Osteocrin: identification, expression, and CRISPR/Cas9 analysis in scale development in teleost

by

Wendy Tri Prabowo

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 5, 2017

Keywords: osteocrin, scale development, CRISPR/Cas9, teleost

Approved by

Zhanjiang Liu, Chair, Alumni Professor of Fisheries and Allied Aquaculture
Rex A. Dunham, Alumni Professor of Fisheries and Allied Aquaculture
Charles Y. Chen, Associate Professor of Crop, Soil, and Environmental Sciences

Abstract

The completion of channel catfish genome sequencing provides a possibility of performing comparative genome studies. The comparative genome study between scaleless channel catfish and armor-scaled (bony dermal plate) catfish, such as common pleco (*Pterygoplichthys pardalis*) and striped Raphael catfish (*Platydoras armatulus*) reveal one gene namely osteocrin (*ostn*) was absent from channel catfish genome but present in armored-scale catfish genome. Additional BLAST analysis indicated that the osteocrin, however, was found in several scaled fishes, but was absent in the scaleless fish genome, suggesting that it could be a potential candidate gene involved in scale development, but detailed analysis of the status of *ostn* in various fishes has not been conducted. This initial observation is in line with the known functions of osteocrin in the development of calcified tissues such as bone. In this work, we conducted extensive comparative genome analysis to validate the correlation of the presence osteocrin with scaled fishes, and the absence of osteocrin with scaleless fishes. We found scaled fish model in phylogeny tree clustered based on their scale types: placoid, ganoid, cosmoid, elasmoid, and dermal bony plate, respectively. Next, we determined the expression patterns of the osteocrin genes during scale regeneration in common carp. Finally, we conducted a knockout experiment using CRISPR/Cas9 system to determine the phenotypes of osteocrin knockouts in zebrafish as the model. The disruption of *ostn* resulted in a number of phenotype defects, including pigment dispersion, imperfect operculum and asymmetrical gill rakers, bone deformities, slow scale growth, and partial scale lost on the heterozygous mutant zebrafish.

Keywords: osteocrin, genome comparative, scale regeneration, CRISPR/Cas9

Acknowledgments

First and foremost, I would like to express my profound and sincere gratitude to my advisor Dr. Zhanjiang Liu for his constant guidance, support, and patience throughout my master degree program. I would like to express my sincere gratitude to my committee: Dr. Rex Dunham and Dr. Charles Chen for their advice and critical reading of my thesis. My sincere thanks go to Ning Li, Tao Zhou, Yulin Jin, Xiaoyan Xu, Yang Liu, Changxu Tian, Zihao Yuan, Yujia Yang, Huitong Shi, Xiaozhu Wang, Qiang Fu, Lisui Bao, Suxu Tan, Wenwen Wang and Shikai Liu for their assistance and all the colleagues in the laboratory for their help, collaboration, and friendship. Many thanks to Dr. Jim Stoeckel, Dr. Ian Butts, and Dr. Michael Miller for their help in supporting the facilities in North Station and Life Science Laboratory. I would like to express my gratitude to the Indonesia Endowment Fund (LPDP) for the financial support for my master program. Finally, I would like to thank my beloved wife Richa, my mother in law Handayati, my parents Harry Sasono and Maryana, my son Alvin Liem and my little daughter Uma Liem for their love and support.

Table of Contents

List of Tables

List of Figures

List of Videos

List of Abbreviations

AU	Auburn University
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeats/ CRISPR-associated protein-9 nuclease
ZFN	Zinc Finger Nuclease
TALEN	Transcription Activator-Like Effector Nucleases
Pifo	Pitchfork protein gene
Ostn	Osteocrin gene
Wt	Wild type
ostn $^{-/}$	heterozygous mutant (knockout *ostn*)
ep	epipleural
en	epineural
an	anterior
pos	posterior

I. INTRODUCTION

Catfish are valuable resources for comparative study because of their morphologic variance. They comprise 12% of all teleost species and 6.3% of all vertebrates which means includes over 4,100 species belong to order Siluriformes (Wilson 2005; Sullivan, 2006; Liu, 2016). Catfish are worldwide distributed with high diversity suggesting a great interest in evolutionary studies (Sullivan 2006; Moyle 2004). In the United States, channel catfish are the most cultured fish species accounting over 60% of aquaculture production (Liu 2016). Several other catfish are important for aquaculture in other parts of the world, especially in Southeast Asia including the walking catfish (*Clarias batrachus*), pangasius (*Pangasius pangasius*), and basa catfish (*Pangasius bocourti*). Biologically, they are characteristic of two prominent features: 1) They have very prominent whisker-like barbels that led to their naming as the catfishes; and 2) For the most part, their bodies are naked without scales. However, barbelless catfish and scaled catfish do exist, making them ideal natural models for the analysis of evolutionary courses leading to such characters. For instance, channel catfish is scaleless, whereas common pleco (*Pterygoplichthys pardalis*) and striped Raphael catfish (*Platydoras armatulus*) harbor armored scales.

