somite formation in chicken embryos. *Mech. Dev.* 70:119–131.

Aono A, Masatoshi H, Taketomi S, Hideki Y, Tsukuda R, Sasaki S, et al. (1995) Potent ectopic bone-inducing activity of bone morphogenetic protein-4/7 heterodimer. *Biochem. Biophys. Res. Commun.* 210:670–677.

Avsian-Kretchmer O, Hsueh AJW (2004) Comparative genomic analysis of the eight-membered ring cystine knot-containing bone morphogenetic protein antagonists. *Mol. Endocrinol.* 18:1–12.

Bacolla A, Larson JE, Collins JR, Li J, Milosavljevic A, Stenson PD, et al. (2008) Abundance and length of simple repeats in vertebrate genomes are determined by their structural properties. *Genome Res.* 18:1545–1553.

Bais M V, Wigner N, Young M, Toholka R, Graves DT, Morgan EF, et al. (2009) BMP2 is essential for post natal osteogenesis but not for recruitment of osteogenic stem cells. *Bone* 45:254–266.

Baldi P, Baisnee P-F (2000) Sequence analysis by additive scales: DNA structure for sequences and repeats of all lengths. *Bioinformatics* 16:865–889.

Balemans W, Van Hul W (2002) Extracellular regulation of BMP signaling in vertebrates: a cocktail of modulators. *Dev. Biol.* 250:231–250.

Banday AR, Azim S, Tabish M (2012) Differentially expressed three non-coding alternate exons at 5' UTR of regulatory type I beta subunit gene of mouse. *Mol. Biol. Rep.* 39:3375–3383.

Bandyopadhyay A, Tsuji K, Cox K, Harfe BD, Rosen V, Tabin CJ (2006) Genetic analysis of the roles of BMP2, BMP4, and BMP7 in limb patterning and skeletogenesis. *PLoS Genet.* 2:e216.

Barrett LW, Fletcher S, Wilton SD (2013) Untranslated gene regions and other non-coding elements. *SpringerBriefs Biochem. Mol. Biol.* 1:1–57.

Barrett R, Chappell C, Quick M, Fleming A (2006) A rapid, high content, in vivo model of glucocorticoid-induced osteoporosis. *Biotechnol. J.* 1:651–655.

Bartel DP, Lee R, Feinbaum R (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297.

Bax BE, Wozney JM, Ashhurst DE (1999) Bone morphogenetic protein-2 increases the rate of callus formation after fracture of the rabbit tibia. *Calcif. Tissue Int.* 2:83–89.

Beck SE, Jung BH, Fiorino A, Gomez J, Rosario E Del, Cabrera BL, et al. (2006) Bone morphogenetic protein signaling and growth suppression in colon cancer. *Am. J. Physiol. - Gastrointest. Liver Physiol.* 63:135–145.

Bellido T (2014) Osteocyte-driven bone remodeling. *Calcif. Tissue Int.* 94:25–34.

Bellusci S, Henderson R, Winnier G, Oikawa T, Hogan BL (1996) Evidence from normal expression and targeted misexpression that bone morphogenetic protein (Bmp-4) plays a role in mouse embryonic lung morphogenesis. *Development* 122:1693–1702.

Berghmans S, Jette C, Langenau D, Hsu K, Stewart R, Look T, et al. (2005) Making waves in cancer research: new models in the zebrafish. *Biotechniques* 39:227–237.

Bessa PC, Casal M, Reis RL (2008) Bone morphogenetic proteins in tissue engineering: the road from the laboratory to the clinic, part I (basic concepts). *J. Tissue Eng. Regen. Med.* 2:1–13.

Bessa PC, Pedro AJ, Klösch B, Nobre A, van Griensven M, Reis RL, et al. (2008) Osteoinduction in human fat-derived stem cells by recombinant human bone morphogenetic protein-2 produced in Escherichia coli. *Biotechnol. Lett.* 30:15–21.

Beveridge NJ, Tooney PA, Carroll AP, Tran N, Cairns MJ (2009) Down-regulation of miR-17 family expression in response to retinoic acid induced neuronal differentiation. *Cell. Signal.* 21:1837–1845.

Blache P, van de Wetering M, Duluc I, Domon C, Berta P, Freund J-N, et al. (2004) SOX9 is an intestine crypt transcription factor, is regulated by the Wnt pathway, and represses the CDX2 and MUC2 genes. *J. Cell Biol.* 166:37–47.

Blair HC, Zaidi M, Schlesinger PH (2002) Mechanisms balancing skeletal matrix synthesis and degradation. *Biochem. J.* 341:329–341.

Blake T, Adya N, Kim C-H, Oates AC, Zon L, Chitnis A, et al. (2000) Zebrafish homolog of the leukemia gene CBFβ: its expression during embryogenesis and its relationship to scl and gata-1 in hematopoiesis. *Blood* 96:4178–4184.

Boström K, Tsao D, Shen S, Wang Y, Demer LL (2001) Matrix GLA protein modulates differentiation induced by bone morphogenetic protein-2 in C3H10T1/2 cells. *J. Biol. Chem.* 276:14044–14052.

Bragdon B, Moseychuk O, Saldanha S, King D, Julian J, Nohe A (2011) Bone morphogenetic proteins: a critical review. *Cell. Signal.* 23:609–620.

Branam AM, Hoffman GG, Pelegri F, Greenspan DS (2010) Zebrafish chordin-like and chordin are functionally redundant in regulating patterning of the dorsoventral axis. *Dev. Biol.* 341:444–458.

Brock M, Samillan VJ, Trenkmann M, Schwarzwald C, Ulrich S, Gay RE, et al. (2012) AntagomiR directed against miR-20a restores functional BMPR2 signalling and prevents vascular remodelling in hypoxia-induced pulmonary hypertension. *Eur. Heart J.* 1:1–9.

Brock M, Trenkmann M, Gay RE, Michel BA, Gay S, Fischler M, et al. (2009) Interleukin-6 modulates the expression of the bone morphogenic protein receptor type II through a novel STAT3-microRNA cluster 17/92 pathway. *Circ. Res.* 104:1184–1191.

Broege A, Pham L, Jensen ED, Emery A, Huang T-H, Stemig M, et al. (2013) Bone morphogenetic proteins signal via SMAD and mitogen-activated protein (MAP) kinase pathways at distinct times during osteoclastogenesis. *J. Biol. Chem.* 288:37230–37240.

Brown MA, Zhao Q, Baker KA, Naik C, Chen C, Pukac L, et al. (2005) Crystal structure of BMP-9 and functional interactions with pro-region and receptors. *J. Biol. Chem.* 280:25111–25118.

Brunet LJ, McMahon JA, McMahon AP, Harland RM (1998) Noggin, cartilage morphogenesis, and joint formation in the mammalian skeleton. *Science* 280:1455–1457.

Von Bubnoff A, Cho KW (2001) Intracellular BMP signaling regulation in vertebrates: pathway or network? *Dev. Biol.* 239:1–14.

De Caestecker M (2004) The transforming growth factor-β superfamily of receptors. *Cytokine Growth Factor Rev.* 15:1–11.

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

Canalis E, Economides AN, Gazzerro E (2003) Bone morphogenetic proteins, their antagonists, and the skeleton. *Endocr. Rev.* 24:218–235.

Cancela ML, Laizé V, Conceição N (2014) Matrix Gla protein and osteocalcin: From gene duplication to neofunctionalization. *Arch. Biochem. Biophys.* 561:56–63.

Capovilla M, Brand M, Betas J (1994) Direct regulation of decapentaplegic by and its role in Drosophila midgut morphogenesis. *Cell* 76:461–475.

Carragee EJ, Chu G, Rohatgi R, Hurwitz EL, Weiner BK, Yoon ST, et al. (2013) Cancer risk after use of recombinant bone morphogenetic protein-2 for spinal arthrodesis. *J. Bone Joint Surg. Am.* 95:1537–1545.

Carreira AC, Lojudice FH, Halcsik E, Navarro RD, Sogayar MC, Granjeiro JM (2014) Bone morphogenetic proteins: facts, challenges, and future perspectives. *J. Dent. Res.* 93:335–345.

Carter TG, Brar PS, Tolas A, Beirne OR (2008) Off-label use of recombinant human bone morphogenetic protein-2 (rhBMP-2) for reconstruction of mandibular bone defects in humans. *J. Oral Maxillofac. Surg.* 66:1417–1425.

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* 21:2933–2942.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17:540–552.

Castro LFC, Lopes-Marques M, Gonçalves O, Wilson JM (2012) The evolution of pepsinogen C genes in vertebrates: duplication, loss and functional diversification. *PLoS One* 7:e32852.

Chalazonitis A, Kessler JA (2012) Pleiotropic effects of the bone morphogenetic proteins on development of the enteric nervous system. *Dev. Neurobiol.* 72:843–856.

Chandler RL, Chandler KJ, McFarland KA, Mortlock DP (2007) Bmp2 transcription in osteoblast progenitors is regulated by a distant 3' enhancer located 156.3 kilobases from the promoter. *Mol. Cell. Biol.* 27:2934–2951.

Chapman BA, Bowers JE, Schulze SR, Paterson AH (2004) A comparative phylogenetic approach for dating whole genome duplication events. *Bioinformatics* 20:180–185.

Chekmenev DS, Haid C, Kel AE (2005) P-Match: transcription factor binding site search by combining patterns and weight matrices. *Nucleic Acids Res.* 33:W432–W437.

Chen D, Zhao M, Mundy GR (2004) Bone morphogenetic proteins. *Growth factors* 22:233–241.

Chen G, Deng C, Li Y-P (2012) TGF-β and BMP signaling in osteoblast differentiation and bone formation. *Int. J. Biol. Sci.* 8:272–288.

Chen NF, Smith ZA, Stiner E, Armin S, Sheikh H, Khoo LT (2010) Symptomatic ectopic bone formation after off-label use of recombinant human bone morphogenetic protein-2 in transforaminal lumbar interbody fusion. *J. Neurosurg. Spine* 12:40–46.

Chen S, Gluhak-Heinrich J, Martinez M, Li T, Wu Y, Chuang H-H, et al. (2008) Bone morphogenetic protein 2 mediates dentin sialophosphoprotein expression and odontoblast differentiation via NF-Y signaling. *J. Biol. Chem.* 283:19359–19370.

Chen W-C, Zhang Y, Ma D, Ma X-J, Shou W-N, Huang G-Y (2012) Bmp2 regulates the interaction between EPDCs and myocytes in cardiac OFT. *Med. Hypotheses* 79:174–177.

Chen Z, Wu C, Gu W, Klein T, Crawford R, Xiao Y (2014) Osteogenic differentiation of bone marrow MSCs by β-tricalcium phosphate stimulating macrophages via BMP2 signalling pathway. *Biomaterials* 35:1507–1518.

Chin AJ, Chen J-N, Weinberg ES (1997) Bone morphogenetic protein-4 expression characterizes inductive boundaries in organs of developing zebrafish. *Dev. Genes Evol.* 207:107–114.

Choi K-Y, Kim H-J, Lee M-H, Kwon T-G, Nah H-D, Furuichi T, et al. (2005) Runx2 regulates FGF2-induced Bmp2 expression during cranial bone development. *Dev. Dyn.* 233:115–121.

Chomczynski P (1987) Single-step method of RNA isolation by acid guanidinium extraction. *Anal. Biochem.* 162:156–159.

Cochran DL, Wozney JM (1999) Biological mediators for periodontal regeneration. *Periodontol. 2000* 19:40–58.

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.* 215:1983–1993.

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.

Conceição N, Simões B, Cancela ML (2013) Functional analysis of the two Runx3 promoters in osseous and non-osseous cells: implications for tissue/differentiation specific transcription of distinct isoforms. In *Bone Abstracts*. p. 273.

Constam DB, Robertson EJ (1999) Regulation of bone morphogenetic protein activity by pro domains and proprotein convertases. *J. Cell Biol.* 144:139–149.

Crane JL, Cao X (2014) Bone marrow mesenchymal stem cells andTGF-ß signaling in bone remodeling. *J. Clin. Invest.* 124:466–472.

Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res.* 14:1188–1190.

Cui Y, Hackenmiller R, Berg L, Jean F, Nakayama T, Thomas G, et al. (2001) The activity and signaling range of mature BMP-4 is regulated by sequential cleavage at two sites within the prodomain of the precursor. *Genes Dev.* 15:2797–2802.

Cui Y, Jean F, Thomas G, Christian JL (1998) BMP-4 is proteolytically activated by furin and/or PC6 during vertebrate embryonic development. *Embo J.* 17:4735–4743.

Cuypers TD, Hogeweg P (2014) A synergism between adaptive effects and evolvability drives whole genome duplication to fixation. *PLoS Comput. Biol.* 10:e1003547.

Dalcq J, Pasque V, Ghaye A, Larbuisson A, Motte P, Martial JA, et al. (2012) RUNX3, EGR1 and SOX9B form a regulatory cascade required to modulate BMP-signaling during cranial cartilage development in zebrafish. *PLoS One* 7:e50140.

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

Dassule HR, McMahon AP (1998) Analysis of epithelial-mesenchymal interactions in the initial morphogenesis of the mammalian tooth. *Dev. Biol.* 202:215–227.

Dathe K, Kjaer KW, Brehm A, Meinecke P, Nürnberg P, Neto JC, et al. (2009) Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2. *Am. J. Hum. Genet.* 84:483–492.

Davit-Béal T, Tucker AS, Sire J-Y (2009) Loss of teeth and enamel in tetrapods: fossil record, genetic data and morphological adaptations. *J. Anat.* 214:477–501.

Deckers MML, van Bezooijen RL, van der Horst G, Hoogendam J, van der Bent C, Papapoulos SE, et al. (2002) Bone morphogenetic proteins stimulate angiogenesis through osteoblast-derived vascular endothelial growth factor A. *Endocrinology* 143:1545–1553.

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3:e314.

Demers C, Hamdy RC (1999) Bone morphogenetic proteins. *Sci. Med.* 1:8–17.

Derveaux S, Vandesompele J, Hellemans J (2010) How to do successful gene expression analysis using real-time PCR. *Methods* 50:227–230.

Derynck R, Zhang YE (2003) Smad-dependent and Smad-independent pathways in TGF-β family signalling. *Nature* 425:577–584.

Doolittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284:601–603.

Doss MX, Chen S, Winkler J, Hippler-Altenburg R, Odenthal M, Wickenhauser C, et al. (2007) Transcriptomic and phenotypic analysis of murine embryonic stem cell derived BMP2+ lineage cells: an insight into mesodermal patterning. *Genome Biol.* 8:R184.

Doss MX, Gaspar JA, Winkler J, Hescheler J, Schulz H, Sachinidis A (2012) Specific gene signatures and pathways in mesodermal cells and their derivatives derived from embryonic stem cells. *Stem Cell Rev.* 8:43–54.

Drissi H, Luc Q, Shakoori R, Lopes SCS, Choi JY, Terry A, et al. (2000) Transcriptional autoregulation of the bone related CBFA1/RUNX2 gene. *J. Cell. Physiol.* 184:341–350.

Ducy P, Karsenty G (2000) The family of bone morphogenetic proteins. *Kidney Int.* 57:2207–2214.

Dumont N, Arteaga CL (2003) A kinase-inactive type II TGFβ receptor impairs BMP signaling in human breast cancer cells. *Biochem. Biophys. Res. Commun.* 301:108–112.

Dunworth WP, Cardona-Costa J, Bozkulak EC, Kim J-D, Meadows S, Fischer JC, et al. (2014) Bone morphogenetic protein 2 signaling negatively modulates lymphatic development in vertebrate embryos. *Circ. Res.* 114:56–66.

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.

Escriva H, Manzon L, Youson J, Laudet V (2002) Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. *Mol. Biol. Evol.* 19:1440–1450.

Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, et al. (2006) Comparative protein structure modeling using Modeller. In John Wiley and Sons, editor *Current Protocols in Bioinformatics*. pp. 5.6.1–5.6.30.

Fainsod A, Steinbeisser H, De Robertis EM (1994) On the function of BMP-4 in patterning the marginal zone of the Xenopus embryo. *EMBO J.* 13:5015–5025.

Fan Q, Tang T, Zhang X, Dai K (2009) The role of CCAAT/enhancer binding protein (C/EBP)-alpha in osteogenesis of C3H10T1/2 cells induced by BMP-2. *J. Cell. Mol. Med.* 13:2489–2505.

Feiner N, Begemann G, Renz AJ, Meyer A, Kuraku S (2009) The origin of bmp16, a novel Bmp2/4 relative, retained in teleost fish genomes. *BMC Evol. Biol.* 9:277.

Feng J, Yang G, Yuan G, Gluhak-Heinrich J, Yang W, Wang L, et al. (2011) Abnormalities in the enamel in bmp2-deficient mice. *Cells tissues organs* 194:216–221.

Feng JQ, Chen D, Ghosh-Choudhury N, Esparza J, Mundy GR, Harris SE (1997) Bone morphogenetic protein 2 transcripts in rapidly developing deer antler tissue contain an extended 5′ non-coding region arising from a distal promoter. *Biochim. Biophys. Acta - Gene Struct. Expr.* 1350:47–52.

Feng JQ, Harris MA, Ghosh-Choudhury N, Feng M, Mundy GR, Harris SE (1994) Structure and sequence of mouse bone morphogenetic protein-2 gene (BMP-2): comparison of the structures and promoter regions of BMP-2 and BMP-4 genes. *Biochim. Biophys. Acta* 1218:221–224.

Ferguson EL, Anderson KV. (1992) Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the *Drosophila* embryo. *Cell* 71:451–461.

Ferraresso S, Vitulo N, Mininni AN, Romualdi C, Cardazzo B, Negrisolo E, et al. (2008) Development and validation of a gene expression oligo microarray for the gilthead sea bream (Sparus aurata). *BMC Genomics* 9:580.

Fong Y-C, Li T-M, Wu C-M, Hsu S-F, Kao S-T, Chen R-J, et al. (2008) BMP-2 increases migration of human chondrosarcoma cells via PI3K/Akt pathway. *J. Cell. Physiol.* 217:846–855.

Fonseca VG, Laizé V, Valente MS, Cancela ML (2007) Identification of an osteopontin-like protein in fish associated with mineral formation. *FEBS J.* 274:4428–4439.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531–1545.

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

Friedman MS, Oyserman SM, Hankenson KD (2009) Wnt11 promotes osteoblast maturation and mineralization. *Mol. Basis Cell Dev. Biol.* 284:14117–14125.

Fritsch C, Lanfear R, Ray RP (2010) Rapid evolution of a novel signalling mechanism by concerted duplication and divergence of a BMP ligand and its extracellular modulators. *Dev. Genes Evol.* 220:235–250.

Fritz DT, Jiang S, Xu J, Rogers MB (2006) A polymorphism in a conserved posttranscriptional regulatory motif alters bone morphogenetic protein 2 (BMP2) RNA: protein interactions. *Mol. Endocrinol.* 20:1574–1586.

Fritz DT, Liu D, Xu J, Jiang S, Rogers MB (2004) Conservation of Bmp2 post-transcriptional regulatory mechanisms. *J. Biol. Chem.* 279:48950–48958.

Furlong RF, Younger R, Kasahara M, Reinhardt R, Thorndyke M, Holland PWH (2007) A degenerate ParaHox gene cluster in a degenerate vertebrate. *Mol. Biol. Evol.* 24:2681–2686.

Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Bioinformatics* 12:543–548.

Gao Y, Ganss BW, Wang H, Kitching RE, Seth A (2005) The RING finger protein RNF11 is expressed in bone cells during osteogenesis and is regulated by Ets1. *Exp. Cell Res.* 304:127–135.

Gates MA, Kim L, Egan ES, Cardozo T, Sirotkin HI, Dougan ST, et al. (1999) A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. *Genome Res.* 9:334–347.

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, Simes DC, Ortiz-Delgado JB, Viegas CSB, Pinto JP, Kelsh RN, et al. (2006) Osteocalcin and matrix Gla protein in zebrafish (Danio rerio) and Senegal sole (Solea senegalensis): comparative gene and protein expression during larval development through adulthood. *Gene Expr. Patterns* 6:637–652.

Gazzerro E, Canalis E (2006) Bone morphogenetic proteins and their antagonists. *Rev. Endocr. Metab. Disord.* 7:51–65.

Gazzerro E, Deregowski V, Vaira S, Canalis E (2005) Overexpression of twisted gastrulation inhibits bone morphogenetic protein action and prevents osteoblast cell differentiation in vitro. *Endocrinology* 146:3875–3882.

Gemayel R, Vinces MD, Legendre M, Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu. Rev. Genet.* 44:445–477.

Ghatpande S, Brand T, Zile M, Evans T (2006) Bmp2 and Gata4 function additively to rescue heart tube development in the absence of retinoids. *Dev. Dyn.* 235:2030–2039.

Ghosh-Choudhury N, Abboud SL, Chandrasekar B, Ghosh Choudhury G (2003) BMP-2 regulates cardiomyocyte contractility in a phosphatidylinositol 3 kinase-dependent manner. *FEBS Lett.* 544:181–184.

Ghosh-Choudhury N, Choudhury GG, Harris MA, Wozney J, Mundy GR, Abboud SL, et al. (2001) Autoregulation of mouse BMP-2 gene transcription is directed by the proximal promoter element. *Biochem. Biophys. Res. Commun.* 286:101–108.

Giraldez AJ, Cinalli RM, Glasner ME, Enright AJ, Thomson JM, Baskerville S, et al. (2005) MicroRNAs regulate brain morphogenesis in zebrafish. *Science.* 308:833–838.

Goldman DC, Donley N, Christian JL (2008) Genetic interaction between Bmp2 and Bmp4 reveals shared functions during multiple aspects of mouse organogenesis. *Mech. Dev.* 126:117–127.

Goldstein AM, Brewer KC, Doyle AM, Nagy N, Roberts DJ (2005) BMP signaling is necessary for neural crest cell migration and ganglion formation in the enteric nervous system. *Mech. Dev.* 122:821–833.

Gong Y, Xu F, Zhang L, Qian Y, Chen J, Huang H, et al. (2014) MicroRNA expression signature for Satb2-induced osteogenic differentiation in bone marrow stromal cells. *Mol. Cell. Biochem.* 387:227–239.

Goode DK, Callaway HA, Cerda GA, Lewis KE, Elgar G (2011) Minor change, major difference: divergent functions of highly conserved cis-regulatory elements subsequent to whole genome duplication events. *Development* 138:879–884.

Gordon KJ, Blobe GC (2008) Role of transforming growth factor-beta superfamily signaling pathways in human disease. *Biochim. Biophys. Acta* 1782:197–228.

Goto T, Matsui Y, Fernandes RJ, Hanson DA, Kubo T, Yukata K, et al. (2006) Sp1 family of transcription factors regulates the human α2 (XI) collagen gene (COL11A2) in Saos-2 osteoblastic cells. *J. bone Miner. Res.* 21:661–673.

Govender S, Csimma C, Genant HK, Valentin-Opran A, Amit Y, Arbel R, et al. (2002) Recombinant human bone morphogenetic protein-2 for treatment of open tibial fractures: a prospective, controlled, randomized study of four hundred and fifty patients. *J. Bone Joint Surg. Am.* 84A:2123–2134.

Granero-Moltó F, Weis JA, Miga MI, Landis B, Myers TJ, O'Rear L, et al. (2009) Regenerative effects of transplanted mesenchymal stem cells in fracture healing. *Stem Cells* 27:1887–1898.

Grimsrud CD, Rosier RN, Puzas JE, Reynolds PR, Reynolds SD, Hicks DG, et al. (1998) Bone morphogenetic protein-7 in growth-plate chondrocytes: regulation by retinoic acid is dependent on the stage of chondrocyte maturation. *J. Orthop. Res.* 16:247–255.

Guarnieri DJ, DiLeone RJ (2008) MicroRNAs: a new class of gene regulators. *Ann. Med.* 40:197–208.

Guo B, Wagner A, He S (2011) Duplicated gene evolution following whole-genome duplication in teleost fish. In PF Friedberg, editor *Gene Duplication*. InTech, pp. 27–36.

Guo L, Zhao Y, Zhang S, Liu K, Gao X (2008) Retinoic acid down-regulates bone morphogenetic protein 7 expression in rat with cleft palate. *Chinese Med. Sci. J.* 23:28–31.

Hang Q, Zhou Y, Hou S, Zhang D, Yang X, Chen J, et al. (2014) Asparagine-linked glycosylation of bone morphogenetic protein-2 is required for secretion and osteoblast differentiation. *Glycobiology* 24:292–304.

Haque R, Chong NW, Chaurasia SS, Klein DC, Iuvone PM (2004) Retinal melatonin biosynthesis: Interactions of A/T–rich regions and CRE–like sequences contribute to cAMP–dependent regulation of the chicken AANAT promoter. *Investig. Ophtalmol. Vis. Sci.* 45:659.

Harris MP, Fallon JF, Prum RO (2002) Shh-Bmp2 signaling module and the evolutionary origin and diversification of feathers. *J. Exp. Zool.* 294:160–176.

Harris SE, Guo D, Harris MA, Krishnaswamy A, Lichtler A (2003) Transcriptional regulation of BMP-2 activated genes in osteoblasts using gene expression microarray analysis: role of Dlx2 and Dlx5 transcription factors. *Front. Biosci.* 8:1249–1265.

Hassan MQ, Javed A, Morasso MI, Karlin J, Montecino M, Wijnen AJ Van, et al. (2004) Dlx3 transcriptional regulation of osteoblast differentiation: temporal recruitment of Msx2, Dlx3, and Dlx5 homeodomain proteins to chromatin of the Osteocalcin gene. *Mol. Cell. Biol.* 24:9248–9261.

Hatakeyama S, Ohara-Nemoto Y, Kyakumoto S, Satoh M (1996) Retinoic acid enhances expression of bone morphogenetic protein-2 in human adenocarcinoma cell line (HSG-S8). *Biochem. Mol. Biol. Int.* 38:1235–1243.

He C, Chen X (2005) Transcription regulation of the vegf gene by the BMP/Smad pathway in the angioblast of zebrafish embryos. *Biochem. Biophys. Res. Commun.* 329:324–330.

He X, Dziak R, Yuan X, Mao K, Genco R, Swihart M, et al. (2013) BMP2 genetically engineered MSCs and EPCs promote vascularized bone regeneration in rat critical-sized calvarial bone defects. *PLoS One* 8:e60473.

Heller LC, Li Y, Abrams KL, Rogers MB (1999) Transcriptional regulation of the Bmp2 gene. *J. Biol. Chem.* 274:1394–1400.

Helvering LM, Sharp RL, Ou X, Geiser AG (2000) Regulation of the promoters for the human bone morphogenetic protein 2 and 4 genes. *Gene* 256:123–138.

Heng S, Paule S, Hardman B, Li Y, Singh H, Rainczuk A, et al. (2010) Posttranslational activation of bone morphogenetic protein 2 is mediated by proprotein convertase 6 during decidualization for pregnancy establishment. *Endocrinology* 151:3909–3917.

Hino J, Takao M, Takeshita N, Konno Y, Nishizawa T, Matsuo H, et al. (1996) cDNA cloning and genomic structure of human bone morphogenetic protein-3B (BMP-3b). *Biochem. Biophys. Res. Commun.* 223:304–310.

Van Hoek MJA, Hogeweg P (2009) Metabolic adaptation after whole genome duplication. *Mol. Biol. Evol.* 26:2441–2453.

Hoeppner LH, Secreto F, Jensen ED, Li X, Kahler RA, Westendorf JJ (2009) Runx2 and bone morphogenic protein 2 regulate the expression of an alternative Lef1 transcript during osteoblast maturation. *J. Cell. Physiol.* 221:480–489.

Hoffman LM, Garcha K, Karamboulas K, Cowan MF, Drysdale LM, Horton WA, et al. (2006) BMP action in skeletogenesis involves attenuation of retinoid signaling. *J. Cell Biol.* 174:101–113.

Hoffmann MF, Jones CB, Sietsema DL (2013) Recombinant human Bone Morphogenetic Protein-2 (rhBMP-2) in posterolateral lumbar spine fusion: complications in the elderly. *J. Orthop. Surg. Res.* 8:1.

Hogan BL (1996) Bone morphogenetic proteins: multifunctional regulators of vertebrate development. *Genes Dev.* 10:1580–1594.

Hollnagel A (1999) Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. *J. Biol. Chem.* 274:19838–19845.

Horbelt D, Denkis A, Knaus P (2012) A portrait of transforming growth factor β superfamily signalling: background matters. *Int. J. Biochem. Cell Biol.* 44:469–474.

Horton JM, Summers AP (2009) The material properties of acellular bone in a teleost fish. *J. Exp. Biol.* 212:1413–1420.

Horvath LG, Henshall SM, Kench JG, Turner JJ, Golovsky D, Brenner PC, et al. (2004) Loss of BMP2, Smad8, and Smad4 expression in prostate cancer progression. *Prostate* 59:234–242.

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496:498–503.

Hsu Y-L, Huang M-S, Yang C-J, Hung J-Y, Wu L-Y, Kuo P-L (2011) Lung tumor-associated osteoblast-derived bone morphogenetic protein-2 increased epithelial-to-mesenchymal transition of cancer by Runx2/Snail signaling pathway. *J. Biol. Chem.* 286:37335–37346.

Hung W-T, Wu F-J, Wang C-J, Luo C-W (2012) DAN (NBL1) specifically antagonizes BMP2 and BMP4 and modulates the actions of GDF9, BMP2, and BMP4 in the rat ovary. *Biol. Reprod.* 86:158.

Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat. Rev. Genet.* 12:99–110.

Hupkes M, Sotoca AM, Hendriks JM, van Zoelen EJ, Dechering KJ (2014) MicroRNA miR-378 promotes BMP2-induced osteogenic differentiation of mesenchymal progenitor cells. *BMC Mol. Biol.* 15:1.

Huttner IG, Trivedi G, Jacoby A, Mann SA, Vandenberg JI, Fatkin D (2013) A transgenic zebrafish model of a human cardiac sodium channel mutation exhibits bradycardia, conduction-system abnormalities and early death. *J. Mol. Cell. Cardiol.* 61:123–132.

Hwang S-PL, Chen CA, Peng M-Y, Chen C-P (2003) Evolutionary conservation of the bone morphogenetic protein 2/4 gene between diploblastic and triploblastic metazoans. *Zool. Stud.* 42:227–234.

Ideno H, Takanabe R, Shimada A, Imaizumi K, Araki R, Abe M, et al. (2009) Protein related to DAN and cerberus (PRDC) inhibits osteoblastic differentiation and its suppression promotes osteogenesis in vitro. *Exp. Cell Res.* 315:474–484.

Imamura T, Oshima Y, Hikita A (2013) Regulation of TGF-β family signalling by ubiquitination and deubiquitination. *J. Biochem.* 154:481–489.

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

Israel DI, Nove J, Kerns KM, Moutsatsos IK, Kaufman RJ (1992) Expression and characterization of bone morphogenetic protein-2 in chinese hamster ovary cells. *Growth Factors* 7:139–150.

Jagadeeswaran P, Liu YC (1997) A hemophilia model in zebrafish: analysis of hemostasis. *Blood Cells. Mol. Dis.* 23:52–57.

Jaillon O, Aury J-M, Brunet F, Petit J-L, Stange-Thomann N, Mauceli E, et al. (2004) Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* 431:946–957.

Jia Z, Wang S, Tang J, He D, Cui L, Liu Z, et al. (2014) Does crystal deposition in genetic hypercalciuric rat kidney tissue share similarities with bone formation? *Basic Transl. Sci.* 83:509–514.

Jiang S, Chandler RL, Fritz DT, Mortlock DP, Rogers MB (2010) Repressive BMP2 gene regulatory elements near the BMP2 promoter. *Biochem. Biophys. Res. Commun.* 392:124–128.

Jiang S, Zhang S, Langenfeld J, Lo S-C, Rogers MB (2008) Mycoplasma infection transforms normal lung cells and induces bone morphogenetic protein 2 expression by post-transcriptional mechanisms. *J. Cell. Biochem.* 104:580–594.

Jin H, Pi J, Huang X, Huang F, Shao W, Li S, et al. (2012) BMP2 promotes migration and invasion of breast cancer cells via cytoskeletal reorganization and adhesion decrease: an AFM investigation. *Appl. Microbiol. Biotechnol.* 93:1715–1723.

Jones CM, Lyons KM, Hogan BL (1991) Involvement of Bone Morphogenetic Protein-4 (BMP-4) and Vgr-1 in morphogenesis and neurogenesis in the mouse. *Development* 111:531–542.

Jung RE, Ha CHF, Sailer HF, Weber FE (2003) Effect of rhBMP-2 on guided bone regeneration in humans: a randomized , controlled clinical and histomorphometric study. *Clin. Oral Implants Res.* 14:556–568.

Kabashi E, Brustein E, Champagne N, Drapeau P (2011) Zebrafish models for the functional genomics of neurogenetic disorders. *Biochim. Biophys. Acta* 1812:335–345.

Kang MH, Kim JS, Seo JE, Oh SC, Yoo YA (2010) BMP2 accelerates the motility and invasiveness of gastric cancer cells via activation of the phosphatidylinositol 3-kinase (PI3K)/Akt pathway. *Exp. Cell Res.* 316:24–37.

Kapinas K, Kessler CB, Delany AM (2009) miR-29 suppression of osteonectin in osteoblasts: regulation during differentiation and by canonical Wnt signaling. *J. Cell. Biochem.* 108:216–224.

Karperien M, Farih-Sips H, Löwik CW, de Laat SW, Boonstra J, Defize LH (1997) Expression of the parathyroid hormone-related peptide gene in retinoic acid-induced differentiation: involvement of ETS and Sp1. *Mol. Endocrinol.* 11:1435–1448.

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.* 382:521–526.

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

Katagiri T, Tsukamoto S (2013) The unique activity of bone morphogenetic proteins in bone: a critical role of the Smad signaling pathway. *Biol. Chem.* 394:703–714.

Katti MV., Ranjekar PK, Gupta VS (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Mol. Biol. Evol.* 18:1161–1167.