Comparative genomics is a valuable tool to determine the genomic basis for the similarities and differences between related species (Wei, 2002). The completion of channel catfish genome sequencing provides a possibility of performing comparative genome studies (Jiang, 2013; Geng, 2015; Liu 2016). In the previous study, comparative genome study between scaleless channel catfish and armor-scaled (bony dermal plate) catfish was performed, such as

common pleco (*Pterygoplichthys pardalis*) and striped Raphael catfish (*Platydoras armatulus*) (Liu 2016). A total 93 genes in common pleco and 91 genes from Raphael catfish were not found in the channel catfish genome sequences. In another side, a total of 169 armor-scaled catfish-specific genes were identified, after subtraction of the 17 genes that overlapped between the two armor-scaled catfishes, which were absent from the scaleless channel catfish (Zhang 2014). One possibility for the absence of these genes is that the genome sequencing of channel catfish may have been incomplete, but quality assessment did not provide strong support for this possibility. Among the 169 armor-scaled catfish-specific genes, two genes were annotated as functional genes related to bone or scale formation, pitchfork protein (*pifo*) and osteocrin-like gene (*ostn*) (Zhang 2014). Additional BLAST analysis indicated that the pitchfork protein gene actually was found also missing from some scaled fish species such as tilapia, suggesting that it may not be uniquely correlated with the status of scales among fishes (Zhang 2014). Osteocrin, however, was found in all scaled fishes, but was absent in the scaleless channel catfish genome, suggesting that it could be a potential candidate gene involved in scale development, in addition to secretory calcium-binding phosphoprotein.

Osteocrin gene (*ostn*), also described as musclin, is a skeletal muscle-derived factor (Liu, 2008). It is not homologous to any known gene except for two conserved sequence motifs reminiscent of dibasic cleavage regions found in peptide hormone precursors (Thomas, 2003). It is expressed in the skeletal tissue, especially in osteoblasts and young osteocytes at bone forming sites (Bord, 2005), suggesting its roles in bone formation (Moffatt, 2007; Bord 2005; Thomas 2003). Ostn was very recently found to be involved in cranial osteogenesis and chondrogenesis in zebrafish (Chiba, 2017). The limited number of studies provided strong evidence of ostn involvement in osteogenesis and bone formation. Our findings that it is present in armor-scaled

catfishes but absent from channel catfish suggested its potential as a candidate gene for scale development as well, but detailed analysis of the status of *ostn* in various fishes has not been conducted. In this work, an extensive comparative genome analysis was conducted to validate the correlation of the presence osteocrin with scaled fishes, and the absence of osteocrin with scaleless fishes. The scaled fish model in phylogeny tree was found to be clustered based on their scale types: placoid, ganoid, cosmoid, elasmoid, and dermal bony plate, respectively. Next, the expression patterns of the osteocrin genes during scale regeneration was determined in common carp, and finally, a knockout experiment using CRISPR/Cas9 system was performed to determine the phenotypes of osteocrin knockouts, providing evidence that osteocrin is involved in bone development as well as in scale development.

II. LITERATURE REVIEW

Genome comparative study

The genome sequence provides fundamental resources for the comparative analysis. Since 2007, the next generation sequencing generating multiple draft genome sequence which blows up in more genome comparative studies (Hu, 2011). Genome comparative study is a comprehensive method to study the similarities and differences between two or more genomes and transcriptomes (Wei 2002). The goal of comparative itself is valuable to understanding the evolutionary process and identifying conserved genes among the species including the unique genes which correlated to specific trait or phenotypes (Touchman 2010). The features of genome used for comparative study including the genome size, chromosome number, DNA sequence, genes order, functional regions, and others structural information (Brown, 2002; Touchman, 2010). The principle of genome comparative starts by making several alignments between the genome sequence to find the orthologous sequence which simply derived from same common ancestry (Ellegren, 2008). Through these results, we can define most of the genome features to reveal the evolutionary stages between the species at molecular level along with the predicted functional sequence of genes. In addition, genome comparative recently become standard analysis in assessing the completeness of newly genome sequence (Koonin 2003; Veeckman, 2016).

Genome comparative analysis using multiple sequences from close related species expand precisely the objectives of comparative genomics itself (Gumucio, 1992; Kumar, 2007; Hardison 2003). Catfish are valuable resources for comparative study because of their morphologic variance. They comprise 12% of all teleost species and 6.3% of all vertebrates

which means includes over 4,100 species belong to order Siluriformes (Sullivan 2006; Liu 2016). Catfish are worldwide distributed with high diversity suggesting a great interest in evolutionary studies (Sullivan 2006; Moyle 2004). In the United States, channel catfish are the most cultured fish species accounting over 60% of aquaculture production (Liu 2016). Major genetic enhancement program has been developed for catfish in recent years through selective breeding (Dunham 1984; Dunham 2006) and genomic resources (Liu 2011). The completeness of whole genome sequence in channel catfish provide possibilities for conducting genome comparative study with other teleost species.

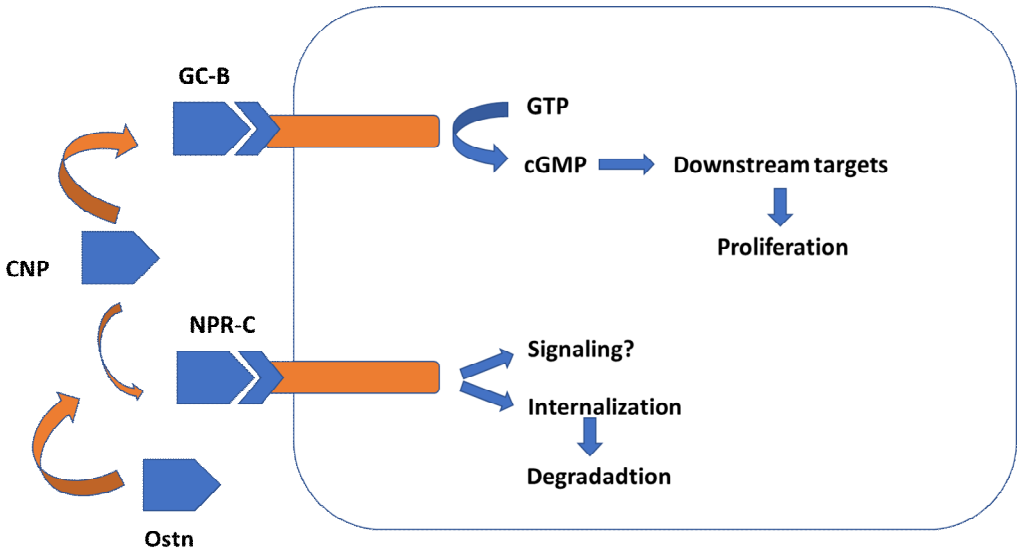
Osteocrin gene

Osteocrin gene (*ostn*), also described as musclin, is a skeletal muscle-derived factor (Liu 2008). Osteocrin gene generates 943 bp mRNA encoding 143 aa with molecular weight 14.8 kDa in zebrafish (http://www.ensembl.org/index.html). It is only homologous to two conserved sequence motifs reminiscent of dibasic serine protease cleavage sites (KKKR and KRR) found in peptide hormone precursors (Thomas 2003). Osteocrin also has NPs motif which is concluded to belong to NP family even though it lacks cysteine residues which conserved among the NP such as an atrial natriuretic peptide (ANP), brain natriuretic peptide (BNP), and type-C natriuretic peptide (CNP) (Potter, 2006; Chiba 2017). The *ostn* role as NPRC (NPR3) specific ligand which blockage the CNP access in chondrocytes (Moffatt 2007; Chiba, 2016). However, osteocrin has not been identified in the puffer fishes *Tetraodon nigroviridis* and *Fugu rubripes* (Moffatt 2007) and channel catfish (Zhang 2014).

Osteocrin is expressed in the skeletal tissue, especially in osteoblasts and young osteocytes at human bone forming sites (Bord 2005), suggesting its roles in bone formation (Moffatt 2007; Bord 2005; Thomas 2003). The *Ostn* was very recently found to be involved in

cranial osteogenesis and chondrogenesis in zebrafish (Chiba 2017). The peak expression of *ostn* is observed right after birth and decrease in older mice (Thomas 2003), again suggesting *ostn* as a marker of early osteoblast forming (Bord 2005). The presence of *ostn* seem leads the mineralizing process in the bone tissues (Lanctot, 2007). The overexpression of *ostn* resulting elongated bone by elevated the cyclic guanosine monophosphate (cGMP) (Thomas 2003). The administration of human *ostn* increased the rat bone mass (Thomas 2003). On the contrary, the loss function of *ostn* provides a possibility CNP to bind with NPRC resulting a decrease in the osteogenesis (Chiba 2017).

Figure 1. Ostn block the NPR-C clearance thus elevating the NPs in the bone. The NPs activities in the skeleton are determined by the distribution of CNP in the signal mediating NPR2 (GC-B) and the NPR3 (NPR-C) clearance receptor. The presence of ostn, block the CNP binding process to NPR3 which increase binding of CNP to the NPR2 resulting an increase of cGMP production, leads to downstream effect of NP and thus leads to bone development (derived from Thomas 2003)



In addition, the cyclic guanosine monophosphate (cGMP) has been reported in associated with pigment aggregation in freshwater shrimp (Ribeiro, 2009). The GMP works in recyprocal manner with cyclic aenosine monophosphate (cAMP), both are members of cyclic nucleotide (Hartzell, 1986; Thibault, 1996). The cAMP is known involved in calcified tissues, including bone (Tintut, 1998), fish skin (Oshima, 2001), and chromatophores (Kotz, 1994).

Fish Scale

Scale (Latin *squama*, Greek *lepid*), in zoology, is defined as a small, plate like tissue grow out of the skin which provides protection in animal (http://www.newworldencyclopedia.org/). These structures commonly emerge from dermal-epidermal integument site in fish, reptiles, birds, and several mammals (Chang, 2009; Chuong, 2000; Alibardi, 2004). Most of reptile, fish, and other vertebrates skin covered with scale or scutes, while amphibians have unique dermal scales which are flat within the dermis (Vickaryous, 2009). In mammals, for example, the mouse has a scale on the tail, and Malayan pangolin has scales on almost all their body surface provide a perfect protection (Gomez, 2013; Meyer, 2013). In birds, the scales emerge on their legs and has some similarity structure resembled in modern crocodile scales (Wu, 2004; Dhouailly 2009). The most probable and unique model for studies is the scale appeared in reptile and teleost (Takagi, 2007; Chang 2009). The study of scales included development, evolution, structure, regeneration in associated with teeth, feather, hair, skin, skeletal muscle, fin, gill, bone, and the cranial bone itself (Harris, 2008; Rohner, 2009; Dhouailly 2009; Chang 2009; Chiba 2017).