Kawabata M, Imamura T, Miyazono K (1998) Signal transduction by bone morphogenetic proteins. *Cytokine Growth Factor Rev.* 9:49–61.

Kawai S, Sugiura T (2001) Characterization of human bone morphogenetic protein (BMP)-4 and -7 gene promoters: activation of BMP promoters by Gli, a sonic hedgehog mediator. *Bone* 29:54–61.

Kawakami Y, Rodriguez-León J, Izpisúa Belmonte JC (2006) The role of TGFβs and Sox9 during limb chondrogenesis. *Curr. Opin. Cell Biol.* 18:723–729.

Keane TM, Creevey CJ, Pentony MM, Naughton TJ, Mclnerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* 6:29.

Keller S, Nickel J, Zhang J-L, Sebald W, Mueller TD (2004) Molecular recognition of BMP-2 and BMP receptor IA. *Nat. Struct. Mol. Biol.* 11:481–488.

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39:1278–1284.

Kim HKW, Oxendine I, Kamiya N (2013) High-concentration of BMP2 reduces cell proliferation and increases apoptosis via DKK1 and SOST in human primary periosteal cells. *Bone* 54:141–150.

Kimura N, Matsuo R, Shibuya H, Nakashima K, Taga T (2000) BMP2-induced apoptosis is mediated by activation of the TAK1-p38 kinase pathway that is negatively regulated by Smad6. *J. Biol. Chem.* 275:17647–17652.

Kingsley DM (1994) The TGF-beta superfamily: new members, new receptors, and new genetic tests of function in different organisms. *Genes Dev.* 8:133–146.

Kirsch T, Nickel J, Sebald W (2000) BMP-2 antagonists emerge from alterations in the low-affinity binding epitope for receptor BMPR-II. *EMBO J.* 19:3314–3324.

Kirsch T, Sebald W, Dreyer MK (2000) Crystal structure of the BMP-2–BRIA ectodomain complex. *Nat. Struct. Biol.* 7:492–496.

Kleeff J, Maruyama H, Ishiwata T, Sawhney H, Friess H, Büchler MW, et al. (1999) Bone morphogenetic protein 2 exerts diverse effects on cell growth in vitro and is expressed in human pancreatic cancer in vivo. *Gastroenterology* 116:1202–1216.

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

Koike N, Kassai Y, Kouta Y, Miwa H, Konishi M, Itoh N (2007) Brorin, a novel secreted bone morphogenetic protein antagonist, promotes neurogenesis in mouse neural precursor cells. *J. Biol. Chem.* 282:15843–15850.

Korchynskyi O, ten Dijke P (2002) Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. *J. Biol. Chem.* 277:4883–4891.

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.* 208:3493–3502.

Kronenberg HM (2003) Developmental regulation of the growth plate. *Nature* 423:332–336.

Kruithof BPT, Fritz DT, Liu Y, Garsetti DE, Frank DB, Pregizer SK, et al. (2011) An autonomous BMP2 regulatory element in mesenchymal cells. *J. Cell. Biochem.* 112:666–674.

Kube D, Laser H, von Knethen A, Tesch H (1999) The AT-rich region between -54 to -66 is important for the promoter activity of interleukin-10 in Epstein-Barr virus positive Burkitt's lymphoma cells. *Genes Immun.* 1:105–114.

Kuhn DE, Martin MM, Feldman DS, Terry AV., Nuovo GJ, Elton TS (2008) Experimental validation of miRNA targets. *Methods* 44:47–54.

Kültz D (2012) The combinatorial nature of osmosensing in fishes. *Physiology* 27:259–275.

Kundu M, Javed A, Jeon J-P, Horner A, Shum L, Eckhaus M, et al. (2002) Cbfβ interacts with Runx2 and has a critical role in bone development. *Nat. Genet.* 32:639–644.

Kuraku S (2010) Palaeophylogenomics of the vertebrate ancestor-impact of hidden paralogy on hagfish and lamprey gene phylogeny. *Integr. Comp. Biol.* 50:124–129.

Kuraku S, Meyer A, Kuratani S (2009) Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol. Biol. Evol.* 26:47–59.

De la Pompa JL, Epstein JA (2012) Coordinating tissue interactions: Notch signaling in cardiac development and disease. *Dev. Cell* 22:244–254.

LaBonne C, Bronner-Fraser M (1999) Molecular mechanisms of neural crest formation. *Annu. Rev. Cell Dev. Biol.* 15:81–112.

Laizé V, Gavaia P, Cancela M (2014) Fish: a suitable system to model human bone disorders and discover drugs with osteogenic or osteotoxic activities. *Drug Discov. Today Dis. Model.*:10.1016/j.ddmod.2014.08.001.

Laizé V, Tiago DM, Aureliano M, Cancela ML (2009) New insights into mineralogenic effects of vanadate. *Cell. Mol. Life Sci.* 66:3831–3836.

Lalani SR, Thakuria JV., Cox GF, Wang X, Bi W, Bray MS, et al. (2009) 20p12.3 microdeletion predisposes to Wolff-Parkinson-White syndrome with variable neurocognitive deficits. *J. Med. Genet.* 46:168–175.

Langenfeld EM, Kong Y, Langenfeld J (2006) Bone morphogenetic protein 2 stimulation of tumor growth involves the activation of Smad-1/5. *Oncogene* 25:685–692.

Langenfeld EM, Langenfeld J (2004) Bone morphogenetic protein-2 stimulates angiogenesis in developing tumors. 2:141–149.

Larraín J, Oelgeschläger M, Ketpura NI, Reversade B, Zakin L, De Robertis EM (2001) Proteolytic cleavage of chordin as a switch for the dual activities of twisted gastrulation in BMP signaling. *Development* 128:4439–4447.

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25:1307–1320.

Legendre M, Ritchie W, Lopez F, Gautheret D (2006) Differential repression of alternative transcripts: a screen for miRNA targets. *PLoS Comput. Biol.* 2:e43.

Li AG, Koster MI, Wang X-J (2003) Roles of TGFβ signaling in epidermal/appendage development. *Cytokine Growth Factor Rev.* 14:99–111.

Li CW, Ge W (2011) Spatiotemporal expression of bone morphogenetic protein family ligands and receptors in the zebrafish ovary: a potential paracrine-signaling mechanism for oocyte-follicle cell communication. *Biol. Reprod.* 85:977–986.

Li GHY, Deng H-W, Kung AWC, Huang Q-Y (2011) Identification of genes for bone mineral density variation by computational disease gene identification strategy. *J. Bone Miner. Metab.* 29:709–716.

Li L, Shi J-Y, Zhu G-Q, Shi B (2012) MiR-17-92 cluster regulates cell proliferation and collagen synthesis by targeting TGFβ pathway in mouse palatal mesenchymal cells. *J. Cell. Biochem.* 113:1235–1244.

Li RH, Wozney JM (2001) Delivering on the promise of bone morphogenetic proteins. *Trends Biotechnol.* 19:255–265.

Liao J, Hu N, Zhou N, Lin L, Zhao C, Yi S, et al. (2014) Sox9 potentiates BMP2-induced chondrogenic differentiation and inhibits BMP2-induced osteogenic differentiation. *PLoS One* 9:e89025.

Liao WX, Moore RK, Otsuka F, Shimasaki S (2003) Effect of intracellular interactions on the processing and secretion of bone morphogenetic protein-15 (BMP-15) and growth and differentiation

factor-9. Implication of the aberrant ovarian phenotype of BMP-15 mutant sheep. *J. Biol. Chem.* 278:3713–3719.

Lickert H, Kutsch S, Kanzler B, Tamai Y, Taketo MM, Kemler R (2002) Formation of multiple hearts in mice following deletion of beta-catenin in the embryonic endoderm. *Dev. Cell* 3:171–181.

Lin SJ, Lerch TF, Cook RW, Jardetzky TS, Woodruff TK (2006) The structural basis of TGF-β, bone morphogenetic protein, and activin ligand binding. *Reproduction* 132:179–190.

Liu A, Niswander LA (2005) Bone morphogenetic protein signalling and vertebrate nervous system development. *Nat. Rev. Neurosci.* 6:945–954.

Liu D, Fritz DT, Rogers MB, Shatkin AJ (2008) Species-specific cis-regulatory elements in the 3'-untranslated region direct alternative polyadenylation of bone morphogenetic protein 2 mRNA. *J. Biol. Chem.* 283:28010–28019.

Liu Y, Liu W, Hu C, Xue Z, Wang G, Ding B, et al. (2011) MiR-17 modulates osteogenic differentiation through a coherent feed-forward loop in mesenchymal stem cells isolated from periodontal ligaments of patients with periodontitis. *Stem Cells* 29:1804–1816.

Locklin RM, Riggs BL, Hicok KC, Horton HF, Byrne MC, Khosla S (2001) Assessment of gene regulation by bone morphogenetic protein 2 in human marrow stromal cells using gene array technology. *J. Bone Miner. Res.* 16:2192–2204.

Londin ER, Niemiec J, Sirotkin HI (2005) Chordin, FGF signaling, and mesodermal factors cooperate in zebrafish neural induction. *Dev. Biol.* 279:1–19.

Luu HH, Song W-X, Luo X, Manning D, Luo J, Deng Z-L, et al. (2007) Distinct roles of bone morphogenetic proteins in osteogenic differentiation of mesenchymal stem cells. *J. Orthop. Res.* 25:665–677.

Luzi E, Marini F, Sala SC, Tognarini I, Galli G, Brandi ML (2008) Osteogenic differentiation of human adipose tissue-derived stem cells is modulated by the miR-26a targeting of the SMAD1 transcription factor. *J. Bone Miner. Res.* 23:287–295.

Lyons KM, Pelton RW, Hogan BL (1990) Organogenesis and pattern formation in the mouse: RNA distribution patterns suggest a role for bone morphogenetic protein-2A (BMP-2A). *Development* 109:833–844.

Ma L, Lu M-F, Schwartz RJ, Martin JF (2005) Bmp2 is essential for cardiac cushion epithelial-mesenchymal transition and myocardial patterning. *Development* 132:5601–5611.

Mackie EJ (2003) Osteoblasts: novel roles in orchestration of skeletal architecture. *Int. J. Biochem. Cell Biol.* 35:1301–1305.

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.

Madabhushi M, Lacy E (2011) Anterior visceral endoderm directs ventral morphogenesis and placement of head and heart via BMP2 expression. *Dev. Cell* 21:907–919.

Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R (2013) Gene duplication as a major force in evolution. *J. Genet.* 92:155–161.

Maguer-Satta V, Rimokh R (2004) FLRG, member of the follistatin family, a new player in hematopoiesis. *Mol. Cell. Endocrinol.* 225:109–118.

Majumdar MK, Wang E, Morris EA (2001) BMP-2 and BMP-9 promotes chondrogenic differentiation of human multipotential mesenchymal cells and overcomes the inhibitory effect of IL-1. *J. Cell. Physiol.* 189:275–284.

Marcellini S, Henriquez JP, Bertin A (2012) Control of osteogenesis by the canonical Wnt and BMP pathways in vivo: cooperation and antagonism between the canonical Wnt and BMP pathways as cells differentiate from osteochondroprogenitors to osteoblasts and osteocytes. *Bioessays* 34:953–962.

Marques CL, Fernández I, Rosa J, Viegas MN, Cancela ML, Laizé V (2014) Spatiotemporal expression and retinoic acid regulation of bone morphogenetic proteins 2, 4 and 16 in Senegalese sole. *J. Appl. Ichthyol.* 30:713–720.

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

Martínez-Barbera JP, Toresson H, Da Rocha S, Krauss S (1997) Cloning and expression of three members of the zebrafish Bmp family: Bmp2a, Bmp2b and Bmp4. *Gene* 198:53–59.

Martinez-Sanchez A, Dudek KA, Murphy CL (2012) Regulation of human chondrocyte function through direct inhibition of cartilage master regulator SOX9 by microRNA-145 (miRNA-145). *J. Biol. Chem.* 287:916–924.

Massagué J (1998) TGF-beta signal transduction. *Annu. Rev. Biochem.* 67:753–791.

Massirer KB, Pasquinelli AE (2006) The evolving role of microRNAs in animal gene expression. *Bioessays* 28:449–452.

Matsubara H, Hogan DE, Morgan EF, Mortlock DP, Einhorn TA, Gerstenfeld LC (2012) Vascular tissues are a primary source of BMP2 expression during bone formation induced by distraction osteogenesis. *Bone* 51:168–180.

Mbalaviele G, Sheikh S, Stains JP, Salazar VS, Cheng S, Chen D, et al. (2005) β-Catenin and BMP-2 synergize to promote osteoblast differentiation and new bone formation. *J. Cell. Biochem.* 418:403–418.

McCauley DW, Bronner-Fraser M (2004) Conservation and divergence of BMP2/4 genes in the lamprey: expression and phylogenetic analysis suggest a single ancestral vertebrate gene. *Evol. Dev.* 6:411–422.

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

McMahon JA, Takada S, Zimmerman LB, Fan C-M, Harland RM, McMahon AP (1998) Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev.* 12:1438–1452.

Merino R, Rodriguez-Leon J, Macias D, Gañan Y, Economides AN, Hurle JM (1999) The BMP antagonist gremlin regulates outgrowth, chondrogenesis and programmed cell death in the developing limb. *Development* 126:5515–5522.

Meunier F, 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.

Mi M, Jin H, Wang B, Yukata K, Sheu T-J, Ke QH, et al. (2013) Chondrocyte BMP2 signaling plays an essential role in bone fracture healing. *Gene* 512:211–218.

Michnick SW, Ear PH, Manderson EN, Remy I, Stefan E (2007) Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat. Rev. Drug Discov.* 6:569–582.

Michos O, Panman L, Vintersten K, Beier K, Zeller R, Zuniga A (2004) Gremlin-mediated BMP antagonism induces the epithelial-mesenchymal feedback signaling controlling metanephric kidney and limb organogenesis. *Development* 131:3401–3410.

Mignone F, Gissi C, Liuni S, Pesole G (2002) Untranslated regions of mRNAs. *Genome Biol.* 3:REVIEWS0004.

Mikio T (2000) Structural and functional studies on the leukemia inhibitory factor receptor (LIF-R): gene and soluble form of LIF-R, and cytoplasmatic domain of LIF-R required for differentiation and growth arrest of myeloid leukemic cells. *Leuk. an lymphoma* 37:517–525.

Miyake S, 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* 17:1517–1525.

Miyoshi K, Nagata H, Horiguchi T, Abe K, Arie Wahyudi I, Baba Y, et al. (2008) BMP2-induced gene profiling in dental epithelial cell line. *J. Med. Investig.* 55:216–226.

Modica S, Wolfrum C (2013) Bone morphogenic proteins signaling in adipogenesis and energy homeostasis. *Biochim. Biophys. Acta* 1831:915–923.

Monzen K, Hiro, Shiojima I, Hiroi Y, Oka T, Takimoto E, et al. (1999) Bone morphogenetic proteins induce cardiomyocyte differentiation through the mitogen-activated protein kinase kinase kinase TAK1 and cardiac transcription factors Csx/Nkx-2.5 and GATA-4. *Mol. Cell. Biol.* 19:7096–7105.

Morikawa M, Koinuma D, Miyazono K, Heldin C-H (2013) Genome-wide mechanisms of Smad binding. *Oncogene* 32:1609–1615.

Moser M, Patterson C (2005) Bone morphogenetic proteins and vascular differentiation: BMPing up vasculogenesis. *Thromb. Haemost.* 94:713–718.

Mu Y, Xu Z, Contreras CI, McDaniel JS, Donly KJ, Chen S (2012) Phenotype characterization and sequence analysis of BMP2 and BMP4 variants in two mexican families with oligodontia. *Genet. Mol. Res.* 11:4110–4120.

Murshed M, Harmey D, Millán JL, McKee MD, Karsenty G (2005) Unique coexpression in osteoblasts of broadly expressed genes accounts for the spatial restriction of ECM mineralization to bone. *Genes Dev.* 19:1093–1104.

Murshed M, Schinke T, McKee MD, Karsenty G (2004) Extracellular matrix mineralization is regulated locally; different roles of two gla-containing proteins. *J. Cell Biol.* 165:625–630.

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

Nakase T, Miyaji T, Tomita T, Kaneko M, Kuriyama K, Myoui A, et al. (2003) Localization of bone morphogenetic protein-2 in human osteoarthritic cartilage and osteophyte. *Osteoarthr. Cartil.* 11:278–284.

Nakayama N, Han CE, Cam L, Lee JI, Pretorius J, Fisher S, et al. (2004) A novel chordin-like BMP inhibitor, CHL2, expressed preferentially in chondrocytes of developing cartilage and osteoarthritic joint cartilage. *Development* 131:229–240.

Nelson JS (2006) Fishes of the world. 4th editio. I Jonh Wiley & Sons, editor.

Nicholls JJ, Brassill MJ, Williams GR, Bassett JHD (2012) The skeletal consequences of thyrotoxicosis. *J. Endocrinol.* 213:209–221.

Nishi T, Itoh S (1986) Enhancement of transcriptional activity of the Escherichia coli trp promoter by upstream A + T-rich regions. *Gene* 44:29–36.

Nishimura R, Hata K, Ikeda F, Ichida F, Shimoyama A, Matsubara T, et al. (2008) Signal transduction and transcriptional regulation during mesenchymal cell differentiation. *J. Bone Miner. Metab.* 26:203–212.

Nishimura R, Hata K, Matsubara T, Wakabayashi M, Yoneda T (2012) Regulation of bone and cartilage development by network between BMP signalling and transcription factors. *J. Biochem.* 151:247–254.

Nishinakamura R, Sakaguchi M (2014) BMP signaling and its modifiers in kidney development. *Pediatr. Nephrol.* 29:681–686.

Nohe A, Hassel S, Ehrlich M, Neubauer F, Sebald W, Henis YI, et al. (2002) The mode of bone morphogenetic protein (BMP) receptor oligomerization determines different BMP-2 signaling pathways. *J. Biol. Chem.* 277:5330–5338.

Nohe A, Keating E, Knaus P, Petersen NO (2004) Signal transduction of bone morphogenetic protein receptors. *Cell. Signal.* 16:291–299.

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

Orgel LE, Crick FHC (1980) Selfish DNA: the ultimate parasite. *Nature* 284:604–607.

Ozaki T, Sakiyama S (1994) Tumor-suppressive activity of N03 gene product in v-src-transformed rat 3Y1 fibroblasts. *Cancer Res.* 54:646–648.

Padgett RW, Stjohnston RD, Gelbart WM (1987) A transcript from a Drosophila pattern gene predicts a protein homologous to the transforming growth factor-β family. *Nature* 325:81–84.

Padgett RW, Wozney JM, Gelbart WM (1993) Human BMP sequences can confer normal dorsal-ventral patterning in the Drosophila embryo. *Proc. Natl. Acad. Sci. U. S. A.* 90:2905–2909.

Pal R, Khanna A (2006) Role of smad- and wnt-dependent pathways in embryonic cardiac development. *Stem Cells Dev.* 15:29–39.

Panopoulou G, Poustka AJ (2005) Timing and mechanism of ancient vertebrate genome duplications - the adventure of a hypothesis. *Trends Genet.* 21:559–567.

Panopoulou GD, Clark MD, Holland LZ, Lehrach H, Holland ND (1998) AmphiBMP2/4, an amphioxus bone morphogenetic protein closely related to Drosophila decapentaplegic and vertebrate BMP2 and BMP4: insights into evolution of dorsoventral axis specification. *Dev. Dyn.* 213:130–139.

Pasquier J, Lafont A-G, Rousseau K, Quérat B, Chemineau P, Dufour S (2014) Looking for the bird Kiss: evolutionary scenario in sauropsids. *BMC Evol. Biol.* 14:30.

Perry JM, Harandi OF, Paulson RF (2007) BMP4, SCF, and hypoxia cooperatively regulate the expansion of murine stress erythroid progenitors. *Blood* 109:4494–4502.

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

Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes , differentially regulated target genes and sample integrity : BestKeeper – Excel-based tool using pair-wise correlations. :509–515.

Piccolo S, Agius E, Leyns L, Bhattacharyya S, Grunz H, Bouwmeester T, et al. (1999) The head inducer cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* 397:707–710.

Piccolo S, Sasai Y, Lu B, De Robertis EM (1996) Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* 86:589–598.

Pinto JP, Conceição N, Gavaia PJ, Cancela ML (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.

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.

Pouliot F, Labrie C (2002) Role of Smad1 and Smad4 proteins in the induction of p21 (WAF1, Cip1) during bone morphogenetic protein-induced growth arrest in human breast cancer cells. *J. Endocrinol.* 172:187–198.

Pourquié O, Fan C-M, Coltey M, Hirsinger E, Watanabe Y, Bréant C, et al. (1996) Lateral and axial signals involved in avian somite patterning: a role for BMP4. *Cell* 84:461–471.

Price PA, Williamson MK (1985) Primary structure of bovine matrix Gla protein, a new vitamin K-dependent bone protein. *J. Biol. Chem.* 260:14971–14975.

Pucéat M (2007) TGFbeta in the differentiation of embryonic stem cells. *Cardiovasc. Res.* 74:256–261.

Pugh BF, Tjian R (1991) Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev.* 5:1935–1945.

Rabier C-E, Ta T, Ané C (2014) Detecting and locating whole genome duplications on a phylogeny: a probabilistic approach. *Mol. Biol. Evol.* 31:750–762.

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* 39:1373–1381.

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.

Raida M, Clement JH, Leek RD, Ameri K, Bicknell R, Niederwieser D, et al. (2005) Bone morphogenetic protein 2 (BMP-2) and induction of tumor angiogenesis. *J. Cancer Res. Clin. Oncol.* 131:741–750.

Raida M, Heymann AC, Günther C, Niederwieser D (2006) Role of bone morphogenetic protein 2 in the crosstalk between endothelial progenitor cells and mesenchymal stem cells. *Int. J. Mol. Med.* 18:735–739.

Rakian A, Yang W-C, Gluhak-Heinrich J, Cui Y, Harris MA, Villarreal D, et al. (2013) Bone morphogenetic protein-2 gene controls tooth root development in coordination with formation of the periodontium. *Int. J. Oral Sci.* 5:75–84.

Ramel M-C, Hill CS (2012) Spatial regulation of BMP activity. *FEBS Lett.* 586:1929–1941.

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

Rao VVNG, Lfffier C, Wozney JM (1992) The gene for bone morphogenetic protein 2A (BMP2A) is localized to human chromosome 20p12 by radioactive and nonradioactive in situ hybridization. *Hum. Genet.* 90:299–302.

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

Reddi A (1997) Bone morphogenetic proteins: an unconventional approach to isolation of first mammalian morphogens. *Cytokine Growth Factor Rev.* 8:11–20.

Remy I, Michnick SW (2006) A highly sensitive protein-protein interaction assay based on Gaussia luciferase. *Nat. Methods* 3:977–979.

Rentzsch F, Zhang J, Kramer C, Sebald W, Hammerschmidt M (2006) Crossveinless 2 is an essential positive feedback regulator of Bmp signaling during zebrafish gastrulation. *Development* 133:801–811.

Rivera-Feliciano J, Tabin CJ (2006) Bmp2 instructs cardiac progenitors to form the heart-valve-inducing field. *Dev. Biol.* 295:580–588.

Rosa J, Tiago DM, Dias J, Cancela ML, Laizé V (2010) Serum-specific stimulation of proliferation and mineralization of fish bone-derived cells. *J. Appl. Ichthyol.* 26:251–256.

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.

Rosen V (2009) BMP2 signaling in bone development and repair. *Cytokine Growth Factor Rev.* 20:475–480.

Rothhammer T, Poser I, Soncin F, Bataille F, Moser M, Bosserhoff A-K (2005) Bone morphogenic proteins are overexpressed in malignant melanoma and promote cell invasion and migration. *Cancer Res.* 65:448–456.

Rucci N (2008) Molecular biology of bone remodelling. *Clin. Cases Miner. Bone Metab.* 5:49–56.

Ruppert R, Hoffmann E, Sebald W (1996) Human bone morphogenetic protein 2 contains a heparin-binding site which modifies its biological activity. *Eur. J. Biochem.* 237:295–302.

Sadikovic B, Thorner P, Chilton-Macneill S, Martin JW, Cervigne NK, Squire J, et al. (2010) Expression analysis of genes associated with human osteosarcoma tumors shows correlation of RUNX2 overexpression with poor response to chemotherapy. *BMC Cancer* 10:202.

Sadler KC, Amsterdam A, Soroka C, Boyer J, Hopkins N (2005) A genetic screen in zebrafish identifies the mutants vps18, nf2 and foie gras as models of liver disease. *Development* 132:3561–3572.

Sailer MHM, Hazel TG, Panchision DM, Hoeppner DJ, Schwab ME, McKay RDG (2005) BMP2 and FGF2 cooperate to induce neural-crest-like fates from fetal and adult CNS stem cells. *J. Cell Sci.* 118:5849–5860.

Sakuta H, Takahashi H, Shintani T, Etani K, Aoshima A, Noda M (2006) Role of bone morphogenic protein 2 in retinal patterning and retinotectal projection. *J. Neurosci.* 26:10868–10878.

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.

Sampath TK, Rashka KE, Doctor JS, Tucker RF, Hoffmann FM (1993) Drosophila transforming growth factor beta superfamily proteins induce endochondral bone formation in mammals. *Proc. Natl. Acad. Sci. USA* 90:6004–6008.

Santini F, Harmon LJ, Carnevale G, Alfaro ME (2009) Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol. Biol.* 9:194.

Sato T, Mikawa S, Sato K (2010) BMP2 expression in the adult rat brain. *J. Comp. Neurol.* 518:4513–4530.

Sato Y, Nishida M (2010) Teleost fish with specific genome duplication as unique models of vertebrate evolution. *Environ. Biol. fishes* 88:169–188.

Scheufler C, Sebald W, Hulsmeyer M (1999) Crystal structure of human bone morphogenetic protein-2 at 2.7 Å resulution. *J. Mol. Biol.* 287:103–115.

Schier AF, Giraldez AJ (2006) MicroRNA function and mechanism: insights from zebrafish. *Cold Spring Harb. Symp. Quant. Biol.* 71:195–203.

Schlange T, Arnold H-H, Brand T (2002) BMP2 is a positive regulator of Nodal signaling during left-right axis formation in the chicken embryo. *Development* 129:3421–3429.

Schmierer B, Hill CS (2007) TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nat. Rev. Mol. Cell Biol.* 8:970–982.

Schultheiss TM, Burch JB, Lassar AB (1997) A role for bone morphogenetic proteins in the induction of cardiac myogenesis. *Genes Dev.* 11:451–462.

Schurgers LJ, Cranenburg ECM, Vermeer C (2008) Matrix Gla-protein: the calcification inhibitor in need of vitamin K. *Thromb. Haemost.* 100:593–603.

Schurgers LJ, Uitto J, Reutelingsperger CP (2013) Vitamin K-dependent carboxylation of matrix Gla-protein: a crucial switch to control ectopic mineralization. *Trends Mol. Med.* 19:217–226.

Sela-donenfeld D, Kalcheim C (1999) Regulation of the onset of neural crest migration by coordinated activity of BMP4 and Noggin in the dorsal neural tube. *Development* 126:4749–4762.

Selever J, Liu W, Lu M-F, Behringer RR, Martin JF (2004) Bmp4 in limb bud mesoderm regulates digit pattern by controlling AER development. *Dev. Biol.* 276:268–279.

Sengle G, Charbonneau NL, Ono RN, Sasaki T, Alvarez J, Keene DR, et al. (2008) Targeting of bone morphogenetic protein growth factor complexes to fibrillin. *J. Biol. Chem.* 283:13874–13888.

Shah RK, Moncayo VM, Smitson RD, Pierre-Jerome C, Terk MR (2010) Recombinant human bone morphogenetic protein 2-induced heterotopic ossification of the retroperitoneum, psoas muscle, pelvis and abdominal wall following lumbar spinal fusion. *Skeletal Radiol.* 39:501–504.

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

Shapiro JA, von Sternberg R (2005) Why repetitive DNA is essential to genome function. *Biol. Rev.* 80:227–250.

Sharapova NE, Kotnova AP, Galushkina ZM, Lavrova N V., Poletaeva NN, Tukhvatulin AE, et al. (2010) Production of the recombinant human bone morphogenetic protein-2 in Escherichia coli and testing of its biological activity in vitro and in vivo. *Mol. Biol.* 44:923–930.

Sharman AC, Holland PW (1998) Estimation of Hox gene cluster number in lampreys. *Int. J. Dev. Biol.* 42:617–620.

Shen M, Sali A (2006) Statistical potential for assessment and prediction of protein structures. 15:2507–2524.

Sheng N, Xie Z, Wang C, Bai G, Zhang K, Zhu Q, et al. (2010) Retinoic acid regulates bone morphogenic protein signal duration by promoting the degradation of phosphorylated Smad1. *Proc. Natl. Acad. Sci. U. S. A.* 107:18886–18891.

Shimasaki S, Moore KR, Otsuka F, Erickson GF (2004) The bone morphogenetic protein system in mammalian reproduction. *Endocr. Rev.* 25:72–101.

Shivdasani RA (2006) MicroRNAs: regulators of gene expression and cell differentiation. *Blood* 108:3646–3653.

Shore EM, Xu MQ, Shah PB, Janoff HB, Hahn G V, Deardorff MA, et al. (1998) The human Bone Morphogenetic Protein 4 (BMP-4) gene: molecular structure and transcriptional regulation. *Calcif. Tissue Int.* 63:221–229.

Shu B, Zhang M, Xie R, Wang M, Jin H, Hou W, et al. (2011) BMP2, but not BMP4, is crucial for chondrocyte proliferation and maturation during endochondral bone development. *J. Cell Sci.* 124:3428–3440.

Silva AC, Filipe M, Kuerner K-M, Steinbeisser H, Belo JA (2003) Endogenous Cerberus activity is required for anterior head specification in Xenopus. *Development* 130:4943–4953.

Simandi Z, Nagy L (2011) Retinoid signaling is a context-dependent regulator of embryonic stem cells. In MS Kallos, editor *Embryonic stem cells*. InTech, pp. 55–78.

Simes DC, Williamson MK, Ortiz-Delgado JB, Viegas CSB, Price PA, Cancela ML (2003) Purification of matrix Gla protein from a marine teleost fish,. *J. bone Miner. Res.* 18:244–259.

Simon M, Feliers D, Arar M, Bhandari B, Abboud HE (2002) Cloning of the 5'-flanking region of the murine bone morphogenetic protein-7 gene. *Mol. Cell. Biochem.* 233:31–37.

Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. *Annu. Rev. Biochem.* 72:449–479.

Soda H, Raymond E, Sharma S, Lawrence R, Cerna C, Gomez L, et al. (1998) Antiproliferative effects of recombinant human bone morphogenetic protein-2 on human tumor colony-forming units. *Anticancer. Drugs* 9:327–331.

Sohaskey CD, Zuckert WR, Barbour AG (1999) The extended promoters for two outer membrane lipoprotein genes of Borrelia spp. uniquely include a T-rich region. *Mol. Microbiol.* 33:41–51.

Soung DY, Dong Y, Wang Y, Zuscik MJ, Schwarz EM, O'Keefe RJ, et al. (2007) Runx3/AML2/Cbfa3 regulates early and late chondrocyte differentiation. *J. Bone Miner. Res.* 22:1260–1270.

Spender LC, Whiteman HJ, Karstegl CE, Farrell PJ (2005) Transcriptional cross-regulation of RUNX1 by RUNX3 in human B cells. *Oncogene* 24:1873–1881.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

Styrkarsdottir U, Cazier J, Kong A, Rolfsson O, Larsen H, Bjarnadottir E, et al. (2003) Linkage of osteoporosis to chromosome 20p12 and association to BMP2. *PLoS Biol.* 1:351–360.

Sudo S, Avsian-Kretchmer O, Wang LS, Hsueh AJW (2004) Protein related to DAN and cerberus is a bone morphogenetic protein antagonist that participates in ovarian paracrine regulation. *J. Biol. Chem.* 279:23134–23141.

Sugatani T, Hruska KA (2007) MicroRNA-223 is a key factor in osteoclast differentiation. *J. Cell. Biochem.* 101:996–999.

Sugi Y, Yamamura H, Okagawa H, Markwald RR (2004) Bone morphogenetic protein-2 can mediate myocardial regulation of atrioventricular cushion mesenchymal cell formation in mice. *Dev. Biol.* 269:505–518.

Sugimori K, Matsui K, Motomura H, Tokoro T, Wang J, Higa S (2005) BMP-2 prevents apoptosis of the N1511 chondrocytic cell line through PI3K/Akt-mediated NF-k B activation. *J. Bone Miner. Metab.*:411–419.

Sugiura T (1999) Cloning and functional characterization of the 5′-flanking region of the human bone morphogenetic protein-2 gene. *Biochem. J.* 338:433–440.

Sun M, Forsman C, Sergi C, Gopalakrishnan R, O'Connor MB, Petryk A (2010) The expression of twisted gastrulation in postnatal mouse brain and functional implications. *Neuroscience* 169:920–931.

Suzuki A, Kaneko E, Maeda J, Ueno N (1997) Mesoderm induction by BMP-4 and -7 heterodimers. *Biochem. Biophys. Res. Commun.* 232:153–156.

Swanhart LM, Cosentino CC, Diep CQ, Davidson AJ, de Caestecker M, Hukriede NA (2011) Zebrafish kidney development: basic science to translational research. *Birth Defects Res. C. Embryo Today* 93:141–156.

Sweatt A, Sane DC, Hutson SM, Wallin R (2003) Matrix Gla protein (MGP) and bone morphogenetic protein-2 in aortic calcified lesions of aging rats. *J. Thromb. Haemost.* 1:178–185.

Szpalski M, Gunzburg R (2005) Recombinant human bone morphogenetic protein-2: a novel osteoinductive alternative to autogenous bone graft? *Acta Orthop. Belg.* 71:133–148.

Takacs CM, Giraldez AJ (2010) MicroRNAs as genetic sculptors: fishing for clues. *Semin. Cell Dev. Biol.* 21:760–767.