The fish scale arises from the integumentary system (Vickaryous, 2009). Like other vertebrates, the fish integument consisting two main part, the inner dermis and the outer epidermis layer (Elliott, 2011). The epidermis consists of several layers of epithelium generated

from ectoderm part of the embryo (Tingaud-Sequeira, 2006; Xu, 2013). The cellular element of the most fish epidermis is the epithelial cell (Xu, 2013). In another side, the dermis is located at the deep inner layer of integument and generated from embryonic mesenchyme of mesodermal (Elliott, 2011). The dermis consists of fibrous tissues made from collagen protein which is also resembled by nerves, pigment cells, and adipose tissue (Vickaryous, 2009). The scales are growing initially from dermis layer (Sire, 1997).

Fish scale creates a significant part of the dermal skeleton of teleost fishes as it is the most observable tissue morphologically along the skin with several types (Gil-Duran, 2016). Fish scale has several functions, such as provide protection for them especially on dermal layer by providing hard layer (Quilhac, 1999) or by adjusting the coloration; a barrier between the species (Vieira, 2011), the environment, and against pathogen (Chai, 2010); moisture retention; movement agility; a source of minerals and nutrient (Palmer, 2008; Scharer, 2012; Niu, 2014; Sire 2003), and provide insight into dental-teeth evolution (Sasagawa, 2013). The fish scale has many types, four at least, such as placoid, ganoid, leptoid, and cosmoid which are classified by Louis Agassiz (Wake, 1992). Leptoid also was known as elasmoid scale included cycloid and ctenoid scale types (Elliott, 2011). Each of fish scale type has its unique characteristic. The placoid, also known as odontodes, is the scale belong to cartilaginous fish such as rays, chimeras, and shark (Brazeau, 2014; Lim, 2015). The placoid scale in the shark is physically homologous to the vertebrate denticle which also constructed by central pulp cavity, a structure of living connective tissue supplied by the blood vessels, cover by dentin dermal layer and transparent enameloid (Sasagawa, 2013). The ganoid scale is the ancestor of the most modern bony fish which still could observe in some primitive ray-finned fish (Elliott, 2011). The ganoid scales which was shown in the spotted gar (*Lepisosteus oculatus*), based on genetic and fossil evidence,

were homologous with the enamel in mammalians and others vertebrate (Qu, 2015). The elasmoid scale is found in most of the modern teleost, the bony fish (Vickaryous 2009). These scales are a lack of dentin and ganoin. Elasmoid scales grow continuously throughout the fish life indicated by circuli mark extent in distance and number (Elliott 2011).

Identification of genes involved in scale development and regeneration through genomics analysis will be useful for understanding the scale evolution in teleost (Costa, 2017). The fish species become a valuable model to provide insight of scale evolution and development (evo-devo) of skeletal elements including scale (Liu, 2016). Fish scale removal includes the loss of the skin, epidermal cells, and scales (Vieira, 2011). The skin wounds in fish quickly will cover by mucus and re-epithelization from the wound margin follow within a few hours (Quilhac, 1999; Iger, 1990). Furthermore, in a few weeks a new mature scale is completely re-formed (Ohira, 2007; Bereiter-Hahn, 1993). This regeneration has been separated into four steps; beginning with the differentiation of scale-forming cells (2 days), continued by a formation of external layer matrix (3 days), the formation of the basal-plate matrix (6-14 days) and lastly mineralization of the plate (14-28 days) (Ohira, 2007). Most of the experiments on fish scale development and regeneration have done on morphology (Le Guellec, 2004; Sire, 2004) or focused on single gene studies. For example, some studies have revealed the function of *esr2a*, *apoeb*, *edar*, *fgfr*, *Tcf7,* and *scpp* in associated with scale formation and evolution (Table 1). In zebrafish and most of the teleost, the dermal skeletal including scale is formed in late post-embryonic periods (Harris, 2008).

Table 1. Genes involved in scale evolution, development, and regeneration in teleost.

Name of gene	Gene symbol	Fish model	Main Biological effect	Reference
sonic hedgehog	*shh*	Zebrafish	Involved in scale morphogenesis and differentiation (epidermal fold in posterior region)	(Sire 2004)
estrogen receptor 2a	*esr2a*	Zebrafish Sea bream	Involved in fin and scale development and regeneration	(Tingaud-Sequeira 2006)
apolipoprotein Eb	*apoeb*	Zebrafish	Involved in fin and scale development	(Vieira 2011)
ectodysplasin-A-(receptor)	*eda* *edar*	Zebrafish	Effect on fin, scale, teeth, and gill rackers	(Kondo, 2001; Harris, 2008)
fibroblast growth factor receptor	*fgfr*	Zebrafish Mirror carp	Involved in partial scale lost	(Rohner, 2009)
calcium binding phosphoproteins	*scpp*	common carp	Involved significantly in scale regeneration	(Liu 2016)