Takigawa Y, Hata K, Muramatsu S, Amano K, Ono K, Wakabayashi M, et al. (2010) The transcription factor Znf219 regulates chondrocyte differentiation by assembling a transcription factory with Sox9. *J. Cell Sci.* 123:3780–3788.

Tamada H, Kitazawa R, Gohji K, Kamidono S, Maeda S, Kitazawa S (1998) Molecular cloning and analysis of the 5′-flanking region of the human bone morphogenetic protein-6 (BMP-6). *Biochim. Biophys. Acta* 1395:247–251.

Taylor JS, Braasch I, Frickey T, Meyer A, Van de Peer Y (2003) Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome Res.* 13:382–390.

Taylor JS, Van de Peer Y, Braasch I, Meyer A (2001) Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos. Trans. R. Soc. B Biol. Sci.* 356:1661–1679.

Thompson DL, Gerlach-Bank LM, Barald KF, Koenig RJ (2003) Retinoic acid repression of bone morphogenetic protein 4 in inner ear development. *Mol. Cell. Biol.* 23:2277–2286.

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.* 582:1381–1385.

Tiago DM, Cancela ML, Laizé V (2011) Proliferative and mineralogenic effects of insulin, IGF-1, and vanadate in fish osteoblast-like cells. *J. Bone Miner. Metab.* 29:377–382.

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.

Tian C, Liu J (2013) Repulsive guidance molecules (RGMs) and neogenin in bone morphogenetic protein (BMP) signaling. *Mol. Reprod. Dev.* 80:700–717.

Tian W, Dong X, Liu X, Wang G, Dong Z, Shen W, et al. (2012) High-throughput functional microRNAs profiling by recombinant AAV-based microRNA sensor arrays. *PLoS One* 7:e29551.

Tomilin N V. (2008) Regulation of mammalian gene expression by retroelements and non-coding tandem repeats. *Bioessays* 30:338–348.

Tressler MA, Richards JE, Sofianos D, Comrie FK, Kregor PJ, Obremskey WT (2011) Bone morphogenetic protein-2 compared to autologous iliac crest bone graft in the treatment of long bone nonunion. *Orthopedics* 34:e877–e884.

Tseng Y-H, He T-C (2007) Bone morphogenetic proteins and adipocyte differentiation. *Cell Sci. Rev.* 3:342–360.

Tsuji K, Bandyopadhyay A, Harfe BD, Cox K, Kakar S, Gerstenfeld L, et al. (2006) BMP2 activity, although dispensable for bone formation, is required for the initiation of fracture healing. *Nat. Genet.* 38:1424–1429.

Umulis D, O'Connor MB, Blair SS (2009) The extracellular regulation of bone morphogenetic protein signaling. *Development* 136:3715–3728.

Upton PD, Long L, Trembath RC, Morrell NW (2008) Functional characterization of bone morphogenetic protein binding sites and Smad1/5 activation in human vascular cells. *Mol. Pharmacol.* 73:539–552.

Urist MR (1965) Bone: formation by autoinduction. *Science (80-.).* 150:893–899.

Urist MR, Huo YK, Brownell AG, Hohl WM, Buyske J, Lietze A, et al. (1984) Purification of bovine bone morphogenetic protein by hydroxyapatite chromatography. *Proc. Natl. Acad. Sci. U. S. A.* 81:371–375.

Valentin-Opran A, Wozney J, Csimma C, Lilly L, Riedel G (2002) Clinical evaluation of recombinant human bone morphogenetic protein-2. *Clin. Orthop. Relat. Res.* 395:110–120.

Vallejo LF, Rinas U (2013) Folding and dimerization kinetics of bone morphogenetic protein-2, a member of the transforming growth factor-β family. *FEBS J.* 280:83–92.

Ventura A, Young AG, Winslow MM, Lintault L, Meissner A, Erkeland SJ, et al. (2008) Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell* 132:875–886.

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* 10:500–509.

Villeneuve LAN, Gisbert E, Moriceau J, Cahu CL, Infante JLZ (2006) Intake of high levels of vitamin A and polyunsaturated fatty acids during different developmental periods modifies the expression of morphogenesis genes in European sea bass (Dicentrarchus labrax). *Br. J. Nutr.* 95:677–687.

Virdi AS, Cook LJ, Oreffo RO, Triffitt JT (1998) Modulation of bone morphogenetic protein-2 and bone morphogenetic protein-4 gene expression in osteoblastic cell lines. *Cell. Mol. Biol.* 44:1237–1246.

Vitt UA, Hsu SY, Hsueh AJ (2001) Evolution and classification of cystine knot-containing hormones and related extracellular signaling molecules. *Mol. Endocrinol.* 15:681–694.

Vukicevic S, Sampath KT (2008) Bone morphogenetic proteins: from local to systemic therapeutics. MJ Parnham, editor.

Walker FO (2007) Huntington's Disease. *Semin. Neurol.* 27:143–150.

Wallin R, Cain D, Hutson SM, Sane DC, Loeser R (2000) Modulation of the binding of matrix Gla protein (MGP) to bone morphogenetic protein-2 (BMP-2). *Thromb. Haemost.* 84:1039–1044.

Walsh DW, Godson C, Brazil DP, Martin F (2010) Extracellular BMP-antagonist regulation in development and disease: tied up in knots. *Trends Cell Biol.* 20:244–256.

Wang E, Rosen V, D'Alessandro J, Bauduy M, Cordes P, Harada T, et al. (1990) Recombinant human bone morphogenetic protein induces bone formation. *Proc. Natl. Acad. Sci. U. S. A.* 87:2220–2204.

Wang L, Park P, Zhang H, La Marca F, Claeson A, Than K, et al. (2012) BMP-2 inhibits tumor growth of human renal cell carcinoma and induces bone formation. *Int. J. Cancer* 131:1941–1950.

Wang L, Park P, Zhang H, La Marca F, Claeson A, Valdivia J, et al. (2011) BMP-2 inhibits the tumorigenicity of cancer stem cells in human osteosarcoma OS99-1 cell line. *Cancer Biol. Ther.* 11:457–463.

Wang S, Furmanek T, Kryvi H, Krossøy C, Totland GK, Grotmol S, et al. (2014) Transcriptome sequencing of Atlantic salmon (Salmo salar L.) notochord prior to development of the vertebrae provides clues to regulation of positional fate, chordoblast lineage and mineralisation. *BMC Genomics* 15:141–158.

Wang S, Wang Q, Crute BE, Melnikova IN, Keller SR, Speck NA (1993) Cloning and characterization of subunits of the T-cell receptor and murine leukemia virus enhancer core-binding factor. *Mol. Cell. Biol.* 13:3324–3339.

Wang X, Chen Z, Zhang R, Liu S, Mei Z, Yu Y, et al. (2012) Construction of a eukaryotic expression vector pEGFP-C1-BMP-2 and its effect on cell migration. *J. Zhejiang Univ. Sci. B.* 13:356–363.

Wang Y, Wu B, Chamberlain AA, Lui W, Koirala P, Susztak K, et al. (2013) Endocardial to myocardial notch-wnt-bmp axis regulates early heart valve development. *PLoS One* 8:e60244.

Wang Y-X, Qian L-X, Liu D, Yao L-L, Jiang Q, Yu Z, et al. (2007) Bone morphogenetic protein-2 acts upstream of myocyte-specific enhancer factor 2a to control embryonic cardiac contractility. *Cardiovasc. Res.* 74:290–303.

Warren AJ, Bravo J, Williams RL, Rabbitts TH (2000) Structural basis for the heterodimeric interaction between the acute leukaemia-associated transcription factors AML1 and CBFbeta. *EMBO J.* 19:3004–3015.

Waxman JS, Yelon D (2011) Zebrafish retinoic acid receptors function as context-dependent transcriptional activators. *Dev. Biol.* 352:128–140.

Weaver M, Yingling JM, Dunn NR, Bellusci S, Hogan BL (1999) Bmp signaling regulates proximal-distal differentiation of endoderm in mouse lung development. *Development* 126:4005–4015.

Wen X-Z, Miyake S, Akiyama Y, Yuasa Y (2004) BMP-2 modulates the proliferation and differentiation of normal and cancerous gastric cells. *Biochem. Biophys. Res. Commun.* 316:100–106.

Wilson PA, Hemmati-Brivanlou A (1995) Induction of epidermis and inhibition of neural fate by Bmp-4. *Nature* 376:331–333.

Winnier G, Blessing M, Labosky PA, Hogan BL (1995) Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev.* 9:2105–2116.

Wise SB, Stock DW (2006) Conservation and divergence of Bmp2a, Bmp2b, and Bmp4 expression patterns within and between dentitions of teleost fishes. *Evol. Dev.* 8:511–523.

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 (2010) The unobtrusive majority: mononucleated bone resorbing cells in teleost fish and mammals. *J. Appl. Ichthyol.* 26:225–229.

Wolfe KH (2001) Yesterday's polyploids and the mystery of diploidization. *Nat. Rev. Genet.* 2:333–341.

Woo EJ (2013) Adverse events after recombinant human BMP2 in nonspinal orthopaedic procedures. *Clin. Orthop. Relat. Res.* 471:1707–1711.

Woods IG, Kelly PD, Chu F, Ngo-hazelett P, Yan Y, Huang H, et al. (2000) A comparative map of the zebrafish genome. *Genome Res.* 1:1903–1914.

Wordinger RJ, Clark AF (2007) Bone morphogenetic proteins and their receptors in the eye. *Exp. Biol. Med.* 232:979–992.

Wozney JM (2002) Overview of bone morphogenetic proteins. *Spine (Phila. Pa. 1976).* 27:2–8.

Wozney JM, Rosen V, Celeste AJ, Mitsock LM, Matthew J, Kriz RW, et al. (1988) Novel regulators of bone formation: molecular clones and activities. *Science.* 242:1528–1534.

Wu J, Kubota J, Hirayama J, Nagai Y, Nishina S, Yokoi T, et al. (2010) p38 Mitogen-activated protein kinase controls a switch between cardiomyocyte and neuronal commitment of murine embryonic stem cells by activating myocyte enhancer factor 2C-dependent bone morphogenetic protein 2 transcription. *Stem Cells Dev.* 19:1723–1734.

Wu L-A, Feng J, Wang L, Mu Y, Baker A, Donly KJ, et al. (2010) Immortalized mouse floxed Bmp2 dental papilla mesenchymal cell lines preserve odontoblastic phenotype and respond to BMP2. *J. Cell. Physiol.* 225:132–139.

Wu M, Chen W, Mi J, Chen D, Wang W, Gao H (2014) Expression analysis of BMP2, BMP5, BMP10 in human colon tissues from Hirschsprung disease patients. *Int J Clin Exp Pathol* 7:529–536.

Wu MY, Hill CS (2009) TGF-β superfamily signaling in embryonic development and homeostasis. *Dev. Cell* 16:329–343.

Xue W, Wallin R, Olmsted-Davis EA, Borrás T (2006) Matrix GLA protein function in human trabecular meshwork cells: inhibition of BMP2-induced calcification process. *Invest. Ophthalmol. Vis. Sci.* 47:997–1007.

Yamaguchi A, Toshihisa K, Suda T (2000) Regulation of osteoblast differentiation mediated by bone morphogenetic proteins, hedgehogs, and Cbfa1. *Endocr. Rev.* 21:393–411.

Yamamoto T, Saatcioglu F, Matsuda T (2002) Cross-talk between bone morphogenic proteins and estrogen receptor signaling. *Endocrinology* 143:2635–2642.

Yan J, Marr TG (2005) Computational analysis of 3'-ends of ESTs shows four classes of alternative polyadenylation in human, mouse, and rat. *Genome Res.* 15:369–375.

Yanagita M (2005) BMP antagonists: their roles in development and involvement in pathophysiology. *Cytokine Growth Factor Rev.* 16:309–317.

Yang H (2013) Conserved or lost: molecular evolution of the key gene GULO in vertebrate vitamin C biosynthesis. *Biochem. Genet.* 51:413–425.

Yang W, Guo D, Harris MA, Cui Y, Gluhak-Heinrich J, Wu J, et al. (2013) Bmp2 in osteoblasts of periosteum and trabecular bone links bone formation to vascularization and mesenchymal stem cells. *J. Cell Sci.* 126:4085–4098.

Yang W, Harris MA, Cui Y, Mishina Y, Harris SE, Gluhak-Heinrich J (2012) Bmp2 is required for odontoblast differentiation and pulp vasculogenesis. *J. Dent. Res.* 91:58–64.

Yao Y, Shahbazian A, Boström KI (2008) Proline and gamma-carboxylated glutamate residues in matrix Gla protein are critical for binding of bone morphogenetic protein-4. *Circ. Res.* 102:1065–1074.

Ying Y, Zhao G-Q (2001) Cooperation of endoderm-derived BMP2 and extraembryonic ectoderm-derived BMP4 in primordial germ cell generation in the mouse. *Dev. Biol.* 232:484–492.

Yu YY, Lieu S, Lu C, Miclau T, Marcucio RS, Colnot C (2010) Immunolocalization of BMPs, BMP antagonists, receptors, and effectors during fracture repair. *Bone* 46:841–851.

Zebboudj AF, Imura M, Boström K (2002) Matrix GLA protein, a regulatory protein for bone morphogenetic protein-2. *J. Biol. Chem.* 277:4388–4394.

Zebboudj AF, Shin V, Boström K (2003) Matrix GLA protein and BMP-2 regulate osteoinduction in calcifying vascular cells. *J. Cell. Biochem.* 90:756–765.

Zhang H, Bradley A (1996) Mice deficient for BMP2 are nonviable and have defects in amnion/chorion and cardiac development. *Development* 122:2977–2986.

Zhang J, Fu W, He M, Xie W, Lv Q, Wan G, et al. (2011) MiRNA-20a promotes osteogenic differentiation of human mesenchymal stem cells by co-regulating BMP signaling. *RNA Biol.* 8:829–838.

Zhang J-L, Patterson LJ, Qiu L-Y, Graziussi D, Sebald W, Hammerschmidt M (2010) Binding between Crossveinless-2 and Chordin von Willebrand factor type C domains promotes BMP signaling by blocking Chordin activity. *PLoS One* 5:e12846.

Zhang R, Edwards JR, Ko S-Y, Dong S, Liu H, Oyajobi BO, et al. (2011) Transcriptional regulation of BMP2 expression by the PTH-CREB signaling pathway in osteoblasts. *PLoS One* 6:e20780.

Zhang YQ, Kanzaki M, Shibata H, Kojima I (1997) Regulation of the expression of follistatin in rat hepatocytes. *Biochim. Biophys. Acta* 1354:204–210.

Zhao M, Xiao G, Berry JE, Franceschi RT, Reddi A, Somerman MJ (2002) Bone morphogenetic protein 2 induces dental follicle cells to differentiate toward a cementoblast/osteoblast phenotype. *J. bone Miner. Res.* 17:1441–1451.

Zheng M, Zhu J, Lu T, Liu L, Sun H, Liu Z, et al. (2013) p300-mediated histone acetylation is essential for the regulation of GATA4 and MEF2C by BMP2 in H9c2 cells. *Cardiovasc. Toxicol.* 13:316–322.

Zhou M, Ma J, Chen S, Chen X, Yu X (2014) MicroRNA-17-92 cluster regulates osteoblast proliferation and differentiation. *Endocrine* 45:302–310.

Zou H, Niswander L (1996) Requirement for BMP signaling in interdigital apoptosis and scale formation. *Science (80-.).* 272:738–741.

APPENDENCES

APPENDIX 1

Supplementary figure 4.2.1. Complete coding sequences (collected or reconstructed) for bone morphogenetic protein (BMP) and decapentaplegic protein (DPP) used to construct the phylogenetic tree presented in Fig. 2. Dme, *Drosophila melanogaster*; Cin, *Ciona intestinalis*; Bfl, *Branchiostoma floridae*; Pma, *Petromyzon marinus*; Cmi, *Callorhinchus milii*; Ler, *Leucoraja erinacea*; Loc, *Lepisosteus oculatus*; Dre, *Danio rerio*; Ipu, *Ictalurus punctatus*; Ame, *Astyanax mexicanus*; Gac, *Gasterosteus aculeatus*; Oni, *Oreochromis niloticus*; Ola, *Oryzias latipes*; Ssa, *Salmo salar*; Tni, *Tetraodon nigroviridis*; Xma, *Xiphophorus maculatus*; Lch, *Latimeria chalumnae*; Amex, *Ambystoma mexicanum*; Eco, *Eleutherodactylus coqui*; Eco, *Xenopus tropicalis*; Aca, *Anolis carolinensis*; Gga, *Gallus gallus*; Cpb, *Chrysemys picta bellii*; Hsa, *Homo sapiens*; Mmu, Mus musculus.

>DmeDPP (Fruit fly; Accession number FBtr0077773)
ATGCGACGTTCGCGAACTGCCAGAATTGGTTGTTGTTGCAGTTGCAGCCAAAGGATTTGTGGATTTTGGAGCGTAACTGAGCG
GCGTGCAGCAGGTGCACCAGAGTTGCAAGCGACCATGCGCGCATGGCTTCTACTCCTCGCAGTGCTGGCGACTTTTCAAACGA
TTGTTCGAGTTGCTAGCACCGAGGATATATCCCAGAGATTCATCGCCGCCATAGCGCCCGTTGCCGCTCATATTCCGCTGGCA
TCAGCATCAGGATCAGGATCAGGACGATCTGGATCTAGATCGGTAGGAGCCTCGACCAGCACAGCATTAGCAAAAGCATTTAA
TCCATTCAGCGAGCCCGCCTCGTTCAGTGATAGTGATAAAAGCCATCGGAGTAAAACAAACAAAAAACCTAGCAAAAGTGACG
CGAACCGACAGTTCAACGAAGTGCATAAGCCAAGAACAGACCAATTAGAAAATTCCAAAAATAAGTCTAAACAATTAGTTAAT
AAACCCAACCACAACAAAATGGCTGTCAAGGAGCAGAGGAGCCACCACAAGAAGAGCCACCACCATCGCAGCCACCAGCCAAA
GCAGGCCAGTGCATCCACAGAATCTCATCAATCCTCGTCGATTGAATCAATCTTCGTGGAGGAGCCGACGCTGGTGCTCGACC
GCGAGGTGGCCTCCATCAACGTGCCCGCCAACGCCAAGGCCATCATCGCCGAGCAGGGCCCGTCCACCTACAGCAAGGAGGCG
CTCATCAAGGACAAGCTGAAGCCAGACCCCTCCACTCTAGTCGAGATCGAGAAGAGCCTGCTCTCGCTGTTCAACATGAAGCG
GCCGCCCAAGATCGACCGCTCCAAGATCATCATCCCCGAGCCGATGAAGAAGCTCTACGCCGAGATCATGGGCCACGAGCTCG
ACTCGGTCAACATCCCCAAGCCGGGTCTGCTGACCAAGTCGGCCAACACAGTGCGAAGTTTTACACACAAAGATAGTAAAATC
GACGATCGATTTCCGCACCACCACCGGTTTCGGCTGCACTTCGACGTGAAGAGCATTCCCGCCGACGAGAAGCTGAAGGCGGC
GGAGCTGCAGCTGACCCGGGACGCACTCAGTCAACAGGTGGTGGCCAGCAGATCGTCGGCGAATCGGACGCGCTACCAGGTGC
TTGTCTACGACATCACGCGCGTCGGGGTGCGTGGTCAGCGGGAGCCGAGCTATCTGCTGTTGGACACCAAGACGGTCCGGCTT
AACAGCACGGACACGGTGAGCCTCGATGTCCAGCCGGCCGTGGACCGGTGGCTGGCGAGTCCGCAGCGCAACTACGGACTGCT
GGTGGAGGTGCGGACGGTCCGCTCCCTGAAGCCGGCCCCACACCACCATGTACGCCTGCGCCGCAGCGCGGACGAGGCGCACG
AGCGGTGGCAGCACAAGCAGCCGCTCCTGTTCACCTACACGGACGACGGGCGGCACAAGGCGCGCTCCATTCGGGACGTGTCT
GGCGGAGAGGGCGGTGGCAAGGGCGGCCGGAACAAGCGGCAGCCGAGACGGCCTACGAGGCGCAAGAACCACGACGACACCTG
CCGGCGGCACTCGCTGTACGTGGACTTCTCGGACGTGGGCTGGGACGACTGGATTGTGGCGCCTCTGGGCTACGATGCATATT
ACTGCCACGGGAAGTGCCCCTTCCCGCTGGCCGACCACTTTAACTCGACCAATCACGCCGTGGTGCAGACCCTGGTCAACAAT
ATGAATCCCGGCAAGGTGCCGAAGGCGTGCTGCGTGCCCACGCAACTGGACAGCGTGGCCATGCTCTATCTCAACGACCAAAG
TACGGTGGTGCTGAAGAACTACCAGGAGATGACCGTGGTGGGCTGTGGCTGTCGATAG

>CinBMP2/4 (Sea squirt; Accession number ENSCINT00000034715)
ATGGTGGCGCTTACGGATTGGACTTTACTCGCGACTCTGGTGTTGCATTTTCTATTGGACGTGAGCTCGGGCCTCGTGCCCAA
GGTCGGCCGGAATAACTTGATACGGCAGGCCGCCAAATACACGAGCGACACCCGAGAGGAAGAGGACGCGATAGTCAACGAAT
ATGAGAAAAGCCTGTTGAACATGTTCGGGCTTAAGCGAAGGCCCACGCCGGAAAAAGACTTCCAGGTCCCGCGTATTATGCAG
CACCTGTACAAGGCGCACATGGGCGACTACTACGAGCCCGAACCTAAGATCGTGACGTCATGGGAAACCGGTTTTGATCTTCC
GCATGAAGATATTGCGAGCCGGGTTAACACTGCTAGAAGTTATCACCACGAAGATAACGAGGAGCATTACCCCGGGATGCCGA
GCAACCACATCCGGCTACGATTTGACATCAGTACGATGCCGAGTGAAGAGGTTATCGGCGCTGCAGAGCTTCTCCTGCATCGA
GAGGCAGTGTGGCATCATGTAACATCAGGCCAAGGCCACATGCATCGCTTGAATATATACGAGATCGTCAAAATGCCTCCGGA
CGATGGGGGAAGAGCAGGGAGCCGTTTTGGAAGCCGACCCCAACCTATTACTCGGTTACTGGACACGAGACTCATAGACACCC
GGAACACAACTTGGGAACGGTTTGATGTTAGTCCAGCTACCTTGCTATGGTCCAGGGTCGCGCATAAGGAGAACCATGGTTTA
GTGGTCGAGGTGGTGCAAGAAGATGAACAAATACCGCCGCAAGGTGACGCTGAGCGACATGTGCGCCTGAAACGTGATGTGCA
CCACGATGTTTCGGACCGAGAGTGGGTCCATCTAAGACCAATGCTTCTCACTTATTCGCACGACGGAAAAGATACCTCGCTGC
GTCGGCGCAAACGTAGTAGTAGAAACCGCGACTCCTCTAACTCAAGTTCTGGAAGAAAAAACCGTCGGAAGAAACGAAAGCGT
CGAAATGTTTGTCAGCGGCATGACCTCTACGTGGACTTGCTGAGGTCGACTGGACGACTTGGATTGTTGCTCCACCGGGCTAC
GACGCCTACTATTGCCAAGGACAGTGCCCGTTCCCAATGTCGGATCATTTGA

>BflBMP2/4 (Lancelet; Accession number AF068750.1)
ATGATTCGGGCGTTCGAGTCAAGTCTGCTCAACATGTTCGGATTAAACGAGCGACCACGGCCCCGGAAGAACCTGGTGATTCC
TCCCTACATGTTGGAGTTGTACCTCAGCCAGACGAAAGACCCTGAAAACCCCAGTGTCAACTTCAACTTTGCGGGAAAGTCAA
CAAGTACTGCCAACACCGTGCGTAGTTTCCACCATGAAGAAGAATCGGAGGCAGGGCAGCCGGTGTGGGAGGGCGACGACGAG
ATCGACCGGCGATTATGGTTCAACACAAGTGCCGTCCCCAGCGTGGAACTTATAAAGGCGGCCGAGTTACGCTTGTTTCGGGA
ACAGATCGATGTTGATCACGTTCAGTACGGAGACTCCACGGACCACCACTTGTACCGAGTAAACGTGTACGAGGTAATGAGAC
CGAACAGTAGGACTAACACGGATACCATAACGCGGCTTCTGGATACCAAATTGGTGGATGTTCGTAATTCGTCGTGGGAATCC
TTTGATGTCCGGTCGGCTGTTACTAAGTGGAAGAATTCCCCGGAGCGAAACTACGGGTTGGAGGTCGAAGTGGTGTCTCCAAA
GCGCGGGGCCCTATCAAACCACCACGTACGGTTACGACGGTCTACAGACATGGACGACCACTCGTGGCAGCACCGCAGACCGC
TGCTGCTCACGTACACGGATGACGGCAAAGGTTCCAGTAACAGTAATAGGGTCGCCTCCAGGCAAAAACGCGCCAATGGAAGA

AAAAAACAGCGGCGGAGACTCAAAGCCAACTGCAGGAGACATTCGTTGTATGTTGACTTCAGTGATGTTGGCTGGAATGACTG
GATTGTAGCTCCACCTGGCTACCAGGCCTACTACTGCCACGGGGAATGCCCCTTTCCCCTCGCAGATCACCTCAATTCCACAA
ACCATGCAATAGTTCAAACTCTAGTCAACTCCGTAAATCCTCTCGCAGTGCCCAAGGCCTGCTGTGTGCCTACAGATCTCAGT
CCTATTAGTATGCTATATTTGAACGAAAATGATCAAGTTGTTCTCAAAAACTACCAGGACATGGTCGTGGAAGGATGCGGGTG
CCGTTAG

>PmaBMP2/4a (Sea lamprey; Accession number AY602220)
ATGCTTCTGCTCGTAGCCCTTTGCTTCCCGGCCGAGAGCGAAAGCGGCAGGAGGCCGCCACCGCCTCCGGGCGCCACGAGCGA
TGACGGCGACGCGCTGGCCCGAGGGCGCGCCACGCCGCCGCACGGCCTCACGCGGCACGTCGAGGGCCAGCTGCTCGGCATGT
TCGGGCTGCAGAGGCGGCCTCGCCCGGGGCCCGGCGCCGTGGTGCCACCGTACATGCTGCAGCTGTACCGGGCCCTGCACGGG
GCCCACAGTGGGGCCCGTGACGTGGGACGGCCCCTGGACCGTTTGGTGGCGCGGCCCGCGAGCCAGGCCAACACCGTGCGCAG
CTTCCATCACGACGAGTCCGCGGAGCACGTGCCGGCCGACTCGGGCGACAGCACGGCACGGCGGCTGCTGTTCAACGTCAGCA
GCATCCCCGACGGCGAGGTGATCACCTCGGCGGAGCTGCACGTCTACCGCGAGCGCCTGAGCAGCCCCGCGAGCGCCGGCCTC
CACCGCATCAACGTGTACGAGGTGCTGCGGCCCGCGGCCGCCGACGGGACGCCCATCGCCCGCCTCCTGGACACGCGGGTGGT
GCACTCCGGCCGCAGCGAGTGGGARCGCTTTGACGTGAGCCCCGCCGCGGTGCGCTGGGCCGCCACGAAAGAGCCCAACCACG
GCCTGCTGGTGGAGGTGCAGCACCTTGACGGAGGCACGCCCGAGAAGAGGCGGCACGTGCGCATCGGCAGGTCGCTGCACGCC
GAGGCGGTGGCGGCGGCCGCCCGGGACGGCGCCAGCGAAGGGGGCGACGGAGGGGAGGGCTGGCCGCAGCTCAGGCCTCTGCT
GGTGACGTTCGGCCACGACGGGAAAACGAGGGACGAGGGGACGCTGCTGAGGCCGCGCCCCAAGCGCAACTCGCGGCCCAACA
AGGGAGGCCGTAGGGGTCGAGGGCAGTGCGCCCGCTACCCTCTGTACGTGGACTTCAGCGACGTGGGCTGGAACGACTGGATC
GTCGCGCCGCCAGGATACAACGCCTTTTTCTGCCAGGGCGAGTGCCACTTCCCGCTGCCGCAGCACCTCAACTCCACGAACCA
CGCGATCGTGCAGACGCTCGTCAACTCGGTGAATCCCGAGGTGCCGCGCGCGTGCTGCATCCCCACGGAGCTGACGCCGATCG
CGCTGCTGTACCTGGACGAGTACGAGAAGGTGGTGCTCAAGAATTACCAGGACATGGTGGTGGAGGGCTGCGGGTGCCGCTAG

>PmaBMP2/4b (Sea lamprey; Accession number AY602221)
ATGATCGGAAGCGTAGGAGGACTCGGAACGCGACGAGGAGGAGGAGGAGGATCTGAGAGACGCGCAGACCTCCTCCTGGCAGA
AAAGTCTCCGGCGACAACGCTCGGCCGCGATGACGAGGAGAAAGCGCGGAGTGGGAACTCGCGGCGGCTGCTGGAGGAGCTGG
AGGGGAGGCTGCTCGGGGTTTTCGGGCTGCAGAGGCGACCACGACCCGGCACGGGGATCGCAGTGCCCGCCTACCTCGTGCAG
CTCTACCGTGCCCAGCAGCACACGGGGCGCGGAGGCGGCCCCCGACGGCGATGGGCCCGGGCAGACAAGGAGCGTGTGGGCAA
AGACGATGTGGATGTCGAGGAGGAGAAGGAGGAGGTCGCATTCCCACGGGAAGCGCAAGGAAGAGCAAACACCGTCCGAGGGT
TTCATCACGACGAGTCTACGGAGAAGCTCTCACTTGGGCAGAGCACGGAAGACGGAACAACGTGGCACTTCCTCTTCAATCTC
TCGTCCATCCCGGACTCTGAGGAGGTGACTGCGGCTGAACTGCGCGTGCACCACACACGCGTGCACAGCCCATGTCCCTCATC
CTCGCCGGCATGCGAGTTGGCCCCTCGGCTGGAGCGCATCAACGTGTATGAAGTTGTGGCGCCGCCATCGTCTCCATCGGGCG
CTGCCTCCCGGCTGCTGGACACTCGCGTGGTGCGCACAAACGAGTCGCGCTGGGAAGCGTTCGACGTGAGCCCCGCCGTGTCG
CGCTGGACCCGGGGCTCCGCACCTAACCGGGGCTTCGCAGTCGAGGTGCTGCCCGTCCGGCGGCCAAGTGGCGGGGTGGCCGC
CAATGGCGTGCCCAGCGAGGCGGTGCTCGCGCAGCCCCGCTCGGGCGTCGCCTCGCTCTTCCCGGGCGATGGGAGCCACCAGA
CCGAGCCGAGACCGCTCCTCGTCACTTTTGGCAGCGACGGCCGGGCCCCATTCACGCCGCGGAGCCGTGCAAGGAGGAGCATC
GGCGGGGCCCCAAGGCAGGCGGCGCACAAGGCGAGGCGCAAGCCGCGCTACAGCTGCAGGCGCCACGCGCTCTACGTCGACTT
CCGCGAGGTGGGCTGGAATGACTGGATTGTGGCGCCGCCCGGGTACCACGCCTACTTCTGCCACGGCGAGTGCCCCTTCCCTC
TGGCCGACCACCTCAACTCGACGAACCACGCCATCGTGCAGACGCTCGTCAACTCGGTGAACGCGAGCATCCCTCGCGCCTGC
TGCGTGCCCACGGAGCTCAGCCCCATCTCCATGCTCTACCTGGACGAGTACGGCAAGGTGGTGCTCAAGAACTACCAGGACAT
GGTGGTGGAGGGCTGCGGCTGCCGCTAA

>PmaBMP2/4c (Sea lamprey; Accession number AY602222)
ATGAGCGAGTCTCGGCACGGCTTTGTGCTGCTGCTGCTGCTGATTATTATGATGCCGTCAACACTGCCACCGTGCGCGGCAAT
GGATTTATCAACGTTCGGACGAGTCGATATCGTCAAGGAGGAACGCGGCGCACAAGTCCTCTCGGCGTTTTCGTCCGCCACCG
CCGGTGCCCACAAAACATATTCCTCGTCGTCATCATCCGCCGCCGCTGTTGCTGCTTCGTCCGCTTCGTCATCGGAGAGCCAA
CGACTCCTGCAAAAGGACCTCGAGACGCGGCTGCTGGAAGTGTTCGGCCTGAAGAGGCGGCCGAGACCCAGGCCAGGCGTTGC
GGTATCCCAGTACCTCGTGGACGTCTACCGCCTTCTGCAGACCGCCGGGCTGGGGCCAAGCGGGCCCGGGGGGTCCCGAGCGG
CGTCCGCGTCCGAGTTGGAATCCGATTTGGGGTTCGCGGAGCGAGCGTCGGCGATGGCCAACACAGTGAGGGGCTTCCACCAC
GAAGATTCGTTGGAGGAGCAGCCTTACCACGGCCAATCGGCGGCGGCAGCGGCGTCGTCGTCCCACCACCACCACCAACA
CCACCGCCTGGTCTTCGACCTGTCCAGCCTCCCCGCCGACGAGTGGGTGCACGCTGCCGAGCTGCGCCTCCACCGGGCCCGGC
TGCGGCCCCACGGCGCCGTGGGCCGCCCGCTCCGCGTCAACGTCTACGAGGCGGCGGCGACGGCGGCGACGCTCCTCGACTCG
CGGCTCGTGCACGACAACGCCACGCGCTGGGAGAGCTTCGACGTGAGCGCCGCCCTCCAGCGACGGGCGCTCAGCGGGCGGCC
CAGCTCCAGCCTCGGCCTCCTCGTCGAGTTGCTGCCGACCCCCGGGGAGGACGCGGCCGACGGCGACGACGACGGCGGCGGCG
ACGACGGCGGCGGCCACGTTGGCCACGTGCGCGCCAGCCGGGCGGCCGTACCGCCGCGGGCCCGCTCGCCTGAGGACCTCGAC
TCGTGGCCGCGGCTCAGGCCGCTGCTCGTCACCTACGGCCACGACGGGAGGGGCGGCCGAACGTTGGGCGAGGCGAGGGCGAG
GTCCCGGCGGGCGTCGGGGAGGCACGCGGGCGGCCGGCGTCCGCGCGTGCTCTGCCAGCGCCAGCCGCTCTACGTGGACTTCC
GCGAGGTGGGCTGGGACGACTGGATCGTCGCGCCGCCGGGCTACAACGCCTACTTCTGCCAGGGCGAGTGCCCCTTCCCGCTC
GCCGACCACCTCAACTCGACGAACCACGCCATCGTGCAGACGCTCGTCAACTCGGTGAACGCGAGCATCCCGCGCGCGTGCTG
CGTGCCCACCGAGCTGAGCCCCATCTCCATGCTCTACATGGACGAGTATGAGAAGGTGGTGCTCAAGAACTACCAGGACATGG
TGGTAGAGGGCTGCGGCTGCCGCTGA

>CmiBMP2 (Elephant shark; Accession number AAVX02028834)
ATGAAGCGAGGCCGGCCTGGAGGCACAGGGCTGTCATTGCTCTGGGTGCTGCTGTTGCTGTTGCTGTGCCAGGTGTTGCTGGG
CGGCTCTGCGGGGTTGCTCCCCGGGATGGCGAGACGCAGGCTCCCCGAGCAGCGGGCACACAGCCCCGAGCAGCGGGCACACA
GCCCCGAGCTAGTCCAGGAGCTCCAGCGACGACTGCTCATTGCCCTGGGCTTCAGTCGCCGGCCCGAGCCCAGCCCTGGTGTG
GTGGTGCCGCTCTACATGTTGCAACTGTACAGGGAGCGCTCCGGCCAGCCACAACCTCAACCACAACCCCAACAGCCCCAACC
ACAACCCCAACCACAACCCCAACAGCCCCAACCACAACCTCAATACCACCCATCTCAACATCCTCAACACCAACACCAACAGG

CTGAGCCTCGCAGGAGGTCCAGCCTCAGCCTCTCCAACACCATCAGGAGTTTCCATCACGAAGAAACCCTGGAGGAGCTGCCG
GGTGGCGGGGGAAAGAGCGTGCGTCGATTTTTCTTCAACCTGGACTCTATCCCGGCAGAGGAGCTGGTCAGCTCGGCCGAGCT
GAGGATCTTCCGTGAGGAGGTGCGAGAGGCCAGCGCCAACGCCAGCGTTCACCACAGGGTCAACGTGTACGAGATCGTGAAGC
CCCTGCCAGGAGCTGAGATGGGCGGGGAACCCATCACCCGCTTGCTGGACACCAGGCTGGTGCACCAGGCCCAGAGCAAGTGG
GAGAGCTTCGATGTGACCCCGGCAGCCTCGCGCTGGGCTTCCCGCGGGCAGCACAACCACGGCCTGGTGGTGGAAGTCCTGCA
CTCTGACAGCGAGCACTCCAGGCGGCACGTGAGGATCAGCAGGTCCTTGCATCAGGACGACGGGACCTGGGCTCAGATGAGGC
CCCTGTTAGTGACATACAGCCACGACGGGAAAGGGCAGAGCTTGGAGAAGAGGTTCAGGCGCCAGGCCAAACACAAGCAGAGG
AAGAGACTGAGGACCAGCTGCAAGAGGCACACTCTGTATGTGGACTTCAGTGACGTGGGCTGGAATGATTGGATCGTGGCACC
TCCTGGCTACCACGCCTTCTACTGCCAGGGTGAGTGCCCCTTTCCCTTGGCCGATTATCTCAACTCAACCAACCACGCCATTG
TGCAGACTTTGGTCAACTCAGTGAACGCTAACATTCCCAGAGCCTGTTGCGTGCCCACGGACCTCAGTCCCATCTCCATGCTG
TATCTGGATGAGTATGACAAAGTGGTGCTCAAGAACTATCAGGATATGGTGGTGGAGGGCTGTGGGTGTCATTAA

>CmiBMP4 (Elephant shark; Accession numbers AAVX01576064, AAVX01001906, AAVX01310926, 1572727321, 1573959575, 1574162849, 1573921934, 1572917552, 15729969422, 1573506432 and 1576724418)
ATGATTCCTGGTAACCGAATGCTGATGGTAATTTTATTATGCCAAGTCTTGCTGGGAGGTAATGCTAGTCTGATACCCGAGGA
AGGAAGGAAGAAGTTCGCTGAGCAAAACCAGGCGGGTGGACGTCGCTTTCCGCAAAACCATGAACTGCTACGGGAATTCGAGG
CCACTTTACTTAACATGTTCGGATTGCAACGACGCCCACAGCCCAGCAAATCTGCCGTCATTCCCCAGTACATGCTGGATCTG
TACAGGCTTCAGTCTGGGGACGAGATCACCCACGACATCAACTTCGATTATCCCGAGATATCGGCGAGCCGGGCTAACACCGT
GAGGGGCTTCCATCACGAAGACCACTTGGAGTTGGTGCCAGGCTGGAGGCAAGACTTCACGTTTCGTTTTGTGTTTAATATCA
GCGCGGTTCCCGAGAACGAAGTGCTTTCTTCTGCCGAGCTGAGAGTCTACCGGGAGCGGATCGACAGCGCCTTGAACTGGGAT
GAGGGCTTCCACCGCATCAACGTCCATGAGATCATGAGAGCCCCCAGCAAGAAGGGTGACCTCATCACCAGACTTTTGGATAC
CAGGTTGGTCCACCAGAACAGGACGCAGTGGGAGAGCTTTGACGTGAGCCCGGCCGTGCTGAGATGGACCATGCACAAACAAC
CCAACCACGGGCTGGCGGTCGAAGTGATTCACCTGAACGAGACCAGAAGCCAGCAAGGCTCGCACGTTAGGATTAGCAGATCG
CTGCACCAGGGTGGAGTGGGGGGAGGAGGAGGAGGAGGAGGAGAAGAAGGAGGAGGAGGAGGAGGAGATTGGTCGCAGCTCAG
GCCTTTGCTTGTAACTTTCAGCCACGATGGCAAGGGACACGCTCTCACCCGGAGGGTGAAACGAAGCAACAAAAGCCAGAGAC
GCAAGAAAAACAAATCACACTGCAGGAGGCACTCGCTCTATGTGGATTTCAGCGATGTGGGCTGGAATGACTGGATAGTGGCA
CCGCCTGGTTACCAAGCCTTTTACTGTCACGGTGACTGCCCATTCCCACTGGCAGATCACCTGAATTCAACCAACCACGCCAT
CGTGCAGACTCTTGTAAATTCTGTAAATTCCAACATCCCCAAGGCTTGCTGCGTCCCCACAGAACTCAGCCCTATCTCAATGC
TCTACTTGGACGAATATGACAAAGTTGTGCTGAAAAACTATCAGGAGATGGTAGTAGAGGGATGCGGATGCCACTGA

>LerBMP2 (Little skate; AESE011560473 and AESE010115821)
ATGACAGCGGGGAGAAGACCACTCATGATGCTCTTCCTGTGTCAAGTTCTGTTCGGCGGCTCGGCGGGGCTCATCCCCGAGGT
CGGCCGCAAGAAGTTCGCCGAGCAACGGGGCAGCTCGAAGGGGAGCAGCCCGCTGCAGACCGAGGGCATGCTGCAGGAGTTTG
AGATGCGCCTACTCAACATGTTCGGCCTCCGGCGGCGCCCTCAGCCCGGCAAGAACCCTGTCATCCCGCAGTATATGGTGGAT
CTGTACAAGTTGCACGTCGGCGAAGAGGCTCATCGGCACGGAGGCGCTTCTGGCCTCTCCCATTTCCCCGACAAACCGGCCAG
CCGAGCCAACACAGTCAGGAGCTTCCACCACGAAGAACCATTGGAAGAGTTGCCCGGAGCCAGGGGAGAGACGACACGCCGCT
TCCTCTTTAATTTGAACTCGATCCCGAAGGAGGAGCTGGTCACCTCTGCGGAGCTGCGGATCTACCGGGAGGAGGTGCGGGAG
GCCTGCGCCAACGCCAGCGCCAGCGGCGGAGGCCACCACCGCATCAACGTCTACGAGATCCTCAAGCCGACGGCCCCCTCCGG
CCACGACCCCATCACCCGGCTGCTGGACACTAAGCTGGTGCGGCACAACACCAGCCGGTGGGAGAGCTTCGACGTCAGCCCGT
CGGTGATGAGGTGGACGGCGCAGGGCCAGCCCAACCACGGCTTCGTGGCCGAGGTCGTCCACTTGGACCAGCAGGGGTGCCGG
CCCTCCAAGAGGCACGTCAGGGTGAGCCGCTCCTTGCACCGGGACGCGGAGAGCTGGCCCCGGCTGCGGCCTTTGCTGGTAAC
GTTCAGCCACGATGGCAAGGGGCACACCCTGGAGAAACGGGTGCGACGTCAGGCCAAGCACCACAAGCAGAAGAAGAGGCTCA
AGTCCAGCTGCAGGCGACATCCTTTATATGTGGATTTCAACGACGTGGGGTGGAATGACTGGATAGTGGCACCCCCGGGATAT
CATGCTTTTTACTGTCAGGGAGAGTGCCCTTTCCCCCTGGCAGACCATTTAAACTCTACCAACCATGCCATCGTGCAGACTCT
GGTCAACTCAGTGAACGCAAACATTCCCCGGGCTTGCTGTGTCCCCACTGACCTGAGCCCCATCTCCATGCTCTACCTCGACG
AGCAGGACAAAGTGGTGCTGAAGAACTACCAGGACATGGTGGTGGAGGGCTGTGG

>LerBMP4 (Little skate; Accession numbers AESE010060999, EE992056, GD242603, GE298551, CV067625 and GD242034)
ATGATTCCTGGTAACCGAATGCTGATGGTAATTTTATTATGCCAAGCCTTGCTGGGAGGTACTGCTAGTCTGATACCAGAGGA
AGGGAGGAAAAAGTTCGCAGATCAGAACCAGGCGGGTGAACGTCGCTCTCCGCAAAACCATGAACTGCTCCGGGAATTCGAGG
CCACTTTACTCAATATGTTCGGACTCCACCGGCGCCCACAGCCCAGTAAATCTGCCGTCGTTCCCCACTACATGCTGGATCTC
TACAGACTCCAGTCTGGGGACGACGTCCCTCACGACAGCAGCTTCGAATATCCCGAGAGATCCACCAGCCGCGCCAACACCGT
GAGGAGCTTCCACCATGAAGAACACTTGGAGCTGATGCCGGGGCAGAGGCAAGATACCCAGCTTCGCTTTGTGTTCAACATCA
GCGCGGTGCCCGGCAGCGAGGTGCTCTCGTCCGCGGAGCTGAGGCTGTACCGGGAGCAGATCGACACGGTGTTGAACTGGAAG
GAAGGGTTCCATCGCATCAACGTGTATGAAATCATGAAAGCCTCCGGCCGGCAGGGCGACCTCATCACCAGGCTGCTCGACAC
CCGGCTCGTCCACCAGAACAAGAGCCGGTGGGAGAGTTTCGACGTGAGCCCGGCCGTGCTGAGATGGACCCTCCAAAGGCAAC
CCAACCACGGGCTGGCCGTCGAGGTGATTCACCTGAACGAGACGAGGAGCCGCCAGGGTATGCACGTCAGGATGAGCAGGTCG
CTTCGCCAGACGTACGGAGACTGGTCCCAGTTCAGACCCTTGTTAGTGACTTTCAGCCACGACGGAAAGGGGCACGCACTGAT
CCGCAGGGCCAAACGCAGCGGCAAAGGTCCGAAGCGCAAGAAGAACAGGTCTCACTGCAGGAGACACTCTCTCTATGTGGATT
TCAGCGACGTGGGCTGGAATGACTGGATTGTGGCACCCCCGGGTTACCAGGCTTTTTACTGTCACGGAGACTGCCCCTTCCCA
CTGGCAGACCACCTGAACTCGACCAACCACGCGATCGTGCAAACGCTTGTAAATTCTGTAAATTCCAACATCCCCAAAGCTTG
CTGTGTCCCCACGGAACTCAGCCCCATATCAATGCTCTACTTGGACGAATACGATAAAGTTGTGCTGAAAAACTATCAGGAGA
TGGTCGTGGAGGGATGTGGATGCCGCTGA

>LocBMP2 (Spotted gar; Accession number AHAT01009032)
ATGGTGCTGCTGCTCTGTCAGGTGTTGCTCGGAGGTTCGGTCGGACTTATTCCGGAGGTTGGCCGAAGGAAGTTCAGCGAATC

GTCTCGGCAAAGCACGCAGCAATCGGAGGACGTTCTTAACGAATTCGAGCTTCGCCTCCTTAATATGTTTGGACTGAAGCGCA
GACCAACTCCAAGCAAGCTTGCGGTGGTCCCGCAGTACATGGTGGACCTCTATCACATGCATTCAGGGAATGGAGATCAGGAC
CCCAACCGCTCCAGATCAGTGCTGGGCCAGTATTCGGAAAGGTCTGCCAGCAGAGCGAATACAATTAGGAGTTTTCACCATGA
AGAATCCTTGGAGGCACTTTTGAGCATCGGTGGAAAAACAACACAGAAATTCTTCTTCAACCTGACATCCATTCCTGGGGAGG
AAATAGTCTCTGCTGCAGAACTGAGGATTTTTCGAGATCAAGTCCTAGGTGCCATTACAGACAGCAGTGGTGGTTACCATCGC
ATTAACATTTATGAAGTTCTCAAACCAGCCTTTTCCTCAAAGGAACCTATAATAAGACTTTTGGATACTAGATTAGTGCAACA
CAGCCAGAGCAAATGGGAGAGCTTTGATGTGAGCCCTGCTATAATGCGATGGACCATGGAAGGGCTTACCAACCATGGTTTTA
TGGTGGAGATTGTCCATTTGGACAATGAAGGGGGGGAAGCAAAGAGGCATGTGAGGATCAGCAGGTCTTTGCATGAGGATGAG
GAAACATGGCCTCAGATGAGGCCTTTATTAGTGACATTTAGCCATGATGGCAAGGGACACGTTCTCCATAAAAGGGAGAAACG
GCAAACCAAGCCAAGACAGAAGAAGAGGCCTAAAGCCAGCTGCAAAAGACATGAGCTATACGTGGACTTCAGTGATGTTGGGT
GGAATGACTGGATAGTTGCCCCACCAGGGTATCATGCCTTTTACTGCCAAGGAGAATGCCCTTTTCCTTTAGCAGACCACCTA
AACTCCACAAATCATGCTATTGTGCAAACTTTAGTCAACTCAGTTAACACAAACATTCCCAAGGCCTGCTGTGTGCCTACAGA
ACTCAGCGCAATTTCTCTGCTCTACCTTGATGAATATGAAAAAGTTATATTAAAGAACTACCAGGACATGGTTGTGGAGGGCT
GTGGTTGCCGTTGA

>LocBMP4 (Spotted gar; Accession number AHAT01019072)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTATTATGCCAAGTCGTACTAGGAGAGAGCAACCATGCTAGTCTAATACT
GGAGGAAGGAAAGAAAAAAGTGTCCGAGCTGCAGGGCCGTCGCTCTGGGCAAAGCCATGAATTGTTGCGGGACTTTGAGACCA
CGCTGCTGCAGATGTTTGGGCTGCACCAACGTCCCAGGCCCAGCCGCTCCGCAGTGGTGCCACAGTACATGCTGGACCTCTAT
CGCCTGCAGTCCGGGGAGGCAGAGGAGGCCGGCGCTGTTGATGCCAGCTTTGAGTACCCCGAGAGGTCAGCCAGTCGAGCCAA
CACAGTGAGGAGCTTCCATCACGAAGAGCACCTGGAGCGCGTGGAAGGCGAGCAGGAGAACTCGGACACTCCTCTCCGCTTCT
TCTTCAACCTCAGCAACATCCCCGAGGACGAGGTGCTGTCCTCAGCGGAGCTCCGGCTGTACCGGCAGCAGGTGGAGAACGCG
GACCAGCACGGGGAGGCCGAGGGGCTCCACCGGATTAACGTCTACGAGGTCCTCAAGCCCCCCCGGGCCACGCAGCTCATCAC
GCGGCTCCTGGACACCAGACTCGTGCGGCACAACGCCTCGCGCTGGGAGAGCTTCGACGCCAGCCCGGCCGTGCTGCGCTGGA
CACGGGACAAACTGCCCAACCACGGCCTGGCCGTGGAAGTCCTCCACCTCAACCACACCCCCACGCAGCAGGGCCGCCATGTG
CGCGTTAGCCGCTCCCTCCACCGGGACCCCGGCGAGGACTGGGCCCAGCTGCGCCCCCTGCTGGTCACCTTCGGCCACGACGG
CAAGGGCCACCCCCTGACCCGGCGGGCGAAGCGCAGCGCCAAGCCGCGTGGCCGCCGGAAGAACCGGAACTGCAGGCGGCACA
CGCTGTACGTGGACTTCAGCGACGTGGGCTGGAATGACTGGATAGTGGCACCGCCTGGATACCAGGCGTTTTACTGCCAGGGG
GATTGCCCGTTTCCTCTGTCAGACCACCTGAACTCCACCAATCACGCCATCGTGCAGACGCTGGTGAACTCTGTCAACACCAA
GATCCCCAAGGCCTGCTGCGTGCCCACAGAGCTCAGTGCCATTTCCATGCTGTACCTGGATGAACACGACAAGGTGGTCCTCA
AAAACTACCAGGAGATGGTGGTAGAGGGTTGTGGCTGCCGTTAA

>LocBMP16 (Spotted gar; Accession number AHAT01023895)
ATGTTCCCTGCTAATCTCCTGGTCCTGATGGTCCTGTTGCTACCTCAAGTCGTGTCAGCTCGCGAGGGGGCTTCTGACCCGAG
CACACCGCCACAGCTCAGCCTGGCTCAGACCATTCAGAACTTGCTGCTGACCCGGCTGGGCCTGCGGAACCGGCCCGAACCCC
GTCCAGGAACCCCCATCCCGCAGTACCTGCTGGACGTCTACCGGTTCCACTCCCAGGACTTCCACCTGATTCAGGACCCGGGC
TTCAGGTTCCCCACCAGGCATACCCAGGGCGCCAACACCCTCCGCAGCTTCCACCCCCTGGAGACTCCAAGACACCTCCTCCC
GAGTAGCCAGGGGGAAACAGACAGCTTCCACATCGTATTCAACGTCTCGTCCATTCCCAGGGACGAGCGGGTGACATCAGCTG
AGCTTCGCCTCTCCCGCAGGGAGCAGTCTGTGCCCTCCGGGACTCAAAGGGTTAACTTATACCAGCTCCCTGATCCCAGGAGC
CCAGCAGACACCGAGACCCGGCTCCTGGCGTCCAGGCTGCTGGCAGATGACCACGATTCGCACTCCTGGGAAAGTTTCAGCCT
TGGCGCGGACTTGTTTGACACACCCAATGGCCAGACGCACCATGTGGCATTCGTGCTGGAGGTGACGCACCTGAACAGCAGCC
GCCCTCCGGAGCAGCGAAGGGAGCGTTTGAGGGTCAGGAGATCCGTGGGCCAGGACCAGCCCAGCTGGGCCCAGGAGAGACCC
CTCCTGGTGACCTACAGCCACGACGGGAGGGGCCAGCCCCTGGCCCAGAAGAGGGTCGAGCCCAGGAGGACCAGGGGCCCCTG
GGAGAGAAGGGTGAAGCGCAACGGCAGGTTGCACAAGCTCAAGAAGATGGCCAGGGCGCGCTGCAGGAGGCACCCCCTCTACG
TGGACTTCAAGGACGTGGGCTGGAACAAGTGGATCGTGGCTCCCAGTGGCTACCACGCCTTCTTCTGCATGGGGGAGTGCCGC
TTCCCGCTGGCCGACCACATGAACTCCTCCAGCCACGCCATGGTGCAGACGCTGGTCAACTCGGTCAACAGCAACGTGCCGCG
GGCCTGCTGCGTGCCCACCACACTCAGCCCCATCGCCATGCTCTACCTGGACCAGCACGACCGGGTGGTGCTCAAGAACTACC
AAGACATGGTGGTGGAGGGCTGCGGCTGCAGATAA

>DreBMP2a (Zebrafish; Accession numbers NC_007128, CABZ01062106
CAAK05040443, NM_131359, BC163036, BC163048, EH593053, EH610560, CT616046,
CT616047, BM187205 and AI878481)
ATGGTCTCGTCCACCGCCGCGCTAATGGTTCTGATGGTCACTCAGGTGTTTTTTGCGGGCTCCAGTGGACTCGTTCCTCAGGT
CGGCCGGTCCTCTCTGTCTCCGGACGTCCTGCACGCGTTTGAACTTCGGCTTCTGAGCATGTTTGGGCTCCAGCACAGACCCA
CTCCCAGCAGCTCCGCAGTGGTTCCTCAATACATGCTGGACTTGTACTCGGCGCACTCCGTGAACGCAGAGCAGGTTAGCAGA
CCCAGAGCTCACTTAGGGAAAGGCTCCGAGAGATCCGCCAGCAGAGCCAACACTATCAGGAGCTTCCACCATGATGAATCTAC
AGAAGATCTGTCCAGCTCCAGTGTTAGGACGACGCAGCGGTTCCTCTTCAACCTGACCTCCATCCCAGACGAGGAGTTGGTGA
CGTCAGCAGAGCTGCGCGTTTTCAGAGAGCAGATCGTGAGTTCGCTAAACAACGCAAGTGCGGGATTTCACCGCATTAACGTT
CACGAAATCATCAGGCCGTCCGGTTCTCTTCAGGAGCCCATCACTAGACTCTTGGACACCAGGCTTGTTCAACACAGCTTGAG
CAAATGGGAGAGCTTTGACGTAACTCCTGCTGTATTAAAGTGGACCACAGACGGACACCCTAACCATGGGATATTGGTGGAAA
TATCACACCCAGATCAAGACTCCAGAAAACATGTGCGCGTAAGTCGGTCCCTTCACAACAATGAGGACACGTGGTCTCAAATG
AGGCCGCTTTTGGTGACCTACAGCCACGATGGAAAGGGAAACGTGCTGCATTCTCGTGAAAAACGACAAGCGAGGAATAATAA
GCAGAGAAAGAAGCATAAAGCAAACTGCAGGAGACATTCCCTGTACGTGGACTTCAGCGACGTAGGCTGGAATGACTGGATTG
TGGCGCCGCCGGGTTACCATGCATTCTACTGCCAGGGCGAGTGCCCGTTTCCACTGGCAGACCACTTAAACTCCACTAACCAC
GCTATCGTGCAGACACTGGTGAACTCGGTGAACAGCAACATTCCCCGAGCCTGCTGCGTGCCCACTGACCTCAGTCCGGTTTC
TCTGCTATATCTGGACGAGTACGAGCGGGTCATCTTAAAAAACTACCAGGACATGGTGGTGGAGGGATGCGGGTGCCGCTGA

>DreBMP2b (Zebrafish, Accession numbers NC_007131, CABZ01042685, NM_131360, BC114256, BC076076, D49971, DRU82232, EV759768, EV756104, CT619514, DR722644, CK400445, DR719530, CV481651, CV481676, CA588097, DV586788, CK692216, AI959065, BI889767, CK696508, EB853879, EB852282, BI889738, CT619515, FP160298, FP160299, EB835261, EB976237 and CA588098)

ATGGTCGCCGTGGTCCGCGCTCTCACGGTGCTGTTGCTCGGTCAGGTGTTGCTGGGAGGTGCCGTTGGACTCATTCCCGAGAT
CGACCGACGGAAATACAGTGATTCGGGGAGACACACACCGGAGCGAACTGATACAAACTTCCTGAACGAGTTTGAGCTACGCT
TGCTCAATATGTTCGGATTGAAGCGAAAACCCACCCCAAGCAAATCGGCAGTGGTCCCTCAGTACATGCTGGACTTGTATTAT
ATGCACTCTGAAAACGATGACCCGAACATTCGGCGCCCGAGGAGCACTATGGGAAAACATGTAGAAAGGGCAGCCAGCAGAGC
AAACACGATACGAAGTTTTCATCACGAAGAGGCTTTCGAGGCACTGTCCAGCCTGAAAGGAAAAACAACGCAGCAGTTTTTCT
TCAACCTTACCTCCATTCCTGGCGAGGAGCTGATCTCCGCTGCGGAGCTGCGCATTTTCAGGGACCAAGTTCTCGGAGATGCC
AGTACGAGTGGCTTCCACAGAATTAACATTTACGAGGTGTTCAGGCCAGCTTTGGCCCCCTCCAAAGAGCCTCTAACCAGACT
TCTGGACACCCGTCTGGTGCAGGACTCTCACACGCGCTGGGAAAGCTTCGACGTGGGTTCAGCTGTGGCACGCTGGGCCCGCG
AATCCCAGCACAACCATGGGCTCCTTGTAGAGGTGCTCCATCCTAAGGAGTCAGAAGTATCCGAGGAGGCTGAGAGCAACCGG
AGGAAGCACGTGAGGGTCAGTCGTTCCCTTCACGCGGATGAGGACTCGTGGGCACAAGCCCGACCTCTGCTGGTAACCTACAG
CCATGACGGTCAAGGCACAGCCGTCTTGCATTCGAACCGAGAAAAGCGGCAGGCTCGACGAGGGCAAAAGCCGAGGAGAAAGC
ACCACCAGCGCTCGAACTGTAGGCGACATGCTCTCTATGTGGACTTCAGTGATGTCGGCTGGAACGAGTGGATCGTGGCACCG
CCAGGCTATCATGCTTTCTACTGCCATGGCGAGTGTCCGTTCCCTCTGCCGGACCATCTAAACTCCACCAACCATGCCATTGT
CCAGACGCTGGTGAACTCGGTCAACTCCAACATTCCCAAAGCCTGTTGCATCCCGACGGAGCTCAGCCCTATCTCACTGCTGT
ACCTGGACGAGTACGAGAAGGTCATTCTTAAAAACTACCAGGACATGGTGGTGGAGGGCTGCGGGTGCCGATGA

>DreBMP4 (Zebrafish; Accession numbers NM_131342, D49972, DRU82231, DRU90122, DU730983, CT674533, CT619393, EB983652, CT599423, BI892288, BI886011, CK683872, CT599424, CN507552, CT619394 and NC_007128)

ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTATTATGCCAAGTCCTACTGGGAGAAAGCAGCTATGCTAGTCTGATACC
CGAGGAAGGGAAGAAGAAAGCGTCGGCTCTTCACCTGGCTCAGAGTCATGAGCTGCTGCGGGACTTTGAAGCCACGCTGCTGC
ACATGTTTGGCCTGCAGAGGCGTCCCAGACCCAGCCACAGCGCCGTCGTACCACAGTATCTGCTCGACCTCTACCGCCTGCAG
TCAGGGGAGCTGGAGGAGGCAGGAGCGCAGCACGTCAGCTTCGACTATCCTGAAAGATCCACCAGTCGAGCCAACACCGTGAG
AGGATTCCATCATGAAGAGCACCTGGAGGAGCTGCAGTCAGACGGCTCCCAGGAGACTCCTCTGCGATTCGTTTTTAATCTCA
GCAGCATCCCAGAGGACGAACTCATATCCACCGCAGAGCTTCGCGTCTACAGGCAACAAATAGATGACGCCTTCTCAGACCCA
GATCAAACAGGGGACCATGGTTTGCATCGGATAAACATATATGAGGTGTTGAAGGCGCCACGGGAAGGACAGCTCATCACGCA
GCTCCTGGACACACGTTTGGTGAGGCACAACACCTCCAAATGGGAAAGTTTCGACGTTAGCCCTGCAGTGTTGCGCTGGACCC
AAGAAAAACGCTCTAATCATGGCCTTGCTGTGGAGGTTGTACAAATGAAGCGAAACCCAGTTCAAAAGGGACGACATGTTCGT
GTAAGTCGCTCCGTGCATCCTCTTCCGGATGAAGAGTGGGACCAGCTACGCCCCCTGCTGGTCACATTCGGACATGACGGCAA
AAGTCACCCGCTGACTCGGCGAGCGAAACGCAGCCCTAAACAAAGAGGTCGAAAGCGTAATCGTAACTGCCGGAGACATGCGC
TGTATGTGGATTTCAGCGACGTAGGCTGGAACGACTGGATTGTGGCACCGCCTGGATATCAGGCGTATTACTGTCATGGAGAG
TGTCCCTTTCCATTAGCCGATCATCTCAACTCCACCAATCACGCTATCGTACAGACACTGGTGAACTCGGTGAACACCAATAT
CCCCAAAGCCTGCTGCGTGCCCACTGAGCTCAGCGCAATCTCCATGCTTTACCTGGACGAAACGGACAGGGTGGTGCTGAAAA
ACTATCAGGAGATGGTGGTCGAGGGGTGTGGCTGCCGCTAA

>DreBMP16 (Zebrafish; Accession numbers NM_001171776, FN400947, CABZ01023011, CAAK05042509, FN400946, CR381630 and CAAK05042510)

ATGTTCCCTGCTAGCCTACTGGTTCTGATGATCCTGCTGCTACCTCAAGCCTCATCAGGTCACCAGGAGGGTCCCAGTCAGAC
TCATGGAAAGCTGGACAGCCTTCTGGAGCCTAGTTTGGCCCATACCATCCAGAACCTCCTGCTAACCCGCCTGGGCTTACAGT
CGCACCCCAACCCCAGTACAGAGGCACAAGTTCCCCAGTATCTTCTGGACTTGTATCGATTCCACACTCAACAGTACCATCTC
ATCGAGGACCCAGAATTTAGTTACCCCTCAAAGCATGTGCAAGGAGCCAACACAGTCCGATGCTTTCACCACACAGATTCTCC
GACTCCGAGTCTTCCAGAAGAGCAGAAGACCACAGTGGACGGCATCCACATAGGCTTCAACCTGTCCTCCATTCCTTCAGAAG
AAAGTGTGGTGTCCGCAGAGTTACGTCTACTCCATGAAGGGTCAAGTGGAGGTTCTCACGTGGCCAGTCTTTACCTTTCCAAC
CATGACCCTTCATCAAAACCCATTCTGCTCCATTCACGTCAACTGACCAGAGACAGGAAGAGTGCCCAACTCTGGGAAACCTT
TCCCCTGGATAGAGAAGTGTTTCAGAATTTATCCAAGACGTCAGGAAGCCTGTCTTTTATCCTGGATGTTATTCCTGATAATA
ACAGCAGCAGTACTCTTCCTAAAGACAGACATTTGCGAGTACGCAGATCCACATTGCAGGACGGGCCTACTTGGGAAAGGCAA
AGGCCTTTATTAGTGACCTACAGTCACGATGGCCGCAGCGAGCCTTTTGTGAATCTCAATAAGCGTAAAAGCAGCCGAAGCAG
GAGCAGATGGACCAGGTTTAAGGGTGATAGGGGTCAGGGCTCTGACTGGAATGAAAGGCGAGTGAAACGAAACAACAGGAGGG
CCGCAAAACTCAAAAGACTCTCCCGCGCTCGTTGTCGTCGGCACCCGCTGTACGTGGACTTCAAAGATGTCGGCTGGAATAAA
TGGATCATAGCTCCATCTGGTTACGATGCTTTCTTTTGCCTCGGAGAGTGCCGCTTCCCACTGGCTGACCATATGAATTCATC
AAGCCATGCCATGGTGCAGACGCTAGTCAGCTCTGTAAACGGAGCCGTGCCGCGACCATGCTGCGTCCCTACAGCTCTCAGCC
CAATCGCTCTACTCTTTCTGGACCAGGAGGAACGTGTGGTGTTGAAAAACTACCAAGACATGGTGGTGGAGGGATGTGGCTGT
CGATAG

>IpuBMP2a (Channel catfish; Accession numbers JT279003, JT446579, JT413282 and JT376107)

ATGGTGCCGGTGTTCGCTTTGCTCATCGCGCAGCTCTTCTTCGGAGGCTGCACGGGTCTCCTGCCGGAGATCGGGCGCCGCAA
GAACAGCCCTCACCTCTCCCAGGACGCGCTGGACGCTTTCGAGCTTCGTCTGCTCACCATGTTCGGACTGAAGCAGAGGCCGA
GTCCGAGCAAATCAGCAGTGGTGCCGCAGTACATGCTGGATCTGTTCACCACGTATTCAGCCAACAGTGCCGAGCAAAAGCAC
CAGAAAGCGAAGAGCACAGTGGGCAGGAGCGCAGACAGATCCGCTAGTCGGGCGAACACAGTACGAAGCTTCCATCATGATGA
ATCTACAGAGGAGCTGTCACGTCACAGTGGAAAGACGACTCAAAGGTTCCTCTTTAATATGAGCTCCATCCCAGGAGAGGAGT
TTGTCACATCAGCAGAGCTGAGGATTTACAGGGAGCAGGTGATGACGAACAGTAGCAGCAACAGCAGTGCGGGGTACCATCGC
ATTAACATCCACGAGATCATCAAGCCTGCACTTCCGTTTAAAGAGCCCATCACAAGACTTCTAGACACCAGACTGGTTCAGCA