Genome editing CRISPR/Cas9

Latest advances in the targeted alteration of diverse eukaryotic genomes have revealed a new era of genome editing (Hsu, 2014). Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system has become efficiently targeted mutagenesis method in zebrafish. CRISPR/Cas9 has some advantages compared to TALENs and ZFNs methods in generating stable mutant, such as easy to use, simple design process and could target multiple genes (Varshney, 2015). Resulted from parts of a bacterial immune system, the CRISPR-Cas9 system allows targeted gene division and gene editing in a diversity of eukaryotic cells, and because the endonuclease division specificity in CRISPR-Cas9 system is directed by RNA

sequences, editing can be guided to nearly any genomic locus by modifying the guide RNA sequence and directing it besides with the Cas endonuclease to the target cell (Doudna 2014). The CRISPR-Cas9 has huge future in widespread practices such as gene therapy, stem cell, plant and animal disease models, and disease-resistant transgenic (Sander, 2014). The CRISPR-Cas9 system is formed of a short non-coding guide RNA (gRNAs) that has two molecular elements: a target-specific CRISPR RNA (crRNA) and a trans-activating crRNA (tracrRNA) (D'Agostino 2016; Hwang, 2013). The gRNA unit directs the Cas9 protein to a precise genomic locus via base pairing between the crRNA sequence and the target sequence. In bacteria, CRISPR loci are formed of a series of repeats divided by parts of exogenous DNA in ~30 bp in length, called spacers (Irion 2014; Moreno-Mateos 2015). The repeat-spacer array is transcribed as a long precursor and processed within repeat sequences to derive small crRNAs that denote the target sequences (known as protospacers) division by Cas9 protein, the nuclease element of CRISPR system (Hsu 2014; Paix, 2015). CRISPR spacers are later used to identify and silence exogenous genetic parts at the DNA level (Mali, 2013). Crucial for the division is a three-nucleotide sequence motif directly downstream on the 3' end of the target region which is known as the protospacer-adjacent motif (PAM) (Pyne, 2016; Liang, 2016). The PAM is appeared in the target DNA, but not the crRNA that targets it (Liu 2016). Furthermore, after binding to the target sequence, the Cas9 protein causes a specific double-strand break (Ding, 2016). Following DNA cleavage, the break is fixed by cellular repair machinery via non-homologous end joining (NHEJ) or homology-directed repair (HDR) mechanisms (Maruyama, 2015; Vartak, 2015). With target specificity determined by a very short RNA-coding region, the CRISPR-Cas9 system significantly facilitates genome editing (D'Agostino, 2016; Moreno-Mateos 2015; Hwang 2013).

The zebrafish is an ideal model for functional genomics analysis (Meeker 2007). The strength point of the zebrafish model lies in its evolutionary association to mammalians, such as human and mice, and invertebrate model system (Lieschke, 2007). The zebrafish also carries many advantages over the mouse, due to its non-mammalian features (Ota, 2014; Yu, 2014). Because zebrafish reproduce by external fertilization, all levels of zebrafish embryogenesis are possible to the researcher for the study. This advantage combined with the optical transparency of the zebrafish allows for real-time observations of reviewed process by fluorescent reporters with tremendous ease (Ramlee, 2015). These observational qualities, in combination with rapid development, relative size, and fecundity of zebrafish compare to mice, enables low maintenance and rearing infrastructure expenses (Hwang, 2013; Irion, 2014). In teleost, zebrafish is a golden standard in molecular studies (Auer, 2014; Liu, 2016)

III. METHODS & MATERIALS

Gene identification and sequence analysis

The *ostn* and neighborhood genes were searched against the transcriptome and whole genome sequence of channel catfish (Liu, 2012; Liu, 2016); and genome from common pleco, lamprey, blue catfish and walking catfish using TBLASTN. The amino acid sequences from several vertebrate species in amphibian (frog), bird (chicken), fish (zebrafish, tilapia), and mammals (human, mouse) were used as query sequences. The *E-value* was set at an intermediately stringent level of $1e^{-5}$ for collecting as many as possible *ostn*. The sequences were aligned by Clustal Omega (Sievers, 2011) to remove duplicates and generate a unique set of sequences. The open reading frame was determined using ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/) and verified by BLASTP against NCBI non-redundant protein sequence database. BLASTN was conducted against a reference genome to verify sequence accuracy. The genomic scaffold containing the *ostn* were retrieved, and the genes within genomic scaffolds were predicted by AUGUSTUS (Stanke, 2005) and FGENESH of Molquest software from Softberry Int. (Solovyev, 2006). Functional domains were determined using Interpro Scan (Finn, 2017).

Phylogenetic analysis

The amino acid sequences of *ostn* from other vertebrate organisms were retrieved from Ensembl v87 (http:www.ensemble.org/), Uniprot (http:www.uniprot.org/), or NCBI database (http:www.ncbi.nlm.mih.gov/) for phylogenetic analysis. Multiple protein sequences alignments were conducted using the MAFFT (Katoh 2013). The best-fit model of Osteocrin evolution was determined using the ProtTest program according to the Bayesian information criterion. The

best-fit model was the JTT+G+I model, which used JTT matrix (Jones, 1992) and the gamma distribution with invariant sites for modeling rate heterogenity (+G+I). The phylogenetic and molecular evolutionary analyses were conducted using MEGA 6. using parameters Bootstrap test of 1000 replicates and partial deletion method, from which the Maximum Likelihood trees were generated. Phylogenetic analyses were constructed among other representative vertebrates species including mammalians (human, orangutan, pig, panda, cow, sheep, dolphin, mouse, opossum), chicken, frog, turtle, and fish (spotted gar, coelacanth, Atlantic salmon, northern pike, herring, cave fish, pleco, shark, carp, zebrafish, amazon molly, shortfin molly, guppy, mummichog, minnow, yellow croaker, damselfish, tilapia, nyererei, zebra mbuna, mouthbrooder).