CAGCCTGAGCAAGTGGGAAAGCTTTGATGTGAGTCCTGCTGTCTTGCGGTGGACCAGGGAGAGTCACACCAACCATGGCTTTG
TGGTGGAGGTGGTTCACCCAGGCGATGCAGAAGGAGACTACAAGAGACATGTCAGAGTCAGTAGGTCGCTGCATGACAGCACA
AACTCATGGCCTCAGATGAGGCCACTGCTGGTGACGTTCAGCCACGACGGCAAAGGACACGTGCTCCACTCTAGGAAGAAACG
GCAAGCGCGAGTCAACAAGCCAAAGAAGAAACACAAAGCCAACTGCAGGAGGCACTCGCTCTATGTGGACTTCAGCGACGTGG
GCTGGAATGACTGGATAGTGGCACCACCAGGCTACCATGCCTTTTACTGCCAAGGAGAGTGTCCTTTCCCTTTGGCAGACCAC
CTAAACTCCACTAACCATGCCATTGTACAGACACTGGTCAACTCGGTGAATAGTAACATTCCCAAGGCATGCTGTGTGCCCAC
AGACCTGAGCCCCATTTCCCTGCTCTATTTAGATGAGTATGAGCGAGTGATCCTGAAAAACTACCAGGACATGGTGGTGGAAG
GCTGTGGGTGCCGCTGA

>AmeBMP2b (Mexican tetra; Accession number DQ915172)
ATGGTGGCCGCGGTGCGCTCCGTCACGGCGCTGCTGCTGCTCGCTCAGGTGCTGCTGGGAGGCGTTGCCGCCGCGCTCGTGCC
CGACGTGGAGCGGCGCAGGCCCGGAGAGCAGAGCTCGAGCTTCTTGCGCGAGTTCGAGCTGCGCCTGCTCAACATGTTCGGCC
TGAAGCGCAGACCCACGCCCAGCGGCGCCGCCGTCGTGCCCGACTACATGCTCGAGCTCTACCACCTGCACGCGGGGGGTGAC
GGAGGAGCGCACGAGCAGCACCTTCAGCAGGGCGCGCGGAGGAGGCGCGCGGACAGGGCGGCCAGCCGGGCCAACACCATCAG
GAGCTTCCATCACGAAGAAGCTTTTGAGGCCCTCTCCAGCCTGAAAGGAAAGACAACACAGCAGTTCTACTTCAACCTCACCT
CAGTGCCAGCAGAGGAGCTAATCACAGCCGCGGAGCTGCGCRTCTTCAGAGACCAAGTTCTAGGTGACTTTACAAAGAGCAAC
AGCAGTGTTGCAGGTGGTTTTCACCGCATAAACATTTTYGAGGTGTTCAGGCCAGCTCATTCCCTATCAGAGGAGCCTCTCAC
CAGGCTGGTGGATACTCGTCTGGTGCAGGACTCTCATTCGCGGTGGGAGAGCTTCGATGTGGGCTCAGCAGTGGCGCGCTGGA
CCGCACACTCGCACCACAATCATGGCCTCCTGGTGGAGGTGCTTCATCCTGAGGAGAGCAGTGGAGAGGAGCAGGCCGAGAGG
AACAGGAGGAGGCACGTAAGGGTGAGCCGTTCCCTGCACGCGGACGAGGACTCTTGGTCGCAGGCCCGGCCGCTGCTGGTTAC
TTACAGCCACGATGGTAAGGGTACGCCCCTCCTGCAGCCACACAGAGAGAAACGGCAGGCGCGGCGTGGGCCCAAGCAGCGGC
GGAAACCACACCAGCGCACCAACTGCCGCCGGCACCCGCTTTACGTGGACTTCAGCGACGTGGGATGGAATGAGTGGATCGTG
GCTCCGCCTGGCTACCATGCCTTCTACTGCAACGGCGAGTGCCCCTTTCCGCTTTCTGACCACCTCAACTCCACAAACCACGC
CATAGTCCAGACTTTGGTGAACTCGGTAAACTCCAACATTCCTCGGGCTTGCTGCGTGCCCACAGAACTCAGCCCCATCTCTC
TGCTCTACCTGGATGAGTATGAAAAAGTTGTACTGAAAAACTATCAGGACATGGTTGTAGAGGGCTGCGGTTGCCGGTGA

>IpuBMP4 (Channel catfish; Accession numbers JT406104 and FI862280)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTATTATGCCAAGTTGTACTGGGAGAGAGCAGCCATGCTAGTCTAATACC
CGAGGAAGGGAAGAAAAAGGCACTGGGTAATAACCCAGGTCAAAGTCATGAACTGCTGCGGGACTTTGAGGCCACGCTGCTGC
ACATGTTCGGCCTGCAGAGGCGGCCACGGCCCAGCCGCTTGGCAGTGGTGCCCCAGTATCTGCTCGACCTCTATAGGCTGCAG
TCAGGAGAAGTGGAAGAGGCTGAAGCTCACGATGTCAGCTTTGAATATCCTGAGAAGTCCACTAGCCGAGCCAACACTGTGAG
AGCATTCCATCATGAAGAGCACATGGAGCAAGTGCCGTCAGACAGACCGTACAGCATCGAGACTCCTCTGCGCTTCTTCTTTA
ACCTCAGCGGCATCCCTGAGGATGAGCTGCTTTCCACGGCAGAACTGCGGCTGTACCGGAAGCAGATTGAGGAGCCTGGTTTG
GAGTCCCAGCATGTGGGAAATGAGGACCTCCATAGGATCAATGTGTATGAGGTGCTGAAGCCACCACAGCCCAGGCAGCTGAT
CACACGCCTCCTGGACACACGCCTTGTTAGCCACAACACCACCAGCTGGGAAAGCTTTGACGTGAGCCCGGCTGTATTGCGTT
GGACCCACGAGCGGTTACCAAACTATGGACTTGCTGTAGAGGTTCTGCATCTGAACCAAACACCGCGGCACCAAGGCAAGCAC
GTGCGCATCAGCCGCTCACTGCACCACACTTCAGAAGAAAACTGGGACCAGCTGCGTCCTTTGTTGGTCACTTTCGGCCATGA
TGGCAAAGGCCACCCACTGACACGTAGGACCAAACGCAGCCCCAAGCAGCGAGGCCGTAAGCGCAACCGCAACTGCCGGCGAC
ATGCGCTGTATGTGGACTTTAGTGACGTAGGCTGGAATGACTGGATAGTGGCACCACCGGGCTATCAGGCCTATTACTGCCAC
GGGGACTGCCCTTTCCCTCTAGCAGACCATCTGAATTCCACGAACCATGCTATTGTACAAACACTGGTCAACTCGTTGAACAG
CAATATTCCCAAGGCCTGCTGTGTGCCCACCGAGCTGAGTGCCATCTCAATGCTCTACCTGGATGAGACTGACAGAGTAGTCT
TGAAAAACTACCAGGAAATGGTGGTGGAGGGCTGTGGCTGCCGCTAG

>IpuBMP16 (Channel catfish; Accession numbers JT415761, JT155730, JT424125 and
JT148139)
ATGTTCCTTGCTAGTCTTCTAGTTCTGATGACCCTGCTGCTTCCTCCAGCCTTGTCAGCTCAACAGGTGGATTCCAGTCACTC
TGATGGGAACCCAGGAATCTCTCTCGAGCCCAGTTTGGCCCAGACCATCCAGAACCTCCTCTTGACCCGACTGGGCCTACAGT
CCTATCCAAACCCTCACCCAGGAGCAGTGGTACCTCAGTATCTCCTGGACCTCTATCGCTTCCACACACAGCAGTACCACCTA
GTAGAAGATCCAGATTTCAGCTTTCCTGTGGAGCATGTAGAAGGTGCCAACACTGTTCGCAGCTTCCACCATGCAGTGTCTTC
AGACCTTCACATCCCTGATGAGAAGAGTGACAGGATCCATATAGCTTTTAACCTGTCCTCTATTCCTTCCTACGAGAATGTGG
TCTCAGCAGAGCTGCGTTTTCTGCGTGGAGGAACACGCTGGGGTTCAGAGCCTCACACTGTCAGCCTCTATGTCTCTGGTGGA
AACATTGAGCCAACACTTCTCCATTCACGACAGTTAGTTAGAGAGCCCTCAGGGTCTTGGGAGGCCTTTCCTCTGAACGCAAA
ACTGTTTCAGAAGTTGTCTGATAGTGCAGGGAAAGTGGCCTTTATCTTGGAAGTGACCTCAGATAACAGCAGTCACGCTGTGG
AGGACATGGACACAGAGGGACACTTGAGAGTGCGTAGGTCTGCAGGTGAAGATGACTACAGCTGGGCACGCCAAAGACCACTA
CTGGTCACTTACAGCCATGATGGGCGCAGTGAACCATTCATTCCCTTTGGAAAGAAGAAGACCCCCAGGGGTAGGAGGGGCGG
TAACAAGGCTAAAAAGCTCAAAGGTCAGGCGTCACAGGTAAGCTGGCGTGATGTGTGGACAGACAAGCGGGTAAAGAGGAATA
ACGGCGGTCGAGCAGCCAAGCTGAAACGCCTGTCTCGCTCTCGATGCCGCCGCCATCCCCTCTACGTGGACTTCAAAGATGTC
GGGTGGAACAAATGGATCGTTGCACCGACTGGTTACGACGCCTTCTTCTGTCTGGGTGAGTGCCGCTTTCCTCTCGCCGACCA
CATGAACTCGTCCAGCCACGCCATGGTGCAGACACTGCTGAACTCGGTTAATGGAGCCGTGCCACGAGCCTGCTGTGTGCCTA
CGGCCCTCAGTCCAATCGCGCTCCTTTATCTGGACCAGGAAGATCATGTGGTTCTGAAAAACTATCAGGATATGGTAGTGGAG
GGCTGCGGCTGCCGGTAG

>GacBMP2 (Three-spined stickleback; Accession numbers DW637392, DW637393 and
AANH01008393)
ATGGTCGCCGTGGTCCGCTCTTTCATGGTACTGCTGATCGCTCAGGTGTTGCTGGAAGGTGCTACGGGACTCATCCCCGAGGT
CGGCCGGAGGAAATACAGCGAGTCCGGGAAGCAGACCCCCGAGCGGTCGGAGAGCTTCCTCCACGAGTTCGAGCTCCGGCTTC
TCAACATGTTCGGACTGAGGCGCAGGCCGACCCCGAGCAGGCAAGCCGTGGTGCCGCAGTACATGGTGGATCTTTACCGCATG
CACTCAGCGAACGGAGACCACAGCGCCAAACGACCCAAGAGCATGGGGAGGCACGCAGAGAGGGCTGCGAGCAAGGCCAACAC

GATTAGAAGCTTTCACCACGAAGAGTCCATGGAGGCCCTGGCCCGGCTGAAGGGTAAAACAACCCAGCAGTTCTACTTCAACC
TCACTTCGATCCCCGAGGAGGAGCTCATCACCTCCGCAGAGCTCCGCGTCTACAGGGACCAGGTCATGGGAGCCGCGGCCCCC
GACAACAGCGCCGGGAACAGCAGCACTAGTGTCAGCGCGCCGGCGGGTGGCTTCCATCGCATCAACATTTATGAGATATTTGG
AGTTCCCTCCGCCCAAGGTGGGGAGCCCCTGGCGCGTCTGCTGGACACTCGGCTAGTGCGGGACTCTTTGAGCCGCTGGGAGA
GCTTCGACGTCAGCCCCGCCGTGTCTCAGTGGACCGCCGGGAAAGGTCGCAACCATGGCTTCATGGTGGAGGTGCATCACCCA
GAGGAGGGGGCGACGGACGGGGAGCACGCCCAGAGACGCAGCAGGCACGTCAGGGTGAGCCGGTCCCTGCACCAGGACCGGGA
CTCGTGGCCTCAGGCTCGGCCTCTGCTGGTGACGTACGGTCACGACGGCCGCGGGGACTCGGTGCTCCACACGCGGGAAAAGC
GCCAGGCGGCGCTCCGCAAACAGCGCAGGAAGCACCAGCACAAGGCCAGCTGCAAGAGGCACGCCCTGTACGTGGACTTCAGC
GACGTGGGGTGGAACGAGTGGATAGTGGCGCCGCCCGGCTACCACGCCTTTTACTGCCAGGGGGAGTGTCCGTTCCCCCTGGC
GGAGCATCTCAATTCCACCAATCACGCCATCGTGCAGACGCTGGTCAACTCGGTCAACTCGAACATCCCCAGAGCCTGCTGCG
TGCCCACTGACCTCAGCCCCATCTCCCTGCTCTACCTGGACGAGTACGAGAAGGTCATCCTGAAAAACTACCAGGACATGGTG
GTGGAGGGGTGCGGCTGCCGGTGA

>GacBMP4 (Three-spined stickleback; Accession numbers DT969085 and AANH01003910)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTAATATGCCAAGTCCTGCTGGGAGAGAGCAACCATGCCAGTCTGATACC
TGAAGAAGGGAAAAAGAAAGTACCGGGCCTGCAGGGTCGTTCGGCCTCTCAGAGCCATGAACTGCTGCGGGACTTTGAGGCCA
CGCTGCTGCACATGTTTGGCCTCAAGAAGCGGCCGCGGCCCAGCCGATCGGCCTCGGTGCCTCGGTACCTGCTGGACCTCTAT
CGGCTGCAGTCAGGGGAGGCGGAGGAGGCCGGGGTGCACGACATCGCTTTCGAGTACCCCGAGAGGTCAGCCAGCCGGGCCAA
CACTGTGAGGGGCTTCCACCACGAGGAGCACATGGAGCGGGTGCACAATCTGGACGATGGAGAAACCGCCCCCCTTCGCTTCC
TGTTCAACCTCAGCAGCATCCCGGAGGACGAGCTGCTCTCTTCAGCCGAGCTCAGGCTCTACCGCCATCAGATCGACGAGGCC
ATCGGCGGCGCCTCGTCAGACAACCAGGGGCTTCACCGGATAAATGTGTACGAGGTGCTGAAGCCCCCCCGGCCGGGCCAGCT
CATCACGCAGCTCTTGGACACGCGGCTCGTGCGGCACAACGCGTCGCGCTGGGAGAGCTTCGATGTGAGCCCGGCGGTGCTGC
GCTGGACTCGCGAGAGCCTCCCGAACCACGGGCTGGTGGTGGAGGTCCTGCACCTCAACCAGACGCCGCGCCACCAGGGCCGG
CACGTCCGCGTCAGCCGCTCGCTGCACCGGGAGCCCGGCGAGGACTGGGAGCAGCTACGTCCCCTCCTGGTTACCTTTGGCCA
CGACGGGAAAGGTCACCCGCTGACCCGTCGGACCAAGCGCAGCCCCAAGCAGCGGGGCCGCAAGCGCAACCGCAACTGCCGGC
GCCACGCGCTATATGTGGACTTCAGCGATGTAGGCTGGAATGACTGGATAGTGGCGCCCCCTGGTTACCAGGCCTATTACTGC
CACGGGGAATGTCCCTTTCCTCTGGCGGATCATCTGAACTCAACCAACCATGCCATTGTTCAGACGCTGGTGAACTCTGTGAA
CAACAACATTCCCAAGGCCTGCTGCGTGCCCACAGAGCTCAGCGCCATCTCCATGCTCTACCTAGACGAACACGACAAGGTGG
TCCTCAAAAACTACCAGGAAATGGTAGTGGAGGGCTGCGGCTGCCGCTGA

>GacBMP16 (Three-spined stickleback; Accession numbers DT979515 and AANH01003009)
ATGTTCCCTGCTAACCTCCTGCTCCTCATGGTCCTGCTCCTACCTCAAGCCTCGTCTGGTCGCCAGGGTGGAGTCACCAGCAA
GACCGACCATGGCCGGGTGTCCTCCATGCCTTCCTCCTTGTCGCCCCCACCAGCTGGTTCCCTCCTCCTGGAGCCCAGTCTGG
CACAGACCATCCAGAACCTCCTCTTGAGCCGCCTGGGCCTACAGTCGCAGCCCGACCCTCGGCCTGGAGTGCCGGTGCCTGGG
TACCTCCTTAATCTTTACCACTTCCACCAGCAGCAGTACCATCTAGTGGAGGATCCCTCGTTTAGCTTCCCGAGCCAGCACAT
TGAGCAGGCCAACACTGTACGCAGCTTTCACCACAATGAGCCCCTCGGAGGCAATCCTGTAGCAGTGGATCGTAAGAGGGTGC
ACATCTCCTTCAATGTATCCTCCATCCCTCAGGATGAGAGGGTGCTCTCCGCTGAGCTTCGGCTTCTCCGCAGTGACAGAGCT
TCACTAGGTCCTGGGGTCCACAGACTTAACTTATACCTCACTAAGCACCATGAGGACCCTGAGCCCACCCTGCTGGAGACAAG
GTTACTCACCACTGGCCTCCACGGTCATAAAGAAGGCGGCTTCTGGGAGGCTTTTAGCTTAAATGCAGAGTTCCTCCATGAGG
CCCTTAGTGGGACTGGCAGTCTGGGCTTCCTCCTAGAGGTCGGAATTGAGAACAGCACCACATTGATCCCTGAACAAAACCTC
TACCCTGCCGCAGCTGAGGAGGAGACAGAGAAACACAAACAGGGACACCTGAGGGTATGCAGGTCTGTGGGTCAGGACGAGCA
TAGCTGGGCCCAGGAGAGACCCCTCTTGGTGACGTACAGCCATGACGGCCGTGGAGAGCCCTTGGTCAAACATGGCAGGAGGA
ACCCCAACGGAGGCCAGAGGGTGAGAGGGAGAAAAGGGACCAAAGACAGGACCAGGGGCAGCAGCAATAGCCGCAGTAGAGAC
CCGGCTTCGGGCAGGGCCAAAAAAACAGGGTATACAGTGCAGGGCAGGGGGAGTGTTAAAAGAGGAGTCAGCTGGAGCGATGC
CGGAAGGGTGAAAAGAAACGGCGGTCGTGCTGCAAAACTGAAGCGCCTTTCTCGTAACAGATGTCGCCGCCATCCTCTCTATG
TAGACTTCAACGATGTGGGGTGGCACAAGTGGATCATTGCACCAAGCGGCTACGACGCCTTCTTCTGCTTGGGCGAGTGTCGC
TTTCCTCTGGCCGACCACATGAACTCCTCCAGCCACGCCATGGTGCAAACTCTGGTGAACTCGGTGAATGGAGCTGTGCCGCG
GGCCTGCTGCGTCCCCACCTCCCTCAGTCCCATCGCCCTGCTCTACCTAGACCCTCAAGACCGAGTGGTGCTGAAGAACTACC
AGGACATGGTGGTGGAAGGCTGCGGCTGCCGGTAG

>OniBMP2 (Nile tilapia; Accession number AERX01005559)
ATGGTCGCCGTGGTCCGCTCTCTCATGGTACTGCTGCTCGCTCAGGTGTTGCTGGAAGGTGCTACGGGACTTATCCCCGAGGT
CGGCCGGAGGAAATACAGCGAATCCGGGAAGCAGACCCCAGAGCAGTCAGAGAGCTTCCTCAACGACTTTGAGCTTCGGCTTC
TCAATGTGTTTGGACTGGGGCGCAGGCCGACCCCGAGTAAGCAAGCCGTGGTGCCGCAGTATATGGTGGACCTTTACCGTATG
CATTCAGCAAACGGAGACCATAGCACCAAACGACCCAAGAGCATGGGGAAACACGCAGAGAGAGCCGCCAGCAAGGCCAACAC
GATTAGAAGCTTTCACCATGAAGAGTCTATGGAGGCCCTGGCCAGCCTAAAAGGCAAAACGACACAGCAGTTCTACTTTAATC
TCACTTCTATTCCAAATGAGGAACTCATCACCTCTGCAGAGCTACGCATTTACAGGGATCAGGTCATGGGATCTGCAAGCCCC
ACCGAAAGCAGCATTAGTGACAGTACTCCTGCTAGTGGCCTCCATCGTATCAACATTTATGAGATATTTGGAGTTCCTGCCAC
TCAGGGCAAGGAACCTCTACTACGTCTGCTGGACACTCGAATGGTGCAGGACTCTTTGAGTCGCTGGGAAAGCTTTGACGTCA
GTCCCGCTATATCTCATTGGACCTCTGGGAATGGCCACAATCACGGCTTCTTGGTGGAGGTGATTCACCCAGAGGAGGGGGAG
ATGGACAGCAAGGACGCCCACAGACGCAGGAGACATGTCAGGGTGAGCCGGTCGCTACACCAGGACCAGGACTCGTGGCCTCA
GGCTCGACCGCTGCTTGTGACATATGGCCACGACAGCCGCGGGGGCTCAGTGCTCCACAAACGGGAAAAAAGGCAAGCAGCAC
CCCGCAAACAACGTAGGAAGCACCAGCACAAGGCAAGCTGCAAGAGGCATGCCTTGTACGTGGACTTCAGTGATGTGGGCTGG
AATGAGTGGATAGTGGCACCCCCTGGCTACCATGCCTTTTATTGCCATGGGGAATGTCCTTTCCCCCTAGCAGACCACCTCAA
TTCTACTAACCACGCCATTGTGCAGACACTTGTCAACTCAGTCAACTCCAACATCCCCAGAGCCTGTTGTGTGCCCACTGACC
TCAGCCCCATTTCTTTGCTCTACCTGGACGAATTTGAGAAAGTAATCCTGAAAAACTACCAAGACATGGTGGTGGAGGGATGC
GGCTGCCGGTGA

>OniBMP4 (Nile tilapia; Accession numbers AB084664 AERX01008958 AB084664 and GR604846)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTAATATGCCAAGTCCTGCTGGGAGAGAGCAACCATGCTAGTCTGATACC
TGAAGAAGGAAAAAAGAAAGTACCTGGCCTGCAGGGTCGTTCGGCCGCTCAAAGCCATGAACTGTTGCGGGACTTTGAGGCCA
CGCTGCTGCATATGTTCGGCCTCAAAAGGCGGCCGCGGCCCAGTCGTTCAGCCACTGTGCCCCGCTACCTGCTGGACCTTTAT
CGGCTACAGTCGGGCGAGGCTGAGGAGGCCGGAGCGCACGACATTGCTTTTGAGTACCCAGAGAGGTCAGCCAGTCGGGCCAA
CACTGTGAGGGGCTTCCACCATGAAGAGCACATGGAGCAGCTACATGAGCTGGATGACGGAGGACCCATGCCCATTCGTTTTA
TGTTTAACCTCAGCAGCATCCCAGAGGACGAGCTTCTGTCTTCCGCAGAACTGAGGCTCTACCGTCATCAGATTGATGAGGCC
ATCGCTGACGCCATCTCACACGACCAGGGGCTTCACCGAATAAACGTGTACGAGGTGCTCAAAGCCCCCCGACCTGGGCAGCT
TATCACGCAGCTTCTGGATACGCGGCTAGTGCGCCACAACGCATCGCGCTGGGAGAGCTTCGACGTCAGCCCCGCCGTGCTGC
GCTGGACTCGTGAGCGCCTTCCCAATTACGGGCTGGCTGTGGAGGTTCTGCACCTCAACCAGACTCCACGTCACCAGGGCCGA
CACGTTCGCATAAGCCGCTCGCTGCACCAAGAGCCCGGTGAAGACTGGGAGCAGCTACGCCCCCTCCTGGTAACCTTTGGTCA
CGACGGAAAGGGTCACCCGCTGACTCGCCGGACCAAGCGCAGCCCGAGACAGCGGGGCCGTAAACGCAACCGCAACTGCCGGC
GCCATGCGCTGTACGTGGACTTCAGCGATGTAGGCTGGAATGACTGGATAGTGGCGCCTCCTGGTTACCAGGCCTATTACTGC
CACGGGGAATGCCCCTTTCCTCTGGCAGATCATCTGAATTCAACCAACCACGCCATTGTTCAGACACTGGTGAACTCTGTGAA
CAACAATATTCCCAAGGCCTGCTGTGTGCCAACAGAGCTCAGCGCCATCTCCATGCTCTACCTAGATGAACACGACAAGGTGG
TCCTAAAAAACTACCAGGAAATGGTAGTGGAGGGCTGCGGCTGCCGCTAA

>OniBMP16 (Nile tilapia; Accession number AERX01025009)
ATGTTCCCTCCTAACCTCCTGCTCCTCATGGTCCTGCTGCTACCTCATGGCTCGTCTGGTCGCCAGGGTGGAGACACCAGCGA
GATCGATCATGGCAGGGTGCACCCCACACCTTCCGCTTTGTCACCGTCACTCCGGGATCCCGGTCTGGCTCAGACCATCCAGA
GTCTCCTTTTGAGCCGACTGGGGCTGCAGTCTCAGCCTGACCCAAGGCCTGGAGTGCCGGTGCCACGGTACCTCCTGGATCTT
TACCACTTCCACCAGCAGCAGTACCATCTAGTAGAGGACCCTTCATTTAGTTTCCCCAGTCAGCACATCCAGGAGGCCAACAC
TATACGCAGCTTCCACCACAATGAGCCCCTTGGAGACAGTCTTGGTGCAGAGGATCATACCAGAGTGCACATTTCCTTCAATA
TCTCTTCCATCCCTCAAGATGAGAGTGTGCTCTCTGCTGAGCTCCGCCTCCGTCGCAGTGACGGAGCATCTCTGGGTCCTGGA
CCCCACAGACTAAACCTATACCTCTCTGAGCACCATGAGGACCCTGAGCCCACTCTGCTGGAAACAAGGTTACTCACCAATGG
CCTCCGTAGTTATAAAACAAGTGGTTTCTGGGAGGCTTTTAGTCTTAGCGCAGAACTCCTACATGAGGCCCATGCTGGGACTG
GCAGCCTAGGATTCCTCCTGGAGGTCAGACCTGAAAACACCACCTCACTTCCTGAGCAGAACCTCTCTTCTGCTGCAAGTGAG
GGGACGGAGAAATCAAAAGAGGCACACCTGAGGGTGTGCAGATCTGTGGGACAAGACGATCACAGCTGGGCCCAAGAGAGACC
CCTCTTGGTTACTTACAGTCATGATGGTCGTGGACAGCCTTTAGTCAAACATGGTAGAAGAAACCCTGGTAATGGCCAGAGGA
GAAAAGGAGCAAAGGCAAGGGTCAAGAATAAAAACAAGGGTCTCAATAGAAACCAAATCTGGGGAAGGGACAAAAGAATGGGG
TATACTGTGCCGAGCTGGGCGGATGACAAAGGAGGAATCAGTTGGAGCGAACGTGGCAGGGTGAAAAGAAACGGTGGTCGTTC
TGCAAAACTGAAACGCCTTTCCCGCAATAGATGCCGCCGCCATCCACTTTATGTAGATTTCAATGACGTGGGCTGGGAAAAGT
GGATCATTGCACCCAGTGGCTACGACGCCTTCTTCTGCCTGGGTGAGTGTCGCTTTCCTTTGGCTGACCACATGAACTCCTCA
AGCCACGCCATGGTGCAAACTCTGGTAAACTCTGTGAATGGAGCAGTGCCCCGGGCCTGCTGCGTCCCAACCTCTCTCAGTCC
CATCGCCCTGCTCTACCTGGACTCTCAAGACCGAGTTGTGCTGAAAAATTATCAGGACATGGTTGTGGAGGGCTGCGGCTGTC
GGTAG

>OlaBMP2 (Japanese medaka; Accession numbers BAAF04048581, DQ915174, AM340447, DK232580, DK251511, AM380364, DE095303, DH343454, DE100795, DE077111 and BAAE01085765)
ATGGTCGCCGTGGTCCGCTGTCTCATGGTACTGCTGCTCGCTCAGGTGTCGCTGGAAGGAGCTGCGGGACTCATCCCCGAGGT
CGGCCGAAGGAGGTACAGCGAGTCAGGGAGGCAGAGCGCAGGGCAGTCGGACAGCTTCCTCCACGAGTTCGAGCTTCGGCTTC
TCAACATGTTCGGGCTGGGGCGCAGGCCCACCCCCAGCAAGCAGGCTGTGGTGCCGCAGTACATGCTGGACCTTTACCACATG
CACTCTGCGCACGGGGAGCACAGCGCCAGGCGTCCCAGGAGCATGGGAAAACACGCGGAGAGAGCCGCCAGCAGGGCCAACAC
GATTAGAAGCTTTCACCATGAAGAAGCGATGGAGGCCCTGGCCAGTCTGAAGGGCAAAACAACCCAGCAGTTCTACTTTAACC
TCACCTCCGTTCCCGAGAAGGAGCTCATCACCTCTGCAGAGCTGCGCATTTACAGGGATCAGGTGAAAGGGGGCGCAAACAGC
AGCAACAGCAGCACGCCTGGTGCCGTTGGCGTCCACCGCATCAACATTTATGAGATGTTTGGACTTCCTGCTTCAAAAGGCGC
AGACCCTCTGGTGCGTCTGCTGGACACCCGGCTGGTGCAGGACTCTTTAAGCCGCTGGGAAAGCTTCGACGTCAGCCCCGCCG
TGTCCAAGTGGACCTCGGGGAAAGGCCACAACCACGGCTTCATGGTGGAGGTGCTTCACCCAGGAGAAGGGGACGGTGAGCAC
GGGCAAAGACAAAAAAGGCACGTCAGGGTGAGCCGGTCCTTGCACCAGGACCAGGACTCGTGGCCTCAGGTGCGGCCGTTGTT
GGTGACGTACGGTCACGACGGCCGCGGGGACTCTGTGCTCCATACAAGGGAAAAAAGGCAAGCATCCCTCCGCAAGCAGCGCA
GGAAAAACCATCACAAAGCCAGCTGCAAGAGGCACGTTTTGTATGTGGATTTCAGCGATGTGGGATGGAACGAGTGGATAGTG
GCACCCCCTGGCTACGATGCCTTTTATTGCCACGGGGAATGCCCCTTCCCCTTACCAGACCACCTCAATGCAACCAATCATGC
CGTAGTGCAGACTCTTGTCAACTCGGTCAACTCGAACATCCCCAAAGCCTGCTGTGTGCCCACTGAACTCAGCCCCATCTCCC
TGCTCTACCTGGATGAGTTCGAGAAAGTGACTCTGAAAAACTACCAGGACATGGTGGTGGACGGCTGCGGCTGCCGGTGA

>OlaBMP4 (Japanese medaka; Accession numbers DK004345, AM140778, AM139646, DC268248, AM299276, AM339195, DQ915175, DK197479, BJ014414, DC244878, DC276748, AM384609, DC253831 and AV669968)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTAATATCCCAAGTCCTGCTGGGAGAGAGCAACCATGCTAGTCTGATACC
CGAAGAAGGGAAAAAGAAAGTACCTGGCCTGCAGGGTCGCTCGACCGCTCAGAGCCATGAACTTTTACGGGACTTTGAGGCCA
CACTGTTGCACATGTTTGGCCTCAAAAAGCGCCCGCGTCCAAGCAAAGCTGCCGCCGTGCCTCGTTACCTGCTGGATCTCTAT
CGACTGCAGTCAGGGGAGGCTGAAGACGTAGAGGCACACGAAATTCCTTTTGAGTATCCCGAGAAGTCAGCCAGCCGAGCCAA
CACCGTGAGGGGCTTCCACCATGAAGAGCACATGGAGAGGATACACGAAGTGGATGATGCAGAATCCATCCGTCTTCGCTTCC
TGTTCAACCTTAGTAGCATTCCAGAGGACGAGCTGCTGTCTTCCGCAGAGCTGCGACTCTACCGTCAGCAGCTGGGCGAGGCC
AACGATGATAGCCCCCTAAACGACCAGGGCCTTCATCGGATAAACATATACGAGGTGCTTAAACCCCCGCGGCCCGGGCAGCT
TATCACGCAGCTTCTGGACACCCGGCTTGTGCGCCACAACGCTTCGCGTTGGGAGAGCTTTGACGTCAGCCCCGCTGTGCTAC

GTTGGACTCGGGAACGGCTCCCTAATTATGGACTGGCTGTGGAGATTCTCCACCTTAATCAGACTCCACACAACCAGCACCGA
CATGTCCGCATCAGCCGTTCCTTACACCAGGAGCCAGGAGAGGACTGGGACCAGGTACGCCCTCTTCTGGTTACCTTTGGCCA
CGATGGAAAGGGTCACTCATTGACACGCCGGACAAAACGCAGCCCCAAACCACGAGGGCGTAAACGAAACCGCAACTGCCGGC
GCCATACTCTGTATGTGGACTTTAGCGATGTAGGCTGGAATGACTGGATAGTGGCACCTCCTGGTTACCAGGCCTATTACTGC
CACGGAGAGTGCCCTTTTCCGCTGGCAGATCATCTGAACTCAACTAACCACGCCATCGTGCAGACACTGGTGAACTCTGTGAA
CAACAACATTCCCAAGGCCTGCTGCGTACCAACAGAGCTCAGCGCCATCTCAATGCTATACCTAGACGAATATGACAAGGTGG
TCCTAAAAAACTACCAGGAAATGGTAGTGGAAGGCTGCGGCTGCCGCTGA