Syntenic analysis

The syntenic analysis was conducted by analyzing most conserved regions harboring *ostn* from several vertebrates based on genome information retrieved from Ensembl and NCBI, to provide additional evidence for gene identification and orthology. Some vertebrates *ostn* sequences were used as queries to search against the draft catfish genome sequences to obtain the genomic scaffolds containing *ostn* from several species. If the *ostn* gene was not found, the similar method was used to search the most conserved gene beside *ostn* to described the syntenic in those species. The neighboring genes were identified by a Fgenesh software program and BLASTP as mentioned above. The orders of these neighboring genes were compared with those from zebrafish and human using the software Genomicus (Louis, 2015).

Expression analysis using available RNA-seq datasets

The meta-analysis was conducted to quantify the *ostn* expression among wild type zebrafish tissues using RNA-seq datasets retrieved from NCBI: caudal fin (SRR5125774), sperm

(SRR4449583), gill (SRR1609738), kidney (SRR1609747), liver (SRR1609750), muscle (SRR1609753), spleen (SRR1609756), hearth (SRR1609741), brain (SRR1609735), intestine (SRR1609744), testis (SRR1265758), eye (SRR1265761), ovary (SRR1265755), blood (SRR891512), pharyngeal arches (SRR5320512), bone (SRR3109807), and skin (ERR1294271). The gene expression analysis was performed using CLC Genomics Workbench Software (V5.5.2). The trimmed high-quality reads will be first mapped onto zebrafish osteocrin transcript sequence. The total mapped reads number for the transcript will be determined, and the read per kilobase per million (RPKM) will be calculated

The meta-analysis was also conducted based on the previous study on common carp RNA-seq dataset during skin regeneration (SRX1201398). The complete procedures were described in the previous study (Liu 2016). The trimmed high-quality reads will be first mapped onto common carp osteocrin transcript sequence. Mapping parameters will set as ≥ 95% of the reads in perfect alignment and ≤ 2 mismatches. The total mapped reads number for the transcript will be determined, and the fold change will be calculated. Transcript with absolute expression fold change value ≥ 1.5, *P-value* < 0.05, and total gene reads ≥ 5 will be concluded in the analysis as significantly differentially expressed genes.

qRT PCR

To quantify *ostn* expression in some calcified-related tissues in zebrafish, the qPCR was performed using SYBR Green PCR Master Mix on a CFX96 RT-PCR detection system (Bio-Rad, USA). The cycling conditions of qPCR are denaturation, 95 ºC/30 s, 40 cycles of 95 ºC /5 s, 60 ºC /5 s, and 72 ºC /5 s. The Elongation factor 1 alpha (EF-1α) gene was used as an internal reference gene. All gene-specific primers were designed by Primer3 software (Rozen, 2000). The specific primer sequences for *ostn* were F) 5’ ATTGACCGGATTGGTGTAGG and R) 5’

GTTGGCTGGTGTGCTTATCA. The primers for reference gene (EF-1α) were F) 5' CTTCTCAGGCTGACTGTGC and R) 5' CCGCTAGCATTACCCTCC. The specificity of primers was assessed by aligning with the zebrafish reference genome sequence using BLASTN with E-value of 1e-10. The melting curve analysis was performed to demonstrate that each primer set amplified a single product. The expression levels of each gene were expressed relative to those of EF-1α in each sample using the $2^{-\Delta\Delta CT}$ method (Livak, 2001). Tissue with lowest Ct value of each gene was used as the reference for normalization of the relative expression of *ostn* genes such that all the values are greater than 1. The data of real-time PCR were expressed as the standard error of the mean (SEM). The qPCR was repeated in triplicate runs to confirm expression patterns.

CRISPR/Cas9 preparation

To generate sgRNA, free cloning method was used in the experiment (Varshney 2015). Three small guide RNAs were generated from CRISPR target sequence in *ostn* within 485 bp coding sequence. The sgRNAs were designed in exon 3 using UCSC Genome Browser (http://genome.ucsc.edu) (Figure S3). Thus, T7 promotor was added to 5' and 20-nt crRNA-TracrRNA overlapping sequence to 3' as oligo 1. The design was as follow: 5′-TAATACGACTCACTATA-GGN (18-20nt target sequence)-GTTTTAGAGCTAGAAATAGC-3′. I assembled oligo 1 with oligo 2 (sgRNA core sequence: (5′-AAAAGCACCGACTCGG TGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAAC-3′) using Accuzyme polymerase (Bioline) in following conditions: 98ºC/3 min, 98ºC/15 s, 58ºC/15 s, 72ºC/1 m for 5 cycles, and 72ºC/5 m. The quality of sgRNAs was checked by running 1 µl on 2.5% gel agarose gel. In vivo transcription was conducted to generate sgRNA from assembled oligo as a template using the HiScribe T7 High Yield RNA Synthesis Kit (New

England BioLabs). The obtained sgRNAs were purified passage through Microspin S-200 HR Column (GE Healthcare). For Cas9 mRNA, the zebrafish codon was used to optimize cas9 plasmid (pCS-nCas9n) as a template (Jao, 2013). The template was linearized by Xbal then purified using Zymo RNA Clean & Concentrator-5 kit (Zymo Research, CA). Total 500-1000 ng linearized and purified template was used to synthesized RNA using mMESSAGE mMACHINE (sp6) kit (Life Technologies) and precipitated using PCI. For microinjection preparation, the phenol red was added to gRNA/Cas9 to one-third volume to color the solution. Approximately 2 nL final volume of mix sgRNAs and Cas9 which were used for the co-injected into one cell stage embryos at concentration: 30 pg of sgRNA with 150 pg Cas9.

Zebrafish rearing

Zebrafish embryos were maintained in optimum condition (Westerfield 2007) to generate F0 fish. All experiments procedures used in this study were approved by Auburn University Institutional Animal Care and Use Committee (AU-IACUC). The embryo and larvae were kept in EM3 solution during the early stage (15 dpf) then transfer to close recirculation system (28^0C). MS-222 solution (100 ppm) was used to anesthetize the fish during observation and sampling.

DNA extraction

The muscle was examined as a tissue sample for DNA extraction, amplification, and sequencing as *ostn* known to be expressed as skeletal muscle active molecule. The muscle was collected from 3 month's adult fish in liquid nitrogen. The DNA was extracted using the protocol as followed, the muscle digested in cell lysis buffer and K proteinase for overnight. The DNA pellets were purified using 100% isopropanol followed by ethanol 70%. The concentration of the DNA was checked using Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific).

Fluorescence PCR and endo T7 I assay detection

To detect all possible mutation site in the target region, I design primers for PCR: F) 5' TTTGTGGGTTTTTGACTAATTTGA and R) 5' TTCCAAATTAGAAGAACAGCATTT. The PCR was performed using Hi Fidelity Plus PCR System (Roche). The cycling process was included denaturation at 98ºC/30 s; followed by 35 cycles at 98ºC/30 s, 60ºC/10 s; 72ºC/20 s, and a final extension at 72ºC/2 min. The PCR result was confirmed in 1% agarose gel. The PCR reactions were purified using Zymo DNA Clean and Concentrator (Zymo Research, Ca) according to the manufacturer guidelines. Thus, I used endo T7 I assay detection Kit (New England Biolabs) to detect the mutation. First, I assembled 50 µl reactions as follow: 19 µl annealing reaction, 2 µl 10X NEBuffer, and nuclease free water. The PCR result were denatured and re-annealed through process: 95ºC/5 min; 95-85ºC at -2ºC/s, 85-25ºC at -0.1ºC/s, and cooling at 4ºC. The last, the T7 endonuclease I was adding to recognized and cleaved (digested) the mismatch sequence. The digested products were confirmed in 2% agarose gel and compared to the Wt control for analysis.

Cloning and sequence analysis

The mutation was identified by amplifying the DNA from positive mutated individuals. The amplicons were cloned using TOPO® TA Cloning® Kit (Invitrogen) and then sequenced it. Thus, the sequences were aligned to the wild type and analyzed to describe the indels among the mutated region.

Scale stain and Measurement

Adult zebrafish scales and bone were stained using alizarin red. The fish were immersed in formaldehyde neutral buffer 3.7% for 24 hours and briefly dehydrate in ethanol 70%. Thus, the fish stained in alizarin red 0.1% and potassium hydroxide (KOH) 1% overnight in room

temperature (Rohner 2009). The fish then transfer into 1% trypsin overnight to make bone visible. I used the mix of 1.5% hydrogen peroxide and 1% potassium hydroxide in equal volume to bleach the fish pigment for 10 minutes. The stained fish then immerse in a mix of 20% glycerol and 1% potassium hydroxide for an overnight before observation. The fish and scale measurement were using a dissecting microscope and ImageJ 1.x software (Schneider 2012). The total fish length was measured from the tip of the head to the tip of the fish tail.

V. RESULT

The *ostn* gene is present in scaled fishes but absent in scaleless fishes

The initial observations of osteocrin in armor-scaled catfish but not in channel catfish (scaleless) suggested its potential involvement in scale formation. This observation is in line with the known functions of osteocrin in the development of calcified tissues such as bone. To establish the correlation of the loss of osteocrin gene with the loss of scales, I first determined the status of osteocrin in the genomes of 34 vertebrate species including representatives of mammals (10), amphibian, bird, reptile, and fishes (21). As shown in Table 1, the osteocrin gene was found from the 21 scaled fish species, but was found missing from scaleless fish species including channel catfish, blue catfish, walking catfish, green-spotted pufferfish, and stickleback. One possibility for the observed absence of the osteocrin gene in the genome sequences is that the osteocrin gene may exist but may have been missed in genomic sequencing. This possibility may be difficult to be excluded with mega data sets of whole genomic sequences from many species. I then conducted syntenic analysis in order to provide stronger evidence that osteocrin gene was indeed lost in these scaleless fish rather than it was not there because it was missed during genomic sequencing.