>OlaBMP16 (Japanese medaka; Accession numbers BAAE01034973, BAAF04019993, BAAF04019995, BAAE01034974, BAAE01219338, AM318552, DK098052, DH423513)
ATGCTCCCTGCTAACATCCTGCTTCTCATGTTCTTGCTGCTACCTCAAGTCTCGTCTGGTCGCCAGGGTGAAGACACCAGCGA
GATTGACCATGGCAGGGTGCCGTCCACGTCTTCATCCTCGACACCCACGCCATCGATGGCGGATTCCAACCTTTCTCAAACCA
TTCAGAATCTTCTCTTGAGCCGACTGGGCCTGCAGTCTCAGCCTAACCCTCAGTCTGGAGTACCAGTCCCACAGTACCTCCTT
GATCTGTACCGCTTCCACCAAAAGCAGTACCATCTAGTGGAAGACCCATCATTTAGCTTCCCCAGCCAGCACATTCAACAGGC
CAACACAGTACGCAGTTTTCATCACCGCGAGTCCCATGAAGACAGTCACACAGAGAAGGATCATACTAGAGTACACATCTCAT
TTGACATCTCGTCCATCCCTCAAGAGGAGAGGGTGCTCTCAGCCGAGCTTCGCCTGCTGCGCAGCGACAAAGCTTCCCTGGGT
CCTGGATCACACAGACTGAACCTATACCTCTCAAAGCATCGCAGAGATCCTAAGCAGGCTTTGCTGGAGACAAGGTTACTCGC
CAACAACCTTCATAGCAATAAACCAAGTGGTTTCTGGGAAGCATTTACCATAAGTGCAGAGCTTCTGCAAAAAATCCAAGCTG
GTACTGGCAGCCTGGGCTTCCTTCTTGAAATCAAACCTGAAAATACTACCAACCCACTTCCTTACCAGATGTTTGCCTCTGCA
TCATATGAAGACAAGTCAGATGGACAAAAAGAGAGACACCTGAGAGTGTGCAGGTCTGTGGGCCAAGATGATCACAGCTGGGC
CAGGGAAAGACCCCTGCTGGTGACCTACAGTCATGATGGCCGAGGAGAGCCTTTGATTAAACATAGCAGGAAAACCTCTGGTA
ACAGTCAGAGGATTAGGGGGAAAAAGGGGTCAAAGGTGAGGGCCAGAAACAGCAGTAAAGGTCACGATACAGCTAAAGCCTGG
GGTAGGGCCAAAAAAAAGGGGTATTCTGTGCAGGACTGGGCGACTAAAACAGGAGGGGTCAGCTGGAGCGAACGAGGAAGGGT
AAAAAGAAATGGTGGGCACTCCGCAAAATTAAAGCGTATTTCGCGTGGCAGATGCCGCCGTCATCCTCTACATGTAGATTTTA
GAGATGTGGGTTGGCAAAAGTGGATCATTGCGCCCAGTGGCTATGACGCCTTCTTCTGCCTGGGAGAGTGCCGTTATCCCCTG
GTCGACCACATGAACGCCTCGAGCCATGCTGTGGTTCAAACCATGGTAAACTCTGTGAACGGAGCAGTGCCCCGTCCCTGCTG
CGTCCCCACCAGCCTCAGCCCCATTGCGCTGCTTTACATGGACTCTCAAGATCGAGTTGTGCTGAAAAATTACCAGGACATGG
TGGTGGAGGGATGTGGCTGCCTGTAG

>SsaBMP2 (Atlantic salmon; Accession numbers BT059611, CA061573, GE795400, GE795401, EG901180, EG901179 and CX353173)
ATGGTCGCCGTGTTTCGTTCTCTCATGGTACTGCTGCTGGCTCAGGTGTTGCTGGGAGGCGCTGCGGGACTAATCCCCGAGGT
TGGCCGGAGGAAGTACAGTGAGTCCGAGAAGCAGAGCCCGGAGCAGTCGGATAACTTTCTCAACGAGTTCGAGCTGCGGCTCC
TCAATATGTTCGGACTTAAGCGGAGGCCGAACCCGAGCAGACAGGCTGTGGTGCCTCAGTACATGGTGGACCTGTACCGTATG
CATTCGTTGAACGGAGACCACAGCACTAAACGGCCCCGGAGCATGGGGAGGCACGCAGAGAGAGCCGCCAGCAAGGCCAACAC
GATTAGAAGCTTTCACCATGAAGAGTCCATGGAGGCTCTGGCCAGTCTGAAGGGCAGGACAACCCAGCAGTTCTTCTTCAACC
TCACCTCCATCCCTGGAGAAGAGCTTATCACCTCTGCAGAGCTCCGGGTATACAGGGACCAGGTCCTGGGGGCCACAGCAGCT
CCCACCAGTAACACCAGTCACAACAGCAGTAACAGCTCCAGTACCAGTGCTGGTGGCTTCCATCGTATCAACGTGTATGAGGT
GTTTGGAGCCCCTGCGTCTCCCCGTGGGGAACCCCTGACACGCCTCTTGGACACGCGGCTGGTGCAGGACTCTCTGAGCCGCT
GGGAGAGCTTTGACGTCAGCCCCGCCGTGTCACAGTGGGCCTCCGGGGGGCGCCACAACCACGGCTTCCTGGTTGAGATGCTC
CACCCAGACACCCTGGGGGGCGAAGAGGGCCAAAGACAGAGACGACACGTCAGGGTAAGCCGCTCCCTCCACGGGGACCAGAA
CTCATGGCCCCAGGCTCGCCCCCTGCTGGTGACCTACGGTCATGACAGACAGGGGAATGCGGTGCTGCACAGGGACAAGAGAC
AGGCAGCAGGCCACAGGAAACAGAGGAAGAAGCACCAACACAAGGCCAACTGCCGCAGACATGCCCTCTACGTGGACTTCAGC
GACGTGGGCTGGAATGAGTGGATAGTTGCACCCCCTGGCTACCACGCCTTCTACTGCCACGGGGAGTGCCCCTTCCCATTGGC
CGACCACCTCAACTCCACCAACCACGCCATCGTTCAGACGCTGGTCAACTCTGTGAACTCCAACATCCCCCGGGCGTGCTGCG
TGCCCACAGAGCTTAGCCCCATATCACTGCTCTATCTGGATGAGTATGAGAAGGTCATCCTCAAAAACTACCAGGATATGGTG
GTGGAGGGATGTGGCTGCCGGTGA

>SsaBMP4a (Atlantic salmon; Accession numbers NM_001139844 DY702674, CX355133, GE769237, CA040511, DW541811, DW551593, CK893345, CB500013, GE769238, CA055499 and DW551594)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTATTATGCCAAGTCCTGCTGGGAGAGAGCAACCATGCTAGTCTGATACC
TGAGGAAGGGAAGAAGAAAGCCCCGGGCCTGCAGAGTCGGACGGCTGGTCAGAGCCATGAACTGCTGCGGGACTTTGAGGCCA
CGCTGCTGCACATGTTTGGGCTGCAGAGGCGACCGCGGCCCAGCCGCTCGGCCACCGTGCCACGCTACCTACTGGACCTCTAC
CGACTGCAGTCGGGTGAGGCCGAGGAGGCCGGTGCCCATGACGCCGCCTTTGAGTACCCCGAGAGGTCGGCTAGCCGCGCCAA
CACCGTGAGGGGCTTCCACCACGAAGAGCACATGGAGCGGGTTCATCTGGGCGATCTTCAGACGAAGAATGGCGAGGCCGCCC
CGCTGCGCTTCCTCTTCAATCTCAGCAGCATCCCAGAGGACGAGCTGCTCTCGTCGGCCGAGCTGCGTCTGTTCCGGCAGGAG
ATCGACGAGGCCATTTCGGGGAGCGGGGGGGACAAGGAGGGGCTCCACCGGATAAACGTGTACGAGGTTCTGAAGCCCCCGCG
GGCCGGGCAGCTAATCACGCGGCTCTTGGACACGCGACTGGTTCGCCATAACGCCTCGCGCTGGGAGAGCTTCGACGTGAGCC
CCGCCGTGCTGCGTTGGACACGGGAGCGCTTGCCCAACCTCGGGCTCGCTGTGGAGGTCCTGCACCTCAACCAGACACCGCGC
CACCAGGGCCGCCACGTCCGCATCAGCCGCTCACTGCACCAGGAACCTGGGGAGGACTGGGAGCAGCTACGCCCCCTGCTGGT
CACCTTCGGCCACGACGGCAAGGGCCACCCGTTGACGCGCCGGGCCAAGCGCAGCCCAAAGCAGCGGGGGCGTAAGCGCAACC
GCAACTGCCGCCGTCACGCGCTGTATGTGGACTTCAGTGATGTGGGCTGGAATGACTGGATAGTGGCGCCCCCAGGGTACCAG
GCATACTACTGCCATGGGGAGTGCCCCTTCCCTCTGGCAGACCACCTGAACTCTACCAACCACGCCATCGTGCAGACGTTGGT
GAACTCGGTGAACACCAACATTCCCAAGGCCTGCTGCGTGCCCACGGAGCTCAGTGCCATCTCCATGCTCTACCTGGATGAAC
ACGACAAGGTAGTCCTAAAAAACTACCAGGAAATGGTAGTGGAAGGGTGCGGTTGCCGCTAA

>SsaBMP4b (Atlantic salmon; Accession numbers CX354839, DW573815, DW573814, CA056395, FJ195610)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTATTATGCCAAGTCCTGCTGGGAGAGAGCAACCATGCTAGTCTGATACC
TGAGGAAGGGAAGAAGAAAGCACCGGGCCTGCAGAGCCGGACTGCTGGTCAGAGCCATGAACTGCTGCGGGACTTTGAGGCCA
CGCTGCTACACATGTTTGGACTGCAGAGGCGACCGCGGCCCAGTCGCTCGGCCACTGTGCCACGGTACCTGCTGGACCTCTAC
CGGCTGCAGTCGGGCGAGGCTGAGGAGGCCGGCACCCACGACACCGCCTTTGAGTACCCCGAGAGGTCGGCTAGCCGCGCCAA
CACCGTGAGGGGCTTCCACCACGAAGAGCACATGGAGCCGGTTCATCCGGGCGAGTCTCAGACGAAGAACGGAGAGGCCATCC
CACTGCGCTTCCTCTTCAATCTCAGCAACATCCCAGAGGACGAGCTGCTCTCGTCGGCCGAGCTGCGTCTGTTCCGGCAGCAG
ATCGACGAGGCCATTACGGAAAGCGAGGGGGACAAGGATCAGCTCCACCGGATAAACGTGTACGAGGTGCTGAAGCCCCCGCG
GGCCGGGCAGCTAATCACGCGGCTTTTGGACACACGGCTGGTGCACCATAACGCCTCGCGCTGGGAGAGCTTTGACGTGAGCC
CCGCCGTGCTGCGCTGGACACGGGAGCGCCTGCCCAACTACGGGCTGGCCGTGGAGGTCCAGCACCTCAACCAGACACCACGC
CACCAGGGCCGCCACCTGCGCATCAGCCGCTCGCTGCACCAGGAGCCCGGGGAGGACTGGGAGCAGCTACGCCCCCTGCTGGT
CACCTTCGGCCATGACGGCAAGGGCCACCCGTTGACGCGTCGGACCAAGCGCAGCCCAAAGCAGCGGGGGCGTAAGCGCAACC
GCAACTGTCGCCGTCATGCTCTCTACGTGGACTTCAGCGATGTAGGCTGGAATGACTGGATAGTGGCGCCCCCAGGGTACCAG
GCATACTACTGCCATGGGGAATGTCCCTTCCCCCTGGCAGACCACCTGAACTCTACCAACCACGCCATTGTTCAGACGTTGGT
GAACTCGGTGAACACCAACATTCCCAAGGCCTGCTGCGTGCCCACGGAGCTCAGTGCCATCTCCATGCTCTACCTGGACGAAC
ACGACAAGGTGGTCCTAAAAAACTACCAGGAAATGGTTGTGGAAGGGTGTGGCTGCCGCTAA

>SsaBMP16 (Atlantic salmon; Accession numbers AGKD01064384, CK879241, CK879040)
ATGTTCCCTGCTAAACTCTTGGTCCTGATGGTCCTACTGCTACCTCAAGCTTCGTCTGGTCGCCAGGGCTTTGCCACCGAGCC
CGATGGCAGTGTGCCGTCGTTGTCACCGTTATCACCCACCCCCCTGGAGCCCAGCCTGGCCCAGACAATCCAGAACCTGTTGC
TGAGCCGCCTGGGCCTGCAGTCCCACCCCAACCCCCGGCCAGGTGCGCCCATCCCCCAGTACATACTGGACCTCTATCGCTTC
CACTCCCAGCAGTACCACCTAGTCCAGGATCCACACTTCAGCTACCCCAGCCAGCATGTCCAGGAGGCCAACACCGTACGCAG
CTTCCACCACACAGAGTCTACCAGCCCCAAACTCCCCCTTGAAAAGAGGCTGACCACCGTGAGCAACAACAAGGTCCCCATCT
CCTTCAACGTGTCCTCCATCCCCCAGGATGAGTGTGTGGTTTCTGCCGAGCTGCGTTTCTTCCGAGGTGGTGGGGCTAGCCTA
GGCCCCGGCGCCCACAAGGTCAGCCTGTTCCTCGATGGAGGCACTGGGGACTCTGAGTCTGAGCCCACTCTGCTGGAGTCACG
GCTCCTCACAGGGGCCCCCAGCACTAAGCCAGCTGGCTCCTGGGAGGCCTTCAGCCTCAGCACTGAACTCTTCGTAGGGGCCC
ACGCTCAGACCGGCAGCCTGGCCTTCCTCCTTGAGGTGACCCCTCTGGGTAACACCACCCTCCTCACCCCCCCTGACCATGAT
GCTCTACCACTCTCCACTGGGGAGGAAGGGCGCAGAGAGGGACACCTGAGGGTGCGCAGGTCTCTGGGACAGGACGAGCACAG
CTGGGCACGGGAGAGACCCCTGCTGGTTACTTACAGCCATGACGGGCACGGGGAGCCACTAGCCGCCCATGGCCGGAGAACCT
CTGACACCGCCCAGAGGATGAGGGGGAGGAAGCGGGCCAGAGAAAGGGGCAGGAGCAACAGCAGGGACCGGGAAAGGGACAGA
GACAAGGACAGGGACTGGAGCGCCAGCCCTGGCTGGGGCGGCAGCTGGAACGAGGGTGGAAGGGTGAAGCGCAATGGTGGCCG
GGCAGCGAAACTGAATCGTCTGTCCCGCGCCCGCTGCCGCCGCCACCCGCTCTACGTGGACTTCAAAGACGTGGGCTGGAACA
AGTGGATCGTGGCGCCTAGTGGCTACGACGCCTTCTTCTGCCTGGGGGAGTGCCGCTTCCCACTCACCCACCACATGAACTCA
TCCAGCCACGCCATGGTGCAGACGCTGGTGAACTCGGTGAACGGAGCTGTGCCGCGGGTCTGCTGCGTACCCACCGCTCTCAG
CCCCATCGCCATGCTCTACCTGGACCCGCAGGACCGCGTGGTGCTCAAAAACTACCAGGACATGGTGGTGGAGGGCTGTGGAT
GCCGGTAA

>TniBMP2 (Spotted green pufferfish; Accession number AC141355)
ATGGTCGCCGTGGTCCGCTCTCTCATGGTACTGCTGCTCGCTCAGGTGTTGTTGGAAGGTGCTACTGGACTTATCCCTGAGGT
CGGCCGGAGGAGATACAGCGAATCTGGGAAGCAGAGCCCACAGCAGTCAGAGGGCTTCCTCAACGAGTTTGAGCTTCGGCTTC
TCAATATGTTCGGACTGAAGCGGAGGCCGACCCCGAGCAAGCAGGCTGTGGTGCCGCAGTACATGGTGGACCTTTACCGCATG
CACTCGGCGAACGGAGATCACAGCGCTAAACGACCCAAGAGCATGGGGAGACATGCAGAGAGAGCCGCCAGCAAGGCCAACAC
GATTAGAAGCTTTCACCATGAAGAGTCGATGGAGGCACTGGCCAGGCTCAAAGGCAAAACAACTCAACAGTTCTATTTTAATC
TCACATCCATACCTAAAGAAGAGCTCATCACCTCTGCAGAGCTCCGTATCTACAGGGATCAAGTCATGGGAGCTGCATCCACC
AACAACAGCTCTACCAACAGCAGCACCAGTGATAAGGCTCATGCTGGTGGCTTCCATCGCATCAACATCTACGAGATATTCGG
AGTTCCTCAAGGCAGGGAACCTCTGGCACGTCTGTTGGACACTAGGCTTGTTCAGGACTCTCTGACCCGGTGGGAGAGCTTCG
ATGTCAGCTCTGCTGTATTTCAGTGGACCTCTGGGAAAGGCCACAATCACGGCTTTATGATAGAAGTGCTTCATCCAGAGGAG
GGCGAGGTAGAACAGGAGCGAGCCGACAAACATAGCAGACATGTCAGAGTGAGCCGGTCCTTGCATCAGGACCGGGACTCATG
GCCTCAGGCCCGGCCTTTGCTAGTAACGTACGGCCACGACGGCCGCGGGGACTCGGTGCTCCACACCAGGGAAAAACGTCAGG
CAGCACTCCGCAAACAGCGCAGGAAACACCAACACAAGGCCAGCTGCAAAAGGCATGCCCTGTATGTAGACTTTAGTGACGTT
GGGTGGAACGAGTGGATTGTGGCACCCCCTGGCTACCATGCCTTTTATTGCCAGGGAGAATGTCCATTCCCCCTGGCAGACCA
CCTCAACTCTACCAATCATGCCATTGTGCAAACACTGGTCAATTCAGTCAACTCGAATATTCCCCGAGCTTGTTGCGTGCCCA
CTGACCTCAGCCCCATCTCCTTACTCTACCTGGATGAATATGAAAAAGTCATCCTTAAAAACTACCAAGACATGGTGGTGGAG
GGATGTGGCTGCCGGTGA

>TniBMP4 (Spotted green pufferfish; Accession number CAAE01015019, AC141355 and AL292165)
ATGATTCCTGGTAATCGAATGCTGATGGTTATTTTAATATGCCAAGTCCTGCTGGGAGAGAACAACCATGCAAGTCTGATACC
TGAAGAAGGGAAAAAGAAAGTACTGGGCCTGCAGGGTCGTTCGGCCACTCAGAGCCATGAACTTCTACGGGACTTCGAGGCCA
CGCTGCTGCACATGTTCGGCCTCAAGAGGCGGCCGAGGCCCAGCCGCTCGACCGCCGTCCCCCAATACCTGCTGGACCTCTAC
CGGCTGCAGTCGGGCGAGGCTGAGGAGGCTGGAGGGCAGGATATCGCATTTGAGTACCCGGAGAGGGCAGCCAGCCGGGCCAA
CACTGTGAGGGGCTTCCATCATGAAGAGCACATGGAGATGATTTACGCGCCGGATGATGGCGAATTGACGCCCCTTCGATTCT
TGTTCAACCTTAGCAGCATTCCAGAGGACGAGCTGCTGTCCTCCGCCGAACTGCGGCTCTACCGCCATCAGGTTAATGAGGAC
AACGTCACGCTCTCGGCAGACCACGGGCTCCATCGGATAAACGTGTATGAGGTGCTCAAGCCCCCTAGGCAAGGGCAGCTAAT
CACGCAGCTCCTGGATACGCGGCTCGTCCGCCACAACGCGTCACGCTGGGAGAGCTTCGATGTGAGCCCTGCTGTGCTACGCT
GGACTCGTGAAGGCCAGCCCAATTATGGATTGGCCATAGAGGTTCTGCACCTCAACCAAACACCGCGCCACCAGGGCCGACAT
GTCCGAGTCAGCCGGTCTTTACACCAAGAGCCTGGGGAGGACTGGGAGCAGCTACGCCCGCTCCTGGTCACCTTTGGTCACGA

CGGGAAAGGTCACCCGCTCACCCGGCGTTCCAAGCGGAGCCCCAAGCAGCGAGGACGTAAACGCAACCGCAGCTGCCGGCGCC
ACTCGCTGTACGTGGACTTTAGCGACGTGGGCTGGAATGACTGGATAGTGGCGCCCCCTGGTTACCAGGCCTACTACTGCCAT
GGGGATTGTCCGTTCCCCCTGGCGGATCATCTGAACTCGACCAACCACGCCATTGTTCAGACACTGGTGAATTCCGTGAACAG
CAACATTCCCAAGGCCTGCTGCGTGCCCACCGAGCTCAGTGCCATTTCCATGCTCTACCTAGATGAACACGACAAGGTGGTCC
TAAAAAACTACCAGGAAATGGTGGTGGAGGGCTGCGGCTGTCGCTAG

>TniBMP16 (Spotted green pufferfish; Accession number CAAE01014537.1 and AL336493)
ATGCTCCCTGCTAACCTCCTGCTCCTCATGGTCCTGCTGCTACCTCAAGTCTCGTCTGGTCGCCAGGGTGGAGACGCCAGCCA
GACTGACCATGGCAGAGTGCCCCCCGTGCCCTCGTCCCTTCTGAGCCCACCAGCCGCTCCCAGCCTGAAGCCCAGCCTGGCCG
AGACTATCCAGAGTCTCCTCCTGAGTCGCCTGGGCCTGCAGTCGCAGCCGGACCCTCGGCCTGGAGTCCCGGTGCCTCGGTAC
CTCCTGGATCTGTACAGATTCCATCAGCAGCAGTACCACCTAGTGGAGGACCCTTCATTCAGCTTCCCCAGGCAGCACATCCA
GCAGGCTAACACCGTTCGCAGCTTTCACCACAATGAACTTTCTGATGACAATTCCCTACCTGATAATCCTCAGAGGGTGCACA
TCTCCTTCAATATCTCCTCCATCCCTGAAGACGAGAGGGTGCTCTCCGCCGAGCTTCGACTCCTCCGAAGCAGCAGAGCCTTT
TGGGGCCCTGGGGTCCACAGATTGAACATTTACCTTGTTGAGAGTCATGGAGACCCTGAGCCCACCCTGCTGGAGACACGATT
ATTCACCACTGCTAATAATACTCATGAATCAAGTAGTCTCTGGGAGGCCTTTCCTCTAAGCTCAGACATTCTCCTTGAAGTTC
TTGGTGGGACTGGCACCCTGGGCTTCCTGCTGGAGGTCAGACCTGAGAACAGAACCTCCTCAGGCGTCCAGCAGAACGTCTAT
GCTGCAGCAGGTAAAGAGGATGCTGAGAAAGAGGGACACCTGAGGGTCCGTAGGTCCGTGGGTCAGGATGAGTACAGTTGGAC
CCAGGAGAGACCTCTCTTGGTGACCTATAGCCATGATGGGCGCGGAGAACCTTTAGCCAACCATGGCAGGAGGACCCCCGCTG
GCGCCCAGAGGATCAGAGGGAGAAAAAAGACCAAAGACAGAGAAGCCAGGAGCAGCAAGAAAAGCCGCGCTAGCGATCCAAAC
AGTAGGAAGGACAAAACAGGATATTCTGTGCCAGGATGGGGTCGTGATAAAGGAGGGGTGGTTTGGAAGGAGAACGGAAGGGT
GAAAAGAAACAACGGCCGTGCCGCAAAACTAAAACGCCTTTCCCGTAACAGATGTCGCCGGCATCCCCTGTACGTAGATTTCA
ATGACGTAGGCTGGCACAAGTGGATCATTGCCCCCAGCGGCTACGATGCCTTCTTCTGCCTGGGAGAGTGCCGCTTTCCCCTG
GCTGACCACATGAACTCCTCCAGCCACGCCATGGTACAAACTCTGGTCAACTCTGTGAACAGAGCCGTTCCGCGGGCTTGCTG
CGTCCCCACCTCCCTCAGCCCCATCGCCCTGCTCTACCTGGACCCGCAAGACAGAGTCGTGCTGAAGAACTACCAGGACATGG
TGGTGGAGGGATGTGGCTGCCGGTAG

>XmaBMP2 (Southern platyfish; Acession number AGAJ01019336)
ATGGTCGCCGTGGTCCGCTCTCTCATGGTACTGCTGCTCGCTCAGGTGTTGCTGGAAGGTGCTGCGGGACTCATCCCCGAGGT
CGGCCGGAGGAAGTACAGCGAATCTGGGAAGCAGAGCCCGGAGCGGTCGGAGAGCTTCCTCAACGAGTTCGAGCTTCGGCTTC
TCAACATGTTCGGGTTGGGGCGCAGACCCACCCCCAGCAAGCACGCCGTGGTGCCACAATACATGGTGGACCTTTACCACATG
CACTCTGCAAACGGAGACCACAGCACTAAACGACCCAAGAGCATGGGCAAGCACGCAGAGAGGGCCGCCAGCAAGGCCAACAC
GATTAGAAGCTTTCACCATGAAGAATCAATGGAGGCACTGGCCAGCCTGAAAAGCAAAACAACCCAGCAATTCTACTTCAACC
TCACGTCTATCCCTAAAGAGGAGCTCATAACCTCTGCAGAGCTCCGCATTTACAGGGATCAGGTCATCGGAGCGGCACCTCCT
AAGAACAGCAGCTCTACTAGTGGTGCTCTTGCTGGAGGCCTCCATCGTATCAACATCTACGAGATCTTCAGAGTACCAATAGG
CACAGAACCTCTGGTCCGTCTGCTAGACACCAGGCTGGTTCAGGACTCTTTGAGCCGCTGGGAGAGTTTTGACGTGAGTCCTG
CTGTATCTGAATGGACTTCGGGGAATGGCCCGAATCATGGCTTCATGGTGGAGGTGCTTCATCCAGAGCAGACAGAGAAGGCT
GCAGAGCATTCCCAGAGAAGAAACAGGCACGTCAGGGTGAGCCGGTCCCTGCACCAGGACCAGGACTCATGGCCTCAGGCTCG
GCCCCTGCTGGTAACATATGGTCACGACGGTCGTGGGGACTCAGTACTCCATACACGGGGGAAAAGGCAGGCAGCAGTCCGGA
AACAACGCAGGAAGCATCAGCAAAAGGCAAGCTGCAAAAGGCATACCTTGTACGTAGATTTCAGCGACGTGCAATGGAACGAG
TGGATAGTGGCACCCCCAGGCTACCATGCCTACTACTGCCATGGGGAATGCCCCTTTCCATTACCAGACCACCTTAATGCAAC
TAATCATGCCATTGTGCAGACACTTGTCAACTCAGTAAACTCAAACATCCCCAGAGCCTGTTGTGTGCCCACCGAACTCAGCC
CCATTTCGTTGCTCTACTTGGACGAATATGAGAAGGTGATTCTGAAAAACTACCAGGATATGGTGGTGGAGGGCTGTGGCTGC
AGATGA

>XmaBMP4 (Southern platyfish; Acession number AGAJ01041398)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTATTATGCCAAGTCCTGCTGGGAGAGAGCAACCATGCCAGTCTGATACC
TGAAGAAGGGAAGAAGAAAGTGCCTGGCCTGCAGGGTCGTTCGGCTGCTCAGAGCCATGAACTGTTGCGGGACTTTGAGGCCA
CGCTGCTGCACATGTTCGGCCTCAGGAGGCGGCCGCGGCCCAGTCGCTCCGTCACCGTGCCCCGCTACCTGTTGGACCTCTAT
CGGCTGCAGTCGGGGGAGGCCGAAGAGGCCGGAGCGTACGACATCGACTTTGAGTATCCCGAGAGGTCGGCCAGCCGGGCCAA
CACTGTGAGGGCCTTCCACCATGAAGAGCACATGGAGAGGATCCATGAGCTGGGTGATGGAGAATCCATGCCTCTCCGCTTCC
TGTTCAACCTCAGCAGCATTCCAGAGGACGAGCTGCTTTCCTCTGCAGAGCTAAGACTCTACCGTAGTCAGATTGATGAGGCC
ACCGCTGACGACCAGAGTCTGCACCGCATAAACATCTACGAGGTCCTGAAGCCCCCGCGGCCCGGGCAGCTCATCACCCAGCT
TCTGGACACACGCCTCGTTCGGCACAACGCGTCACGGTGGGAGTCCTTTGATGTCAGCCCTGCCATCTTGCGCTGGAGTCGCG
AGGGCCTCCCTAATTATGGTCTGGCCGTGGAGGTTCTGCACCTCAACCAGACCCCACACCACCAGGGCCGACACGTCCGCATC
AGCCGCTCGGTGCACCAAGAGCCCGGGGAGGACTGGGAACAGTTACGCCCTCTTTTGGTTACTTTTGGTCATGATGGCAAGGG
TCACCCACTGACCCGGCGGACCAAGCGCAGCCCCAAACAACGGGGCCGCAAAGGCAGGCGCAACTGCCGGCGCCATGCACTGT
ACGTAGACTTCAGTGATGTAAACTGGAATGACTGGATTGTGGCACCGCCTGGCTATCAGGCATATTACTGCCACGGGGAGTGC
CCCTTTCCCCTGGCAGATCATTTAAACTCCACCAACCATGCCATTGTTCAGACACTGGTGAACTCTGTGAATAACAACATTCC
CAAGGCCTGCTGCGTACCAACAGAGCTCAGCGCCATCTCCATGCTCTACGTGGATGAACATGACAAGGTGGTCCTGAGAAACT
ACCAGGAAATGGTAGTGGAGGGCTGTGGCTGCCGCTAG

>XmaBMP16 (Southern platyfish; Acession number AGAJ01016356)
ATGCTCCCTGCTAACCTCCTGCTCCTCATGGTCTTGCTGCTACCTCAAGCCTCGTCTGGCCGCCAGGGTGGAGAATCCAGCGA
GACTGACAGTGGCAGGGTGCCCCCGACCCCTTCGTCCCCGTCGCCCGCCGCAGCGCTGCCCGATCCCGGTCTGGCTCAGACCA
TCCAGTCTCTCCTCCTGAGCCGGCTGGGCCTGCAGTCTCAGCCGAACCCCAGGCCCGACGTGCCGGTGCCGCAGTACCTCCTC
GACCTTTATCGCTTCCACCAGCAGCAGTACCATCTGCTGGAAGACCCGACGTTCAGCTTCCCCAGCCACCACATTCAGCAGGC
CAACACTATTCGCAGCTTTCACCACAGTGATTCCTCTGGAGACAACACCAAAATAAGAGATCAGACAAGAGTGCACATTTCCT
TCAACATCTCCTCCATACCTCAAGATGAAACGGTTCTCTCTGCTGAGCTTCGCCTGCTCCGCAGCCAGGCAGTGTCCCTGGGC

CCCGGGCCTCACAGATTAAACCTGTACCTCTCTAAACACTATGAGGATCCTGAGCCCACGTTGCTTGAGACAAGACTACTCAC
CAGTGGCCTCCATGGCATTAAAGCCGGTAGTTTCTGGGAGGCATTTAGCCTAAATGCGGACCTCCTCCGTAAAGCCCATGGTG
GGACTGGCAGCTTGGGCTTTCTCCTGGAGGTCAGACCTGAAAACAGCACCACCTCACCTGACCAGAGCAACTCTTCTGCTGCA
GGTGATAATCAGTCAGAGCAACATCAGGAGAGACACCTGAGGGTGTCCAGGTCGGTGGGACAGGACGATCACAGCTGGGCCCA
GGAGAGACCCCTCCTGGTGACTTACAGTCATGACGGCCATGGAGAGCCTTTAGTCAAACACGGCAGGAGAACATCTGGTAACA
GCAAGTGGATTAAGGCTAGAAAAGGGTCAAACATGAAGACCAAGAGAAGCGGCAAAGGCCACAATAGAGACCAAGGCTGGGGG
AAGACCCGAAAAATGGGTTACTCCATGAAGAGCTGGGGGGACGACAAAGGGGAGGTAAGCTGGCGTGATCGTGGAAGAGTGAA
AAGAAACGGCAGTCCCGCTAAGCTGAAGCGGCTTTCCCGTGGCAGGTGCCGGCGTCATCCTCTACATGTAGATTTCAGAGATG
TGGGCTGGCACAAATGGATCATCGCGCCGAGCGGCTATGACGCCTTCTTCTGCCTGGGAGAGTGCCGTTACCCTCTGGCCGAC
CACATGAACGCCTCCAGCCATGCCGTGGTCCAAACCATGGTAAACTCTGTGAACGGAGCAGTGCCCCGGGCCTGTTGCGTCCC
GACCTCCCTCAGCCCCATTGCCCTGCTTTACCTGGACAATCACGACAGAGTCGTGTTAAAGAATTACCAGGACATGGTGGTGG
AAAGCTGCGGCTGCTGGTAG

>LchBMP2 (Coelacanth; Accession number JH126648.1)
ATGGTTGCCGGAGTCCGCTCTCTCATGGTTCTGCTGCTCTGCCAGGTCCTTCTTGGAGGTTCGGCCGGGCTCATCCCGGAGGT
GGGCCGGCGCAAGTTCACCGAGTCTGGCCGCTCCAGCCCCCAGCAAAGCGAAGAGATCCTCAACGAGTTTGAGCTCCGGCTGC
TCAACATGTTCGGACTGAAGCGGCGGCCGAACCCTGGCAAGAGCGCCGTGATCCCGCAGTACATGCGGGACCTGTACCGCCTG
TATTCCCGGAGTGGTGAGGACAGTCTGGACCAGCACGGATCATCCCTAAATCCACAGCTGGAAAGATCTGCTAGTCGGGCCAA
CACTATCAGGAGTTTCCACCATGAAGAATCTTTGGAGGAACTACCAGAAACTGGTACAAAAACAGCACGGCATTTCTTCTTTA
ATTTATCTTCCATCCCTGGGGAGGAAATGCCTGCCTCAGCTGAACTACGGATTTTTCGTGAAGAGGTCCAACCAACCTTGGCA
AATAATTCCAGTGGCCACCACCGAATTAATATTTATGAAATTATAATGCCAGCCTCATCCACACTGAGGGATCCCATTACAAG
ACTGCTAGACACCAGATTGGTGCACCAAAGCCAGAGCAAATGGGAGAGTTTTGATGTGACCTCTGCTGTCATGAGGTGGCTAA
TGTACCAGGAACCCAATCATGGATTTGTGGTAGAAGTGGTCCACGTGGACAAAGAGGAAAGTGCCTCCAAAAGACATGTTAGG
ATTAGCAGGTCTTTGCATGAAGATGAAGAAAGTTGGCCTCAGATGAGGCCATTATTAGTAACTTTCAACCACGATGGTAAGGG
ACACCCTCTCCATAAAAGGGAGAAGCGCCAGGCAAGACCGAAACACAGGAGACGGCTGAAATCGAGCTGTAGGAGGCATGCTT
TGTATGTGGACTTCAGTGATGTAGGGTGGAATGACTGGATAGTTGCTCCTCCAGGGTATAATGCCTATTATTGCCATGGGGAA
TGCCCATTTCCACTAACGGATCACCTGAACTCAACCAACCATGCCATTGTTCAGACTTTGGTCAATTCTGTCAATTCAAATAT
TCCTAAAGCCTGCTGTGTGCCCACAGAACTCAGTCCCATCTCCATGCTCTATCTTGATGAGTATGACAAGGTTGTATTAAGGA
ATTATCAAGATATGGTTGTGGAGGGTTGTGGTTGTCGCTGA

>LchBMP4 (Coelacanth; Accession number JH126564.1)
ATGATTCCTGGTAACCGAATGCTGATGATAATTTTATTATGCCAAGTCCTATTAGGGGGTAATACCCATACTAGTTTAATACC
TGAGAAGGGGAGAAAAAAAGTTTCAGAGATTCAGGGCCAGGCAGGAGGACGTCGCTCTGGGCAAAGCCATGAACTCTTGCGGG
AATTTGAGGCAAGTTTGCTGCAGATGTTCGGATTGCAGAGACGGCCCCAGCCCAGTAAATCGGTCGTCGTCCCCCGTTACATG
CTGGATCTGTACCGATTCCAGTCCGGGGATGAAAACGTAAAGGATATCAACTTTAAATACCCCGAGAGATCCACGAGCCGGGC
CAACACAGTCAGGAGCTTCCATCACGAAGAACGCTTGGAAAAGGATTCGGCCAACAGCGACGACTCCTTAAGCCACTTTATAT
TTAACATCAGCAGTCTTCCTGAAAACGAAGTGTTTTCCTCTGCAGAGCTGAGACTTTATAGGGAACAGATAGAACAAGGTCCC
AGTTGGGAAGAACATTTTCACAGAATCAACATTTATGAGGTCATGAGACCACCTAGCAGAAATGGACAACTTATTACCAGACT
ACTGGACACTAGGCTGGTGCATCAGAATGCCACACAATGGGAAAGTTTTGATGTCAGTCCAGCTGTTGTGAGATGGATCAGGG
ACAAACAACCATATCGTGGGTTTGCTGTTGAGGTCATCCACCTTAACCAAACAAAAACCCATCAGGGGAAACATGTCAGGATT
AGCCGATCTTTACCTCAAGGGGATACAGATTGGTCCCAGTTGAGACCACTTTTGGTCACTTTTACTCATGATGGCAAAGGACA
TGCACTGACCAGGAGGGTGAAAAGGAGCGCTAAACATCAAAGATCCAGGAAGAAAAACAAAAACTGTAGGAGACATGCTCTTT
ATGTGGATTTTAGCGATGTGGGCTGGAATGACTGGATTGTGGCACCACCTGGTTACGACGCATATTACTGCCATGGGGACTGC
CCTTTCCCCTTGGCTGATCATCTGAACTCTACCAACCATGCTGTCGTTCAAACCTTGGTCAACTCAGTCAACTCCAATATACC
CAAAGCTTGCTGTGTGCCCACTGAGCTCAGCGAAATCTCAATGCTTTACGTAGACGAACATGACAAGGTTGTCCTCAAAAACT
ATAAGGATATGGTAGTACAAGGGTGTGGATGCCGTTGA

>LchBMP16 (Coelacanth; Accession number ENSLACT00000009133)
ATGGTCCCTGCTAACCTGTTATTCCTGATGCTTCTCTTGCCACCCTATACCCTTTTGGGTTCTGAGGACATTTTGCACCAGCA
AGAAGTTCTGCCCAGCTCCCACCCTCCTGTCAAGAAGACCACCAACCGCACCCTTGACCTCAATGTGATCCGGTCTGTTCAGA
CCCTCCTGCTGACCCGCTTGGGGTTCCGGAGCTGCCCCCAGCCAAAACCAGGCCTCCTTATCCCTCAGTATATGTGGGACCTG
TACCGGTACCATTCTGGAGACCTTCCCGCAATCCAAGACTCAGGATTCCAGATCCCAGAAGAACACACAGGGAAAGCAAACAC
CATCAGGGCCTTTCATCATCTTGAAGGACCAGAGAATGTTGGTGAGCCACAGGAAGACCAGATTTTTAATTTTCAGTTCAATA
TTTCCTCTATTCCCGGGGATGAGAAGGTGACTTCAGCCGAACTCCGTTTGCACAAAGAGAAGCCCAAAAGCTATGGAAATTCC
CTGCTACAAAGGGTTAACCTCTACCATATCACAGATCTCCAATCCAGGAGACGCCAACTTCTAGAGACCAGGGTGTTGACTCC
TAACTGGGCCCAGTGGGAGAGTTTTGATGTTACTCCTGCTGTGCTCATGTGGCTTCAAGATAGCATGGCTAATATGGAGTTTA
TGATTGAGGTGGTCCATCAAAATGGAACCCAAGTAGATTGGAACCAGAGTGGATATCTGAGGGTGAGGAGGTCACTTGGTGAA
GGTGAGGAAACCTGGGTTCATCAGAGGCCATTATTAGTCACCTATAGTCATGACGGTCGTGGCCAGCCTCTGGCACGAAGGAC
CAAAATTAAAAGGAATGGGAAGATGCAGAAGTCAAAGAAGCGAGCAAAGCCTAGATGCAAGAGGCATCCATTATTTGTTGACT
TCAAAGATGTTGGGTGGAACAAATGGATTGTTGCCCCAAGTGGTTATCGTGCCTTTTACTGCCATGGTGAATGTCGTTTTCCC
TTAGTGGATCACATGAACTCTTCCAGCCATGCCATGGTGCAGACATTGGTAAACTCAGTAAATTCTAAAGTGCCCAGGGCATG
CTGTGTGCCCACGGACCTAAGCCCAATCGCCATGCTCTATCTGGACCAACATGAGAGGGTGGTCCTCAAGAACTACAATGACA
TGGTAGTGGAAGGCTGTGGCTGCCGATGA

>AmexBMP2 (Axolotl; Accession number EU339232)
ATGGTGGCCGGGATGCGCTCGTGTCTGGTGCTGCTCCTGTACCAAGTGCTGCTGGGCGGCTCGGTGGGGCTCATCCCGGAGGT
GGGCCGGCGCCGCTTCGCCGAGTCGAGCCGGGGCCCTGCCAGCGCGGGCCCCCAGCAGCAGCAGCAGCAGCACTCCGAGGACG
TGCTGCGCGAGTTCGAGCTCCGCCTGCTGCACATGTTCGGCCTGAAGCGGCGGCCCAGCCCGAGCGCCGGCGCCCCCATCCCG

CCCTACATGCGCGACCTGTACCAGGGCCAGGGCCTGGACCTCCGGCAGGAGCGCTCCGCCAGCCGCGCCAACACCGTCCGCAG
CTTCCACCATGAAGAAGTTTTGGAAGGCCTGCCTGAAGCAAGCGGAAAACTGCTGCAGCGTTTCTTCTTCAACTTGACTTCTA
TCCCTAACGAGGAGTTTATCACCTCGGCTGAACTCCGGATTTTTCGAGAACTGGTCACGGAGACCTCTGGCGGGAACAGCAGC
ACGCACCGCATTAATATTTATGAAATCCTCAAGCCCCCCGCCGCGGGGCCCCAGGAGCTGGTCACGCGGCTGCTGGACACCCG
GCTGGTGCAGCCCAACGCCAGCAGCTGGGAGAGCTTCGACGTGACCCCGGCCGTGCTGCGGTGGCAGGTGCACGGGCAGCCCA
ACCACGGCTTCGTGGTGGAGGTGGCCCACCTAGACAGTGAGCGCGGCCTGGCCCAGCGGCACGTCCGCCTCCGCCGCTCCCCG
CAGCAGGACGCGCACAGCTGGGCCCAGCTCCGGCCCCTCTTAGTAACGTTTGGCCACGACGGCAAGGGCCACCCGCTGCACCG
GCGCGAAAAGCGCCAGGCGCGCCAGCGGCAGCGCAAGCGCCTCAAGTCCAACTGCCGGAGGCACCCCCTCTACGTGGACTTCA
GCGACGTGGGCTGGAATGACTGGATCGTGGCGCCCCCGGGATACCACGCCTTCTATTGCCACGGGGACTGCCCCTTCCCTCTG
GCGGACCACATGAACTCCACGAACCATGCCATCGTGCAGACTTTGGTCAACTCTGTGAATGCCAACATCCCCAAGGCTTGCTG
CGTGCCCACCGAGCTGAGCGCCATCTCCATGCTCTACCTGGACGAGAACGAAAAGGTGGTGCTCAAGAATTACCAGGACATGG
TCGTGGAGGGTTGTGGGTGTCATTAG

>EcoBMP4 (Common coqui; Accession number AY525159)
ATGATTCCTGGTAATCGAATGCTGATGGTCATTTTATTATGCCAAGTCCTGCTAGGAGGTACTAGCCATGCTAGTCTGATCCC
TGAGACCGGAAAGAAGAAAGTGGCTGAGATTCAGGGCCAGGCGGGTGGAAGGAGGTCTGCTCAGAGCCATGAGCTCTTGCGGG
ACTTTGAGACAACGCTGCTGCAGATGTTTGGCCTCCGCAAGCGGCCGCAACCCAGTAAGGAAGTGGTGGTGCCTGCGTATATG
CGGGACTTGTACCGACTACAGTCAGCAGAGGAAGAGGATGAGTTTCAGGAGATCAACTTGGAATATCCTGAGAGACCTACAAG
CCGTGCCAATACCGTGAGGAGCTTCCACCATGAAGAATATTTGGAGAACATACCGGGCACAGCTGAAAACGCCGGCATCCGCT
TTTACTTCAATCTCAGCAGCATTCCTGAGAATGAGGTGATTTCTTCAGCTGAGCTAAGACTTTATAGAGAACAAATTGAACAT
GGCCCAGCATGGGAAGAAGGCTTTCATAGAATAAATATATATGAAGTTATGAAGACTCTGACAGCAAATGGACAGATGATTAC
TAGGCTACTGGACACCAGACTAATCCATCACAATGTGACTCGGTGGGAAAGTTTTGATGTAAGTCCAGCAATTATGAGGTGGA
CCCAAGTCAAGGGAATTAACCATGGGCTTGCTGTTGAGGTCATTCATCTCAACCAGACTAAAACTTATCAGGGGAAGCATGTC
AGGATTAGCCGATCGTTATTACCTCAAGATAATGCAGACTGGTCACAGATGAGGCCACTTCTAATTACATTTAGCCATGATGG
CAGGGGACATGCACTGACCAGGAGGTCCAAAAGAAGTCCTAAACCGCAGAGAGCCCGTAAAAAGAACAAAAATTGCCGGAGGC
ATTCTCTGTATGTGGATTTCAGCGATGTAGGCTGGAATGATTGGATTGTGGCACCACCTGGATATCAGGCTTTTTACTGCCAC
GGTGATTGCCCGTTTCCCCTGGCTGACCACCTAAACTCAACCAACCATGCCATTGTACAGACTCTGGTTAACTCCGTAAATTC
AAGCATTCCCAAAGCATGCTGTGTTCCCACAGAACTGAGTGCTATTTCCATGCTCTATTTGGACGAGTATGACAAAGTAGTCC
TTAAAAACTATCAGGAGATGGTAGTCGAGGCATGTGGCTGCGCTAA

>XtrBMP2 (Western clawed frog; Accession number NM_001015963 and AC150281)
ATGGTAGCTGGGATCCACTCTCTGCTCCTGCTGCTGTTTTACCAGATTTTGCTGAGTGGCTGCACCGGACTTATACCAGAGGA
AGGCAAGCGGAAGTACACGGAATCTGGTCGCTCGTCTCCGCAGCACTCCCAGAGTGTACTCAACCAGTTCGAGCTCCGGCTGC
TCAATATGTTTGGCTTAAAGAGGCGTCCAACGCCTGGCAAAAACATTGTGATCCCACCCTACATGCTGGACTTGTACCACCTG
CATTCAGGTCAGCTGGCTGCTGATCAAGACAGTTCCCCCATGGACTACCAGATAGAGCGTGCAGCTAGCCGAGCAAACACCGT
TAGGAGCTTTCACCATGAAGAATCCATGGAAGAAATTCCCGAGTCTGGCGAGAAAACAATCCAACGATTCTTCTTCAACCTTT
CTTCAGTTCCGAACGAGGAGCTGGTCACTTCTGCCGAGCTACGGATTTTTCGAGAGGGGGTCCAGGAGCCATTTGAGGGTGAC
AGCAGCAAACTTCATCGGATTAATATTTATGATATTGTCAAGCCAGCAGCGGCTGCCTCCCGGGGCCCCGTTGTAAGACTATT
GGACACCAGACTGATACATCATAACGAAAGCAAATGGGAAAGTTTTGATGTGACACCGGCAATTACACGGTGGATTGCACATA
AACAGCCGAACCATGGGTTTGTTGTTGAAGTGACTCACTTGGACAATGACAAAAATGTGCCGAAGAAGCATGTGAGGATTAGC
AGGTCTTTAGTCCCAGATAAAGATAGCTGGCCTCGGATACGGCCGTTGTTGGTGACTTTTAGCCACGACGGCAAAGGACATGC
GCTTCACAAAAGAGAAAAGCGCCAAGCAAGGCACAAACAACGGAAACGCCTTAAATCGAGCTGCAGGAGGCATCCGTTGTATG
TCGATTTCAGCGACGTTGGTTGGAATGACTGGATTGTTGCCCCACCGGGGTATCATGCCTTTTACTGCCACGGGGAATGCCCT
TTCCCACTGGCAGACCATTTAAACTCTACAAACCACGCTATCGTGCAGACTTTGGTGAATAACGTCAACCCAAACATCCCCAA
AGCTTGCTGTGTACCCACAGAACTCAGCGCCATCTCCATGCTCTACCTGGACGAGAACGAAAAAGTAGTATTAAAAAATTATC
AGGACATGGTGGTGGAGGGGTGCGGCTGCCGTTAG

>XtrBMP4 (Western clawed frog; Accession number NM_001017034 and AJ315161)
ATGATTCCTGGTAACCGAATGCTGATGGTCATTTTATTATGCCAAGTCCTGCTAGGAGGCACTAACCATGCCAGCCTGATACC
TGAGACGGGCAAGAAGAAAGTAGTGGCCGAGATTCAGGGAGGTAGAAGGTCCGCTCAGAGCAATGAGCTCTTGCGGGATTTCG
AGGTGACGCTGCTGCAGATGTTCGGCCTCCGCAAGCGGCCGCAGCCCAGTAAGGATGTGGTGGTGCCCGCGTATATGCGGGAC
CTGTACAGGCTTCAGTCAGCGGAGGAGGAGGATGAACTACACGATATCAGCATGGAGTACCCGGAGAGACCGACCAGCCGGGC
TAACACCGTGAGGAGCTTCCATCATGAGGAGCATTTGGAGAATCTACCAAGCACAGCAGAAAATGGAAATTTCCGTTTTGTGT
TCAACCTCAGCAGCATTCCAGAGAATGAGGTGATTTCTTCAGCAGAACTAAGACTCTATAGAGAACAAATAGACCATGGTCCA
GCGTGGGAAGAGGGTTTCCACCGGATAAATATATATGAAGTTATGAAACCAATAACAGCAAGTGGACACATGATAAGTAGGCT
GCTGGACACACGGCTAATCCACCACAATGTGACACAGTGGGAAAGTTTTGATGTAAGCCCTGCAATTATAAGGTGGACCCGGG
ATAAACAGATAAACCATGGGCTTGCCATTGAGGTTGTTCACCTCAACCAAACAAAAACTTATCAGGGGAAGCATGTAAGGATA
AGTCGATCATTATTACCTCAAGAGGATGCAGACTGGTCACAGATGAGACCGCTTTTAATTACATTCAGCCATGATGGCAGGGG
ACATGCACTGACCAGGAGGTCAAAAAGAAGTCCAAAACAGCAAAGACCCCGTAAAAAAAAATAAACATTGCCGGAGGCATTCTC
TTTATGTGGATTTCAGCGATGTTGGCTGGAATGACTGGATTGTGGCACCTCCTGGATACCAGGCTTTTTACTGCCATGGAGAT
TGTCCATTTCCCTTGGCTGATCACCTAAACTCAACTAACCATGCTATTGTACAAACTCTGGTAAACTCTGTTAACTCAAGCAT
CCCAAAAGCGTGTTGCGTCCCCACAGATCTGAGTGCTATCTCCATGCTTTATTTGGATGAATATGACAAAGTCGTCCTTAAAA
ACTATCAGGAGATGGTGGTGGAAGGGTGTGGGTGCCGTTAA

>AcaBMP2 (Green anole; Accession numbers XM_003215309 , EU402671, FG668269,
FG668317, FG670828, FG670684 and AAWZ02005417)
ATGACCATGGTTGCCAGGACCTGCTCTCTACTAGCACTGCTGCTTTCTCACATATTGCTTGGTGGTTCGGCCAGTCTCATCCC
AGAGGTGGGGGACCGGCGCAGGTTCAGTAGCGATTTGGGTCAGGCTGCCCCCCTGCAGCTCTCGGAAGGGATCCTGCGTGAAT

TTGAACTTCGTCTGCTCAACATGTTTGGGCTCCAGCGACGGCCGACCCCAAGCAAGAATGCCATAATTCCTCCTTATATGTTG
GAGCTCTACCACCTGCACACGAGTCAGAAGCCAGCTACCATGGACTATAATCTGGAGAGAGCTACCAGCAGGGCCAATACAGT
TAGGAGCTTCCACCATGAGGAATCACTGGAAGAGCTTCCTGAAAGAAATGGGAAAACATCCCGGCGTTTCTTCTTTGACTTAG
CATCCATTCCAAGTGGGGAGTTTATCACTTCAGCTGAGCTCCAGATTTTTCGGGAGCAGGTCCAAGATACCTTTGGTAGCAAC
AGCAGCTCTCACCATCGTATTAATATTTATGAAATAATAAAGCCAGAGGGAGAGGCTTCTAAGGACCCTGTCACAAGACTTTT
GGACACCCGATTAGTGCATCACAATGCCAGCAAATGGGAAACCTTTGATGTAACACCAGCGCTAAGGAGATGGGTTGCGCATG
GACAACCCAATCATGGTTTTGTGGTTGAGGTGTTGCATTTGGACAAGAAGAACAATGTCTCCAAGAGGCATGTCAGGATCAGC
AGGTCATTGCATCAGGACGATGGGAGCTGGTCTCAGCTCAGGCCGCTACTAGTAACCTTTGGACATGATGGGAAGGGCCACCC
TCTCCATAGGCGAGAGAAACGCCAGGCAAAATATAAAGAACGCAAACGCCATAAATCCCATTGCAAAAGGCATCCTTTATACG
TAGACTTCAATGATGTAGGGTGGAATGACTGGATTGTGGCCCCACCGGGCTATAGTGCCTACCACTGCCGTGGAGATTGTCCC
TTTCCATTGGCAGATCACCTCAACTCCACAAATCATGCCATTGTTCAGACTCTGGTCAATTCAGTGAATTCCAAAATCCCCAA
AGCTTGTTGTGTGCCGACGGAACTGAGTCCTATCTCCATGCTCTACCTTGATGAAAATGAAAAGGTTGTACTCAAGAACTATC
AAGACATGGTTGTGGAGGGCTGTGGTTGCCGCTGA

>AcaBMP4 (Green anole; Accession numbers GL343794.1)
ATGAAGCAGCCTAGAATTCCCTTGCTGTCCCATGTTGTAATGCCGGCGCTGTTTTTCCTTTCCTCCCGCTTCCCAGAGCACCT
GGAGCAAGTGCCGGGTTCGGATCCCCGCCTGGGTCGCTTCCGCTTCTTCTTCAACCTGAGCAGCGTGCCGGCGGTTGAGGCCA
TCTCTACCGCAGAGCTACGTCTCTTTCGTGAGCGGATCCCCGATGCTGCGGAGGACGACGAGGGTGACCTCCACCGGATCAAC
ATTTACGAGGTGATGAAGCCGTGGCGGGACGGGGAGGACCCGGTCACCCGGCTGCTGGACACTCGCCTGGTGCGCCACAACGT
GAGCGGCTGGGAGAGCTTCGACGTCAGCCCAGCGGTGTCGCGATGGACGGCGGGCGGTCAGCCCAACCACGGCTTGGCGGTGG
AAGCCCTCCACCTCCACCCCAAGCAGACTCACCAGGGCCGGCACGTCCGGATGAGCCGCTCTTTACCTCAAGGCGGGGGGCGC
CTAGAGGACGGCGCTTGGGCCCACCTCCGTCCACTTCTGGTGACTTTTGGCCACGACGGTCCCAGTGGCTCGGCCCTGACGCG
GCGCCGGAGGGCCAAGCGGAGCCCCAAAGGACCCGGAGGCTCCGGCGGGCGGGCACAACGCAAGGGCAAGAAGAACTGCCGGC
GCCATGCCCTCTATGTGGACTTCAGTGACGTGGGCTGGAACGATTGGATTGTGGCTCCGCCGGGCTACCAAGCCTTCTACTGC
CACGGCGACTGCCCCTTCCCCTTGGCCGACCACCTCAACTCCACCAACCACGCCATCGTCCAGACGCTGGTCAACTCCGTCAA
CGGCAGCATCCCCAAGGCGTGCTGCGTGCCCACCGAGCTGAGCGCCATCTCCATGCTCTATTTGGACGAGTACGACAAGGTGG
TCCTCAAGAACTACCAGGAGATGGTGGTCGAAGGATGCGGGTGCCGTTGA

>AcaBMP16 (Green anole; Accession numbers GAFZ01245723.1)
ATGGTGTCTGGTAACTTGGCTGTTCTGTCGATGCTTTTTGTGTCACTTGCCTTCTGGAGATTTGAAGCCCCCCAGGGAGCATC
ACCTGGAACCGACAAACATGACCCTAGAGAACTGCCAATCCAGGCCCTCCACACTATCCTGTTAAGACGCCTAGGCTTGCAAA
GGCGGCCGGAACCAAAACCAGGCCACGTGGTCCCCCAGTACCTGCTGGATTTGTACCAATTTGCCTTAGGAAGGGTTCCTCCC
TCCTTGCAGGGAACGGATCTACCTTTCTTGGAGGAACAGGCGGGAAGAGCCAACACTGTGCGCACATTTCATCATGTCGAGGA
TCCAGATTACTTTCCTGACCTTGTAGGAAATGAATTTTACTTCTTGTTCAACCTCACCTTGCTGCCTGCAGAGGAGGAGCTGA
CTGCCCTTGAGTTGCGCCTCTACCACAGAGAGAGAGAAAGCAAGGGTTTTCACATCAACGTCTACCACACAATGGATCCCCCA
CTCATTTCCAGAAACAAGAGCAGACTCCTGGCCCGCAAATTCTTGGCTCCAAACCAGTCCAAGTGGGAAAGCTTTGATGTGAC
TCCTGCCTATAAGACAGCCAAGAACCAGAAGCTTCATCTTGGGTTTCTTGTTGAGGTGGAGCACCCAAACAATAGCTATAACC
TTCAAGAGCACCACACATCCGCTCGTGCAAGACGGTCTCTTGGAGAGGAGGCAGCTCAGTGGGCTCTGGAGAGGCCTTTACTA
GTTACTTATAGCCACGACCAAAGAGGGCAACCCCTGACACGTGGAAAAAGACACAACAGAAGTCAGTCAAATGCACTAACCCA
GAAAGGAGGCAGAAAACCTACAAAAGGTGCGGTCTCCAAGAACAAACTCAAGAGCCTAAAGGCAAAGACCAAAATCAGTACCA
AGTGCAAAAGGCACCGCTTGTTTGTAGATTTCAAAGAGGTTGGTTGGAACGATTGGATAGTTGCCCCCAGTGGCTACCACGCA
TTCTACTGCTCTGGGGAATGTCAGTTCCCACTGGCTGACCATATGAATTCCTCCAGTCATGCTGTGGTGCAGACAATGTTGAA
TTCTGTGAACTCCAAAGTGCCTAAGCCATCTTGTGTTCCTACAGACCTCAGTCCCATAGCTATGCTCTATCTGGATCAACATG
ACATGGTGGTCCTTAAAACCTATCAGGATATGGTGGTGCAGGGCTGTGGGTGCCGGTAA

>GgaBMP2 (Chicken; Accession numbers NM_204358, X75914, BU137064, BU114214,
BU444424, AL585496 and AADN02041168
ATGGTTGCCGCCACCCGCTCCCTCCTGGCGCTGCTGCTCTGCCGGGTGCTGCTGGGCGGCGCGGCCGGCCTCATGCCGGAGGT
GGGACGGCGGCGCTTCAGCGAACCGGGCCGCGCCGCCTCGGCCGCGCAGCGCCCCGAGGACCTCCTGGGCGAGTTCGAGCTGC
GCCTGCTCCACATGTTCGGGCTGAAGCGGCGGCCGAGCCCCGGCAAGGACGTCGTCATCCCCCCCTACATGTTGGACCTCTAT
CGCCTGCACGCCGGCCAGCAGCTGGGCTACCCGCTGGAGAGGGCCGCCAGCCGCGCCAACACCGTGCGCAGCTTCCACCACGA
AGAAGTTTTGGAAGAACTGCCAGAAACAAGTGGGAAAACAGCACGACGTTTCTTCTTTAATTTAACTTCCATCCCTAATGAGG
AGTCTGTCACCTCAGCTGAACTCCAGATTTTTCGGGAGCAGGTGCACGAAGCCTTTGAGAGCAACAGCAGCTACCATCACCGT
ATTAATATTTATGAAATTATGAAGCCAGCCACAGCCACCTCCAAGGACCCTGTCACGAGACTTTTGGACACCAGGTTGGTGCA
TCATAATGCAAGTAAATGGGAAAGTTTTGATGTAACGCCAGCTGTTTTGAGGTGGATTGCACACGGACAACCTAACCATGGGT
TTGTGGTGGAGGTGGTTCACTTGGACAAAGAGAACAGTGCCTCCAAGAGGCACGTTAGGATTAGCAGGTCTTTACATCAGGAT
GAAGATAGCTGGTCTCAGCTCAGGCCGTTGTTAGTGACGTTTGGGCATGATGGCAAGGGACACCCGCTCCACAAAAGAGAAAA
GCGTCAAGCGAAACACAAACAGCGTAAGCGCCACAAATACAGTTGCAAAAGGCATCCGTTGTATGTGGACTTCAATGACGTGG
GGTGGAATGACTGGATTGTTGCCCCGCCGGGGTACAGTGCCTTTTACTGCCATGGGGAATGTCCTTTTCCGCTGGCAGATCAC
CTAAACTCAACAAACCATGCCATTGTTCAGACTTTGGTCAATTCGGTGAATTCCAAAATCCCCAAGGCTTGCTGTGTGCCGAC
AGAACTGAGTGCTATCTCAATGCTCTACCTTGACGAGAACGAAAAGGTCGTACTAAAGAACTATCAAGATATGGTTGTGGAGG
GCTGCGGGTGCCGCTGA

>GgaBMP4 (Chicken; Accession numbers NM_205237, AC172371, BM426811, BU473912,
BU244511, BU473786, BU268076, BU355305, BU217696, BU256447, BX259520, BU473508 and
BU227272)
ATGATTCCTGGTAACCGAATGCTGATGGTCATCCTACTATGCCAAGTCCTGCTGGGAGGTACCAACCATGCTAGCCTGATCCC
CGAGACCGGCAGGAAGAAAGTCGCAGAGCTTCAGGGACAAGCCGGATCCGGACGCCGCTCTGCCCAAAGCCATGAACTCTTGC

GGGGTTTCGAAACGACTCTGCTGCAGATGTTTGGGCTGCGAAGGCGGCCCCAGCCCAGCAAATCAGCCGTCATCCCCAGTTAC
ATGCTGGATCTCTACCGGCTCCAGTCCGGAGAAGAGGAGGAGAGCCTCCAGGAGATCAGCCTGCAGTACCCGGAGCGATCGAC
CAGCCGGGCCAACACCGTGAGGAGCTTCCACCATGAAGAGCACCTGGAGAGCGTGCCGGGTCCCAGCGAAGTGCCGCGGATCC
GCTTCGTCTTCAACCTCAGCAGCGTGCCGGACAACGAGGTGATCTCGTCGGCGGAGCTGCGGCTGTACCGGGAGCAGGTGGAG
GAGCCGAGCGCGGCGTGGGAGAGGGGCTTCCACCGGATAAACATTTACGAAGTGATGAAGCCGCTGTCGGAGCGCTCGCAGGC
CATTACGCGCCTGTTGGACACGCGGTTGGTGCACCACAACGTGACGCGCTGGGAGACCTTTGATGTGAGCCCAGCCGTGATCC
GGTGGACCAAGGACAAGCAGCCCAACCACGGGCTGGTGATCGAGGTGACCCACCTCCACCAGGCACAGACTCATCAGGGCAAA
CACGTCAGGATTAGCCGATCTTTACCTCAGGGGTACGGCGGGGACTGGGCGCAGCTCAGGCCGCTCCTGGTCACCTTCGGGCA
CGACGGGCGAGGCCACGCGCTGACCCGCAGGGCCCGCCGCAGCCCCAAGCACCAGCGTTCCCGCAAGAACAAAAAAAACTGCC
GCCGCCACGCTCTCTATGTGGATTTCAGCGACGTGGGTTGGAACGATTGGATCGTGGCCCCCCCGGGCTACCAGGCGTTTTAC
TGCCACGGGGACTGCCCCTTCCCTCTGGCCGACCACCTCAACTCCACCAACCACGCCATCGTGCAGACGTTGGTCAACTCCGT
CAACTCCAGCATCCCCAAGGCCTGCTGCGTGCCCACGGAGCTGAGCGCCATCTCCATGCTCTACCTGGATGAGTATGACAAGG
TGGTGCTGAAAAACTACCAGGAGATGGTGGTGGAGGGGTGCGGGTGCCGCTGA

>CpbBMP2 (Painted turtle; Accession number AHGY01038458)
ATGGTGGCCGCCATCCGCTCCCTGCTAGCGCTGCTGCTGTACCAGGTGCTGCTGGGCGGCTCGGCCGGGCTCATCCCGGAGGT
GGGCCGGCGGCGCTTCAGCGACTCGGGCCGCGCCGCCCCGCAGGAGCTGCTGAGCGAGTTCGAGCTGCGTCTGCTCAACATGT
TCGGGCTGAAGCGGCGGCCGAGCCCCAGCAAGAACGCCGTCATCCCGCCCTACATGCTGGACCTCTACCGCCTGCAGGCCGGC
CAGCAGCTGGGCCACTCCGCCCCGGGCTACCCGCTGGAGAGAGCCGCCAGCCGCGCCAACACGGTCCGGAGCTTCCACCACGA
AGAAGCTTGGGAAGAGCTCCCAGAAGCAAGTGGGAAAACAGCACGGCGCTTCTTCTTTAATTTAACTTCCATCCCTAACGAGG
AGTTTCTCACCTCGGCTGAACTTCAGATTTTTCGGGAGCAGGTGCAGGAAACCTTTGAGAACAATAGCAGCTACCATCACCGT
ATTAATATTTATGAAATTATAAAGCCAGCCACAGCCACTTCTAAGGACCCTGCCGCAAGACTTTTGGACACCAGGTTGGTGCA
TCAGAATGCAAGTAAATGGGAAAGCTTTGATGTAACACCAGCTGTAATGAGGTGGATTGCACATAGACAGCCTAATCATGGGT
TTATAGTAGAGGTGGTTCACTTGGACAATGAGAGCAGTGTTTCGAAGAGGCACGTTAGGATTAGCAGGTCTTTGCATCAGGAT
GAAGATAGCTGGTCTCAGCTCCGGCCGTTGTTAGTAACTTTTGGCCATGATGGCAAGGGACACTCTCTCCATAAAAGAGAGAA
ACGCCAAGCAAAACACAAACAGCGCAAACGCCACAAATCCAGTTGCAAAAGACATCCTTTATATGTGGACTTCAATGACGTGG
GGTGGAATGACTGGATTGTTGCCCCACCAGGATATGGTGCCTTTTACTGCCATGGGGAATGCCCTTTTCCATTGGCAGATCAC
CTAAACACAACAAACCATGCCATTGTTCAGACTTTGGTCAATTCCGTGAATTCCAAAATCCCTAAGGCTTGCTGTGTTCCGAC
AGAACTGAGCGCTATTTCCATGCTCTACCTTGACGAGAACGAAAAAGTTGTATTAAAGAACTATCAAGATATGGTTGTGGAGG
GTTGTGGGTGCCGCTAA

>CpbBMP4 (Painted turtle; Accession number XM_005295477.1)
ATGATTCCTGGTAACCGAATGCTGATGGTCATTCTGTTATGCCAAGTTCTGCTAGGAGGTACTAACCATGCCAGTTTGATACC
CGAGACGGGGAGGAAGAAAGTTGGAGAACTCCAGGCTCACGCAGGATCTGGACGCCGCTCCGGGCAAAGCCATGAACTCTTGC
GGGGTTTCGAGGCGAGTTTGCTGCAGATGTTCGGGCTCCGCAGGCGGCCCCAGCCCAGCAAAGCAGCCGTCATCCCCGCGTAC
ATGCTGGATCTCTATCGGCTGCAGTCCGGGGAAGAGGAGGAAAGCCTGCAGGACATCAGCCTGCGGTATCCCGAGAGATCCAC
TAGCCGGGCCAACACCGTGAGGAGTTTCCACCATGAAGAACATCTGGAGAACCTCCCCGGGCCCAGCGACACCCCCCGCGTTC
GCTTCCTCTTCAACCTCAGCAGCGTGCCGGAGAACGAGGTGATCACCTCGGCCGAGCTGCGGCTGTACCGGGAGCAGCTGGAA
GCGCGGAGCCCCGCGTGGGAACGGGGCTCGCACCGGATCAACATCTACGAAGTGATGAAGCCCCCGCCGGCCCGCGGGCAGGT
GATCACGCGGCTGCTGGACACGCGGCTGGTGCATCACAACGTGTCCCGGTGGGAGAGCTTCGACGTGAGCCCCGCCGTCCTCC
GCTGGACCCAGGACAGGCAGCCCAACCACGGGCTGGCGGTGGAGGTGACCCCGCTGCCCCCGGCCCCCGCGGGCCCCGGGAAC
CACGTCAGGATTAGCCGATCTTTACCTCAAGGAGACGGGGACTGGGCCCAGTTGAGGCCGCTGCTGGTCACTTTCAGCCACGA
CGGCAGGGGCCACGCACTGACCCGCCGGGCCCGCCGCAGCGCCAAGCACCAGCGGCCCCGCAAGAACAAAAAAAACTGTCGAC
GCCACGCGCTCTATGTGGATTTCAGCGACGTGGGCTGGAACGACTGGATCGTGGCCCCCCCGGGCTACCAGGCCTTTTACTGC
CACGGGGACTGCCCCTTCCCTTTGGCCGACCACCTCAACTCCACCAACCACGCCATCGTGCAAACCCTGGTGAACTCCGTCAA
CGCCAGCATCCCCAAGGCCTGCTGCGTGCCCACGGAGCTGAGCGCCATCTCCATGCTCTACCTGGACGAGTATGACAAAGTGG
TCCTGAAAAACTACCAGGAGATGGTGGTGGAGGGGTGCGGGTGCCGTTAA

>HsaBMP2 (Human, Accession numbers NM_001200 and NT_011387)
ATGGTGGCCGGGACCCGCTGTCTTCTAGCGTTGCTGCTTCCCCAGGTCCTCCTGGGCGGCGCGGCTGGCCTCGTTCCGGAGCT
GGGCCGCAGGAAGTTCGCGGCGGCGTCGTCGGGCCGCCCCTCATCCCAGCCCTCTGACGAGGTCCTGAGCGAGTTCGAGTTGC
GGCTGCTCAGCATGTTCGGCCTGAAACAGAGACCCACCCCCAGCAGGGACGCCGTGGTGCCCCCCTACATGCTAGACCTGTAT
CGCAGGCACTCAGGTCAGCCGGGCTCACCCGCCCCAGACCACCGGTTGGAGAGGGCAGCCAGCCGAGCCAACACTGTGCGCAG
CTTCCACCATGAAGAATCTTTGGAAGAACTACCAGAAACGAGTGGGAAAACAACCCGGAGATTCTTCTTTAATTTAAGTTCTA
TCCCCACGGAGGAGTTTATCACCTCAGCAGAGCTTCAGGTTTTCCGAGAACAGATGCAAGATGCTTTAGGAAACAATAGCAGT
TTCCATCACCGAATTAATATTTATGAAATCATAAAACCTGCAACAGCCAACTCGAAATTCCCCGTGACCAGACTTTTGGACAC
CAGGTTGGTGAATCAGAATGCAAGCAGGTGGGAAAGTTTTGATGTCACCCCCGCTGTGATGCGGTGGACTGCACAGGGACACG
CCAACCATGGATTCGTGGTGGAAGTGGCCCACTTGGAGGAGAAACAAGGTGTCTCCAAGAGACATGTTAGGATAAGCAGGTCT
TTGCACCAAGATGAACACAGCTGGTCACAGATAAGGCCATTGCTAGTAACTTTTGGCCATGATGGAAAAGGGCATCCTCTCCA
CAAAAGAGAAAAACGTCAAGCCAAACACAAACAGCGGAAACGCCTTAAGTCCAGCTGTAAGAGACACCCTTTGTACGTGGACT
TCAGTGACGTGGGGTGGAATGACTGGATTGTGGCTCCCCCGGGGTATCACGCCTTTTACTGCCACGGAGAATGCCCTTTTCCT
CTGGCTGATCATCTGAACTCCACTAATCATGCCATTGTTCAGACGTTGGTCAACTCTGTTAACTCTAAGATTCCTAAGGCATG
CTGTGTCCCGACAGAACTCAGTGCTATCTCGATGCTGTACCTTGACGAGAATGAAAAGGTTGTATTAAAGAACTATCAGGACA
TGGTTGTGGAGGGTTGTGGGTGTCGCTAG

>HsaBMP4 (Human; Accession number NM_001202.3)
ATGATTCCTGGTAACCGAATGCTGATGGTCGTTTTATTATGCCAAGTCCTGCTAGGAGGCGCGAGCCATGCTAGTTTGATACC
TGAGACGGGGAAGAAAAAAGTCGCCGAGATTCAGGGCCACGCGGGAGGACGCCGCTCAGGGCAGAGCCATGAGCTCCTGCGGG

ACTTCGAGGCGACACTTCTGCAGATGTTTGGGCTGCGCCGCCGCCCGCAGCCTAGCAAGAGTGCCGTCATTCCGGACTACATG
CGGGATCTTTACCGGCTTCAGTCTGGGGAGGAGGAGGAAGAGCAGATCCACAGCACTGGTCTTGAGTATCCTGAGCGCCCGGC
CAGCCGGGCCAACACCGTGAGGAGCTTCCACCACGAAGAACATCTGGAGAACATCCCAGGGACCAGTGAAAACTCTGCTTTTC
GTTTCCTCTTTAACCTCAGCAGCATCCCTGAGAACGAGGTGATCTCCTCTGCAGAGCTTCGGCTCTTCCGGGAGCAGGTGGAC
CAGGGCCCTGATTGGGAAAGGGGCTTCCACCGTATAAACATTTATGAGGTTATGAAGCCCCCAGCAGAAGTGGTGCCTGGGCA
CCTCATCACACGACTACTGGACACGAGACTGGTCCACCACAATGTGACACGGTGGGAAACTTTTGATGTGAGCCCTGCGGTCC
TTCGCTGGACCCGGGAGAAGCAGCCAAACTATGGGCTAGCCATTGAGGTGACTCACCTCCATCAGACTCGGACCCACCAGGGC
CAGCATGTCAGGATTAGCCGATCGTTACCTCAAGGGAGTGGGAATTGGGCCCAGCTCCGGCCCCTCCTGGTCACCTTTGGCCA
TGATGGCCGGGGCCATGCCTTGACCCGACGCCGGAGGGCCAAGCGTAGCCCTAAGCATCACTCACAGCGGGCCAGGAAGAAGA
ATAAGAACTGCCGGCGCCACTCGCTCTATGTGGACTTCAGCGATGTGGGCTGGAATGACTGGATTGTGGCCCCACCAGGCTAC
CAGGCCTTCTACTGCCATGGGGACTGCCCCTTTCCACTGGCTGACCACCTCAACTCAACCAACCATGCCATTGTGCAGACCCT
GGTCAATTCTGTCAATTCCAGTATCCCCAAAGCCTGTTGTGTGCCCACTGAACTGAGTGCCATCTCCATGCTGTACCTGGATG
AGTATGATAAGGTGGTACTGAAAAATTATCAGGAGATGGTAGTAGAGGGATGTGGGTGCCGCTGA

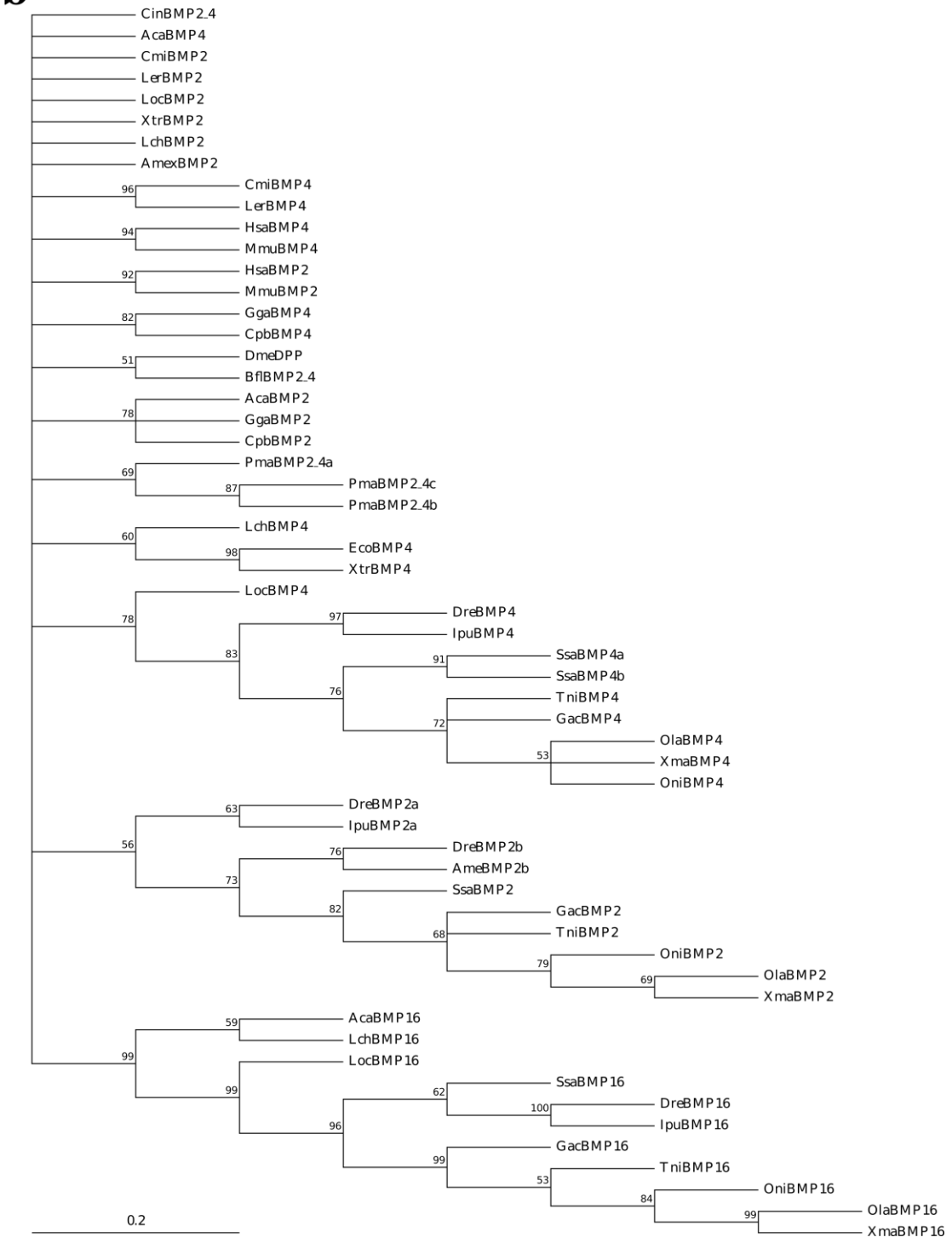
>MmuBMP2 (Mouse; Accession number NM_007553.2)
ATGGTGGCCGGGACCCGCTGTCTTCTAGTGTTGCTGCTTCCCCAGGTCCTCCTGGGCGGCGCGGCCGGCCTCATTCCAGAGCT
GGGCCGCAAGAAGTTCGCCGCGGCATCCAGCCGACCCTTGTCCCGGCCTTCGGAAGACGTCCTCAGCGAATTTGAGTTGAGGC
TGCTCAGCATGTTTGGCCTGAAGCAGAGACCCACCCCCAGCAAGGACGTCGTGGTGCCCCCCTATATGCTAGATCTGTACCGC
AGGCACTCAGGCCAGCCAGGAGCGCCCGCCCCAGACCACCGGCTGGAGAGGGCAGCCAGCCGCGCCAACACCGTGCGCAGCTT
CCATCACGAAGAAGCCGTGGAGGAACTTCCAGAGATGAGTGGGAAAACGGCCCGGCGCTTCTTCTTCAATTTAAGTTCTGTCC
CCAGTGACGAGTTTCTCACATCTGCAGAACTCCAGATCTTCCGGGAACAGATACAGGAAGCTTTGGGAAACAGTAGTTTCCAG
CACCGAATTAATATTTATGAAATTATAAAGCCTGCAGCAGCCAACTTGAAATTTCCTGTGACCAGACTATTGGACACCAGGTT
AGTGAATCAGAACACAAGTCAGTGGGAGAGCTTCGACGTCACCCCAGCTGTGATGCGGTGGACCACACAGGGACACACCAACC
ATGGGTTTGTGGTGGAAGTGGCCCATTTAGAGGAGAACCCAGGTGTCTCCAAGAGACATGTGAGGATTAGCAGGTCTTTGCAC
CAAGATGAACACAGCTGGTCACAGATAAGGCCATTGCTAGTGACTTTTGGACATGATGGAAAAGGACATCCGCTCCACAAACG
AGAAAAGCGTCAAGCCAAACACAAACAGCGGAAGCGCCTCAAGTCCAGCTGCAAGAGACACCCTTTGTATGTGGACTTCAGTG
ATGTGGGGTGGAATGACTGGATCGTGGCACCTCCGGGCTATCATGCCTTTTACTGCCATGGGGAGTGTCCTTTTCCCCTTGCT
GACCACCTGAACTCCACTAACCATGCCATAGTGCAGACTCTGGTGAACTCTGTGAATTCCAAAATCCCTAAGGCATGCTGTGT
CCCCACAGAGCTCAGCGCAATCTCCATGTTGTACCTAGATGAAAATGAAAAGGTTGTGCTAAAAAATTATCAGGACATGGTTG
TGGAGGGCTGCGGGTGTCGTTAG

>MmuBMP4 (Mouse; Accession number BC013459.1)
ATGATTCCTGGTAACCGAATGCTGATGGTCGTTTTATTATGCCAAGTCCTGCTAGGAGGCGCGAGCCATGCTAGTTTGATACC
TGAGACCGGGAAGAAAAAAGTCGCCGAGATTCAGGGCCACGCGGGAGGACGCCGCTCAGGGCAGAGCCATGAGCTCCTGCGGG
ACTTCGAGGCGACACTTCTACAGATGTTTGGGCTGCGCCGCCGTCCGCAGCCTAGCAAGAGCGCCGTCATTCCGGATTACATG
AGGGATCTTTACCGGCTCCAGTCTGGGGAGGAGGAGGAGGAAGAGCAGAGCCAGGGAACCGGGCTTGAGTACCCGGAGCGTCC
CGCCAGCCGAGCCAACACTGTGAGGAGTTTCCATCACGAAGAACATCTGGAGAACATCCCAGGGACCAGTGAGAGCTCTGCTT
TTCGTTTCCTCTTCAACCTCAGCAGCATCCCAGAAAATGAGGTGATCTCCTCGGCAGAGCTCCGGCTCTTTCGGGAGCAGGTG
GACCAGGGCCCTGACTGGGAACAGGGCTTCCACCGTATAAACATTTATGAGGTTATGAAGCCCCCAGCAGAAATGGTTCCTGG
ACACCTCATCACACGACTACTGGACACCAGACTAGTCCATCACAATGTGACACGGTGGGAAACTTTCGATGTGAGCCCTGCAG
TCCTTCGCTGGACCCGGGAAAAGCAACCCAATTATGGGCTGGCCATTGAGGTGACTCACCTCCACCAGACACGGACCCACCAG
GGCCAGCATGTCAGAATCAGCCGATCGTTACCTCAAGGGAGTGGAGATTGGGCCCAACTCCGGCCCCTCCTGGTCACTTTTGG
CCATGATGGCCGGGGCCATACCTTGACCCGCAGGAGGGCCAAACGTAGTCCCAAGCATCACCCACAGCGGTCCAGGAAGAAGA
ATAAGAACTGCCGTCGCCATTCACTATACGTGGACTTCAGTGACGTGGGCTGGAATGATTGGATTGTGGCCCCACCCGGCTAC
CAGGCCTTCTACTGCCATGGGGACTGTCCCTTTCCACTGGCTGATCACCTCAACTCAACCAACCATGCCATTGTGCAGACCCT
AGTCAACTCTGTTAATTCTAGTATCCCTAAGGCCTGTTGTGTCCCCACTGAACTGAGTGCCATTTCCATGTTGTACCTGGATG
AGTATGACAAGGTGGTGTTGAAAAATTATCAGGAGATGGTGGTAGAGGGGTGTGGATGCCGCTGA

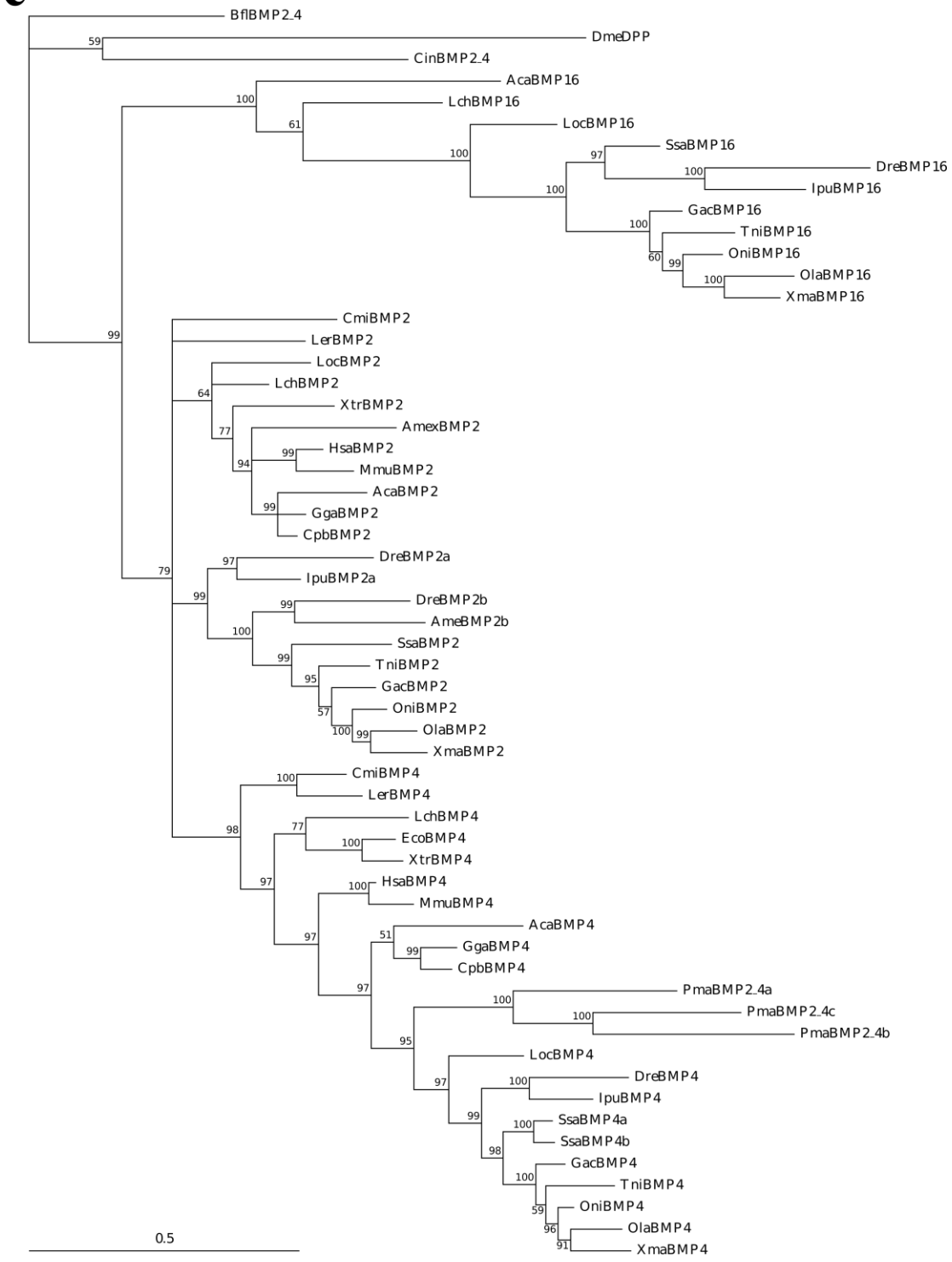
Supplementary figure 4.2.2. Phylogenetic analyses of 58 BMP gene sequences. (A) RAxML 7.8.4-MPI ML optimal tree search GTR+Γ+Fest 400 replicates. Likelihood: −lnL = 17343.6220; (B) RAxML 7.8.4-MPI ML bootstrap GTR+Γ+Fest 400 replicates; (C) MrBayes MCMC GTR+Γ+I marginal likelihood: −lnLh = 17419.5992. 2 chains, 5,000,000 generations, 10,000 samples, 2,000 burnin each chain; (D) P4 MCMC GTR+Γ+I marginal likelihood: −lnLh = 17430.2044. 2 chains, 1,000,000 generations, 5,000 samples, 200 burnin each chain. Posterior predictive simulations of $\chi 2$ statistic of composition homogeneity: original statistic = 469.8018, the distribution goes from 43.1524 to 142.2298, the tail-area probability = 0.0000; (E) P4 MCMC GTR+Γ+I+CV2 Marginal likelihood: −lnLh = 17045.6169. 2 chains, generations, samples, burnin each chain. Posterior predictive simulations of $\chi 2$ statistic of composition homogeneity: original statistic = 469.8018, the distribution goes from 248.3454 to 772.3124, the tail-area probability = 0.4613; (F) CodonPhyML GY+ω+κ+Fest non- synonymous/synonymous ω parameter value of 1.0 fixed for all sites. Likelihood: −lnL = 19333.7, κ = 2.3671

[Appendix 1]

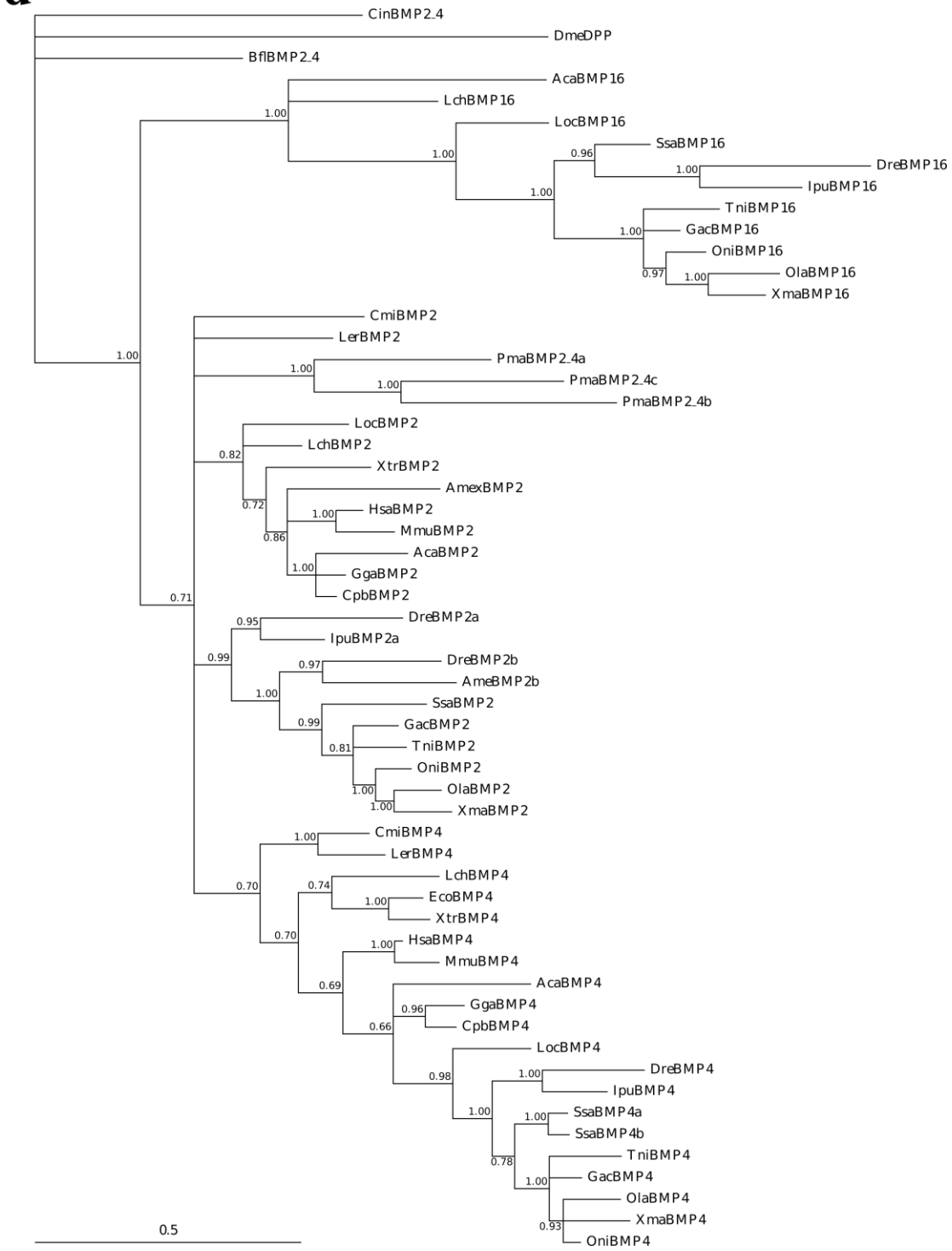
b


CinBMP2_4
AcaBMP4
CmiBMP2
LerBMP2
LocBMP2
XtrBMP2
LchBMP2
AmexBMP2
96
CmiBMP4
LerBMP4
94
HsaBMP4
MmuBMP4
92
HsaBMP2
MmuBMP2
82
GgaBMP4
CpbBMP4
51
DmeDPP
BflBMP2_4
78
AcaBMP2
GgaBMP2
CpbBMP2
69
PmaBMP2_4a
87
PmaBMP2_4c
PmaBMP2_4b
60
LchBMP4
98
EcoBMP4
XtrBMP4
78
LocBMP4
97
DreBMP4
IpuBMP4
83
91
SsaBMP4a
SsaBMP4b
76
TniBMP4
72
GacBMP4
53
OlaBMP4
XmaBMP4
OniBMP4
63
DreBMP2a
IpuBMP2a
56
76
DreBMP2b
AmeBMP2b
73
SsaBMP2
82
GacBMP2
68
TniBMP2
79
OniBMP2
69
OlaBMP2
XmaBMP2
59
AcaBMP16
LchBMP16
99
LocBMP16
99
62
SsaBMP16
100
DreBMP16
IpuBMP16
96
GacBMP16
99
TniBMP16
53
OniBMP16
84
99
OlaBMP16
XmaBMP16
0.2

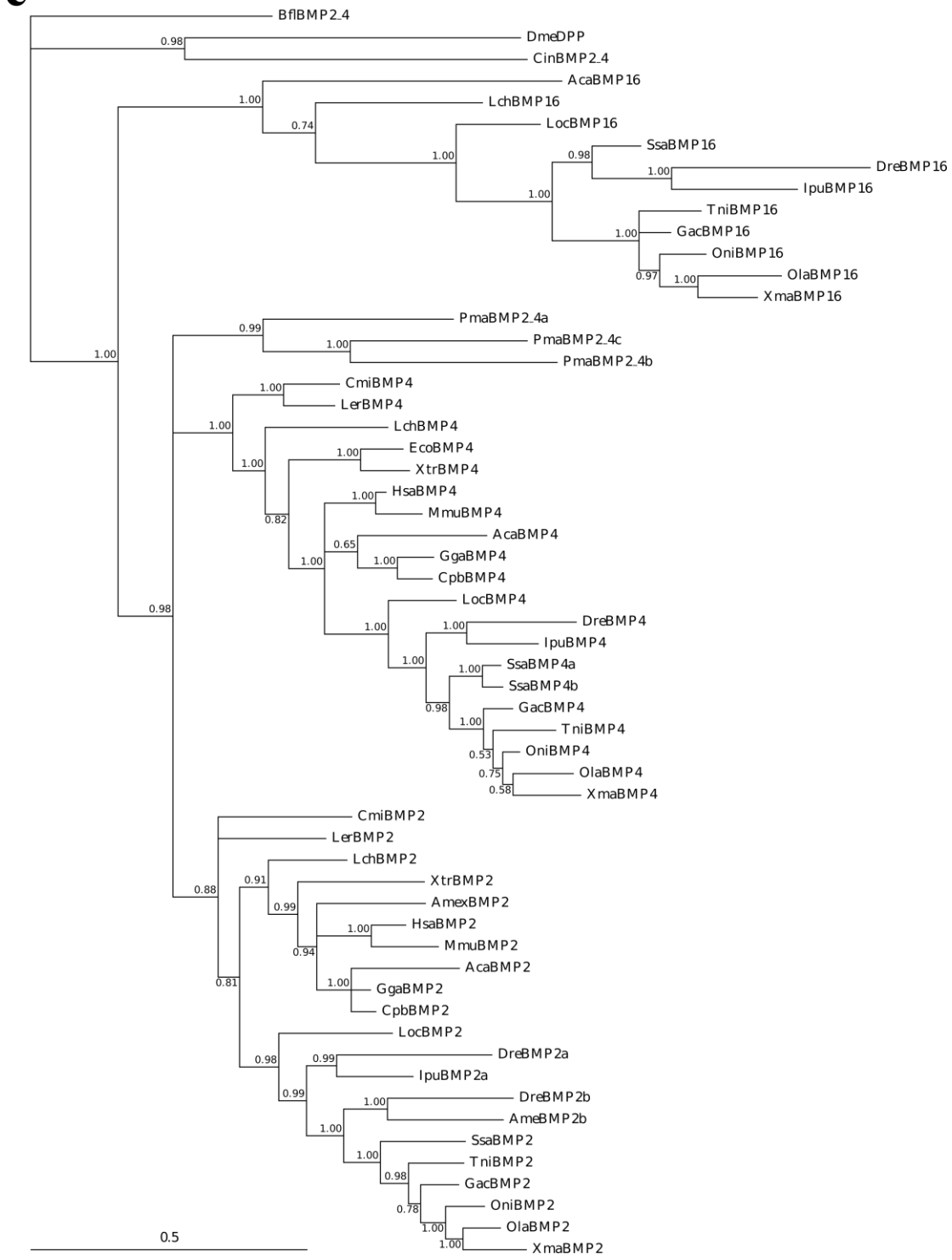
c


BflBMP2_4
DmeDPP
CinBMP2_4
AcaBMP16
LchBMP16
LocBMP16
SsaBMP16
DreBMP16
IpuBMP16
GacBMP16
TniBMP16
OniBMP16
OlaBMP16
XmaBMP16
CmiBMP2
LerBMP2
LocBMP2
LchBMP2
XtrBMP2
AmexBMP2
HsaBMP2
MmuBMP2
AcaBMP2
GgaBMP2
CpbBMP2
DreBMP2a
IpuBMP2a
DreBMP2b
AmeBMP2b
SsaBMP2
TniBMP2
GacBMP2
OniBMP2
OlaBMP2
XmaBMP2
CmiBMP4
LerBMP4
LchBMP4
EcoBMP4
XtrBMP4
HsaBMP4
MmuBMP4
AcaBMP4
GgaBMP4
CpbBMP4
PmaBMP2_4a
PmaBMP2_4c
PmaBMP2_4b
LocBMP4
DreBMP4
IpuBMP4
SsaBMP4a
SsaBMP4b
GacBMP4
TniBMP4
OniBMP4
OlaBMP4
XmaBMP4
0.5

d


CinBMP2_4
DmeDPP
BflBMP2_4
AcaBMP16
LchBMP16
LocBMP16
SsaBMP16
DreBMP16
IpuBMP16
TniBMP16
GacBMP16
OniBMP16
OlaBMP16
XmaBMP16
CmiBMP2
LerBMP2
PmaBMP2_4a
PmaBMP2_4c
PmaBMP2_4b
LocBMP2
LchBMP2
XtrBMP2
AmexBMP2
HsaBMP2
MmuBMP2
AcaBMP2
GgaBMP2
CpbBMP2
DreBMP2a
IpuBMP2a
DreBMP2b
AmeBMP2b
SsaBMP2
GacBMP2
TniBMP2
OniBMP2
OlaBMP2
XmaBMP2
CmiBMP4
LerBMP4
LchBMP4
EcoBMP4
XtrBMP4
HsaBMP4
MmuBMP4
AcaBMP4
GgaBMP4
CpbBMP4
LocBMP4
DreBMP4
IpuBMP4
SsaBMP4a
SsaBMP4b
TniBMP4
GacBMP4
OlaBMP4
XmaBMP4
OniBMP4
0.5

e


BflBMP2_4
DmeDPP
CinBMP2_4
AcaBMP16
LchBMP16
LocBMP16
SsaBMP16
DreBMP16
IpuBMP16
TniBMP16
GacBMP16
OniBMP16
OlaBMP16
XmaBMP16
PmaBMP2_4a
PmaBMP2_4c
PmaBMP2_4b
CmiBMP4
LerBMP4
LchBMP4
EcoBMP4
XtrBMP4
HsaBMP4
MmuBMP4
AcaBMP4
GgaBMP4
CpbBMP4
LocBMP4
DreBMP4
IpuBMP4
SsaBMP4a
SsaBMP4b
GacBMP4
TniBMP4
OniBMP4
OlaBMP4
XmaBMP4
CmiBMP2
LerBMP2
LchBMP2
XtrBMP2
AmexBMP2
HsaBMP2
MmuBMP2
AcaBMP2
GgaBMP2
CpbBMP2
LocBMP2
DreBMP2a
IpuBMP2a
DreBMP2b
AmeBMP2b
SsaBMP2
TniBMP2
GacBMP2
OniBMP2
OlaBMP2
XmaBMP2
0.5

Supplementary Table 4.2.1. Z-DOPE and QMEAN scores for the zebrafish BMP2, BMP4, BMP16 protein models.

Model	Z-DOPE score[1]	QMEAN score[2]
BMP2a	-0.659	0.669 (-0.96)
BMP2b	-0.746	0.623 (-1.42)
BMP4	-0.710	0.621 (-1.44)
BMP16	-0.260	0.646 (-1.21)

1 – The Z-DOPE score (Benkert et al., 2008) should be negative for acceptable quality models, and <-1 is usually seen for native structures

2 – The QMEAN score (Shen and Sali, 2006) ranges from 0 (worst) to 1 (best). Values in parenthesis are Z-scores based on a distribution of QMEAN scores for a set of experimentally determined structures (Z>-1.5 denote good models).

References:

Benkert P, Tosatto SCE, Schomburg D (2008). QMEAN: A comprehensive scoring function for model quality assessment *Proteins: Structure, Function, and Bioinformatics*. 71:261-277.

Shen MY, Sali A (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15: 2507-2524.