Validation of osteocrin loss through syntenic analysis

To validate the absence of the osteocrin gene in scaleless fishes, the syntenic analysis was conducted. The conserved syntenic blocks were first identified, and then the genes surrounding the osteocrin gene were determined. As shown in Figure 2, the conserved syntenies were observed from fish to mammals. The genes surrounding the osteocrin gene such as MASP1 and GMNC genes are apparently conserved among many species, although the exact gene order and

orientation differed in some cases. Osteocrin gene was found in the conserved syntenic region of all scaled fish except doras for which only draft genomic sequences were available. There are two possibilities in may explain the missing of Osteocrin in doras. First, it is likely that osteocrin was simply missed during sequencing in this species. Second, the doras actually derived from the family loricarioidei which is similar to the scaleless channel catfish, while the pleco derived from family silurioidei suggesting a possibility of different armored scales were derived to pleco and doras. However, the osteocrin gene was not found in the conserved syntenic region for all scaleless fish species including channel catfish, blue catfish, walking catfish, green-spotted pufferfish, and stickleback. The syntenic analysis provided additional support for the correlation between the presence of the osteocrin gene in the genome with the presence of scales. Based on copy number analysis among several vertebrate species, *ostn* gene was present mostly as a single-copy gene, with the exception in common carp where there are two copies of *ostn*. Only one copy of *ostn* was found in higher vertebrates ranging from amphibians, reptiles, birds, and mammals (Table 2).

Table 2. Characteristics of Osteocrin genes identified in several representative vertebrate genomes: Species, mRNA size, amino acid sequence length, musclin domain position, copy number, and scale status among the teleost species.

Species	mRNA size (bp)	AA	Position of Musclin domain	Copy	Scale status
Human	3231	133	1-133	1	N/A
Mouse	1268	130	1-130	1	N/A
Opossum	745	133	1-133	1	N/A
Cow	1464	132	1-132	1	N/A
Sheep	1799	132	1-132	1	N/A
Dolphin	1639	132	1-132	1	N/A
Sunda pangolin	950	134	34-134	1	N/A
Orangutan	3145	111	1-111	1	N/A
Panda	450	132	1-132	1	N/A

Guinea pig	757	134	4-134	1	N/A
Turtle	402	133	1-133	1	N/A
Chicken	543	139	1-133	1	N/A
Frog	531	176	44-176	1	N/A
Amazon molly	1047	134	54-134	1	Present
Shortfin molly	1026	134	54-134	1	Present
Guppy	1758	136	51-136	1	Present
Minnow	672	105	22-105	1	Present
Mummichog	908	251	174-251	1	Present
Damselfish	1756	132	38-132	1	Present
Yellow croaker	801	134	6-134	1	Present
Tilapia	2343	139	10-139	1	Present
Zebra Mbuna	1878	139	11-139	1	Present
Moutbrooder	1093	139	11-139	1	Present
Nyererei	1102	138	12-138	1	Present
Atlantic Salmon	919	134	1-134	1	Present
Northern Pike	324	107	1-107	1	Present
Atlantic Herring	309	102	11-102	1	Present
Zebrafish	957	134	1-134	1	Present
Cave fish	1862	135	1-135	1	Present
Common carp	889/898	149/134	51-149/1-134	2	Present
Pleco	468	155	44-145	1	BDP
Elephant shark	902	143	64-143	1	Present
Coelacanth	676	129	1-129	1	Present
Spottted Gar	2321	133	1-133	1	Present

Conservation of *ostn*, its musclin domains and scale types

The characteristics of the osteocrin gene in various species are summarized in Table 2 including the mRNA size, the amino acids it encodes, the position of the musclin domain, copy numbers of the gene, and the status of scales. The length of *ostn* transcripts varied between 309 and 3,231 bp encoding 102-251 amino acids, and the exon number ranged from 3 and 7. The sequence similarities are low among many osteocrin genes from various species, but two regions homologous to the natriuretic peptide-like motifs were well conserved across a broad spectrum

of species except in pleco which only has 1 NP-like motif (Fig.S1). This is noteworthy because only pleco has a scale type that is bony dermal plates, not regular scales.

The osteocrin amino acid sequences were used for phylogenetic analysis (Fig. 3). Very interestingly, the clades of the phylogenetic tree built using osteocrin amino acid sequences reflected the types of scales very well. At least, four fish scale types including leptoid, placoid, ganoid and cosmoid as classified by Louis Agassiz (Wake 1992) and one fishes with a bony dermal plate were identified based on their groups. Scale types of cycloid and ctenoid were both derived from the same leptoid (elasmoid) type (Fig. 3).

Expression of osteocrin gene during scale regeneration

Osteocrin was not expressed in most of non-calcified tissues such as liver, spleen, kidney, sperm, ovary, testis, intestine, eye, and blood except for muscle and brain (Fig. 4A), as analyzed using a meta-analysis of RNA-seq data sets. In calcified tissues, as RNA-Seq datasets were not available for some part, I also analyze the expression level of *ostn* using qRT-PCR among calcified tissues in zebrafish including skin, scale, bone, teeth, dorsal fin, ventral fin and caudal fin. The expression of osteocrin was much higher in calcified tissues, of the highest in the bone, followed by ventral fin, scale, teeth, and caudal fin (Fig. 4B).

Figure 2. Syntenic analysis of *ostn* gene in various species. Conserved syntenic regions were retrieved from NCBI database and Genomicus. The GMNC, MASP1, and FGF12 genes were used as anchor genes for the identification of conserved syntenic regions. Note the absence of *ostn* gene in the genomic neighborhood of scaleless fish such as channel catfish, blue catfish, and walking catfish, green-spotted pufferfish and stickleback. Gene symbols are abbreviated as the following: GMNC, geminin coiled-coil domain containing; MASP1, mannan-binding lectin serine peptidase 1; FGF12, fibroblast growth factor 12.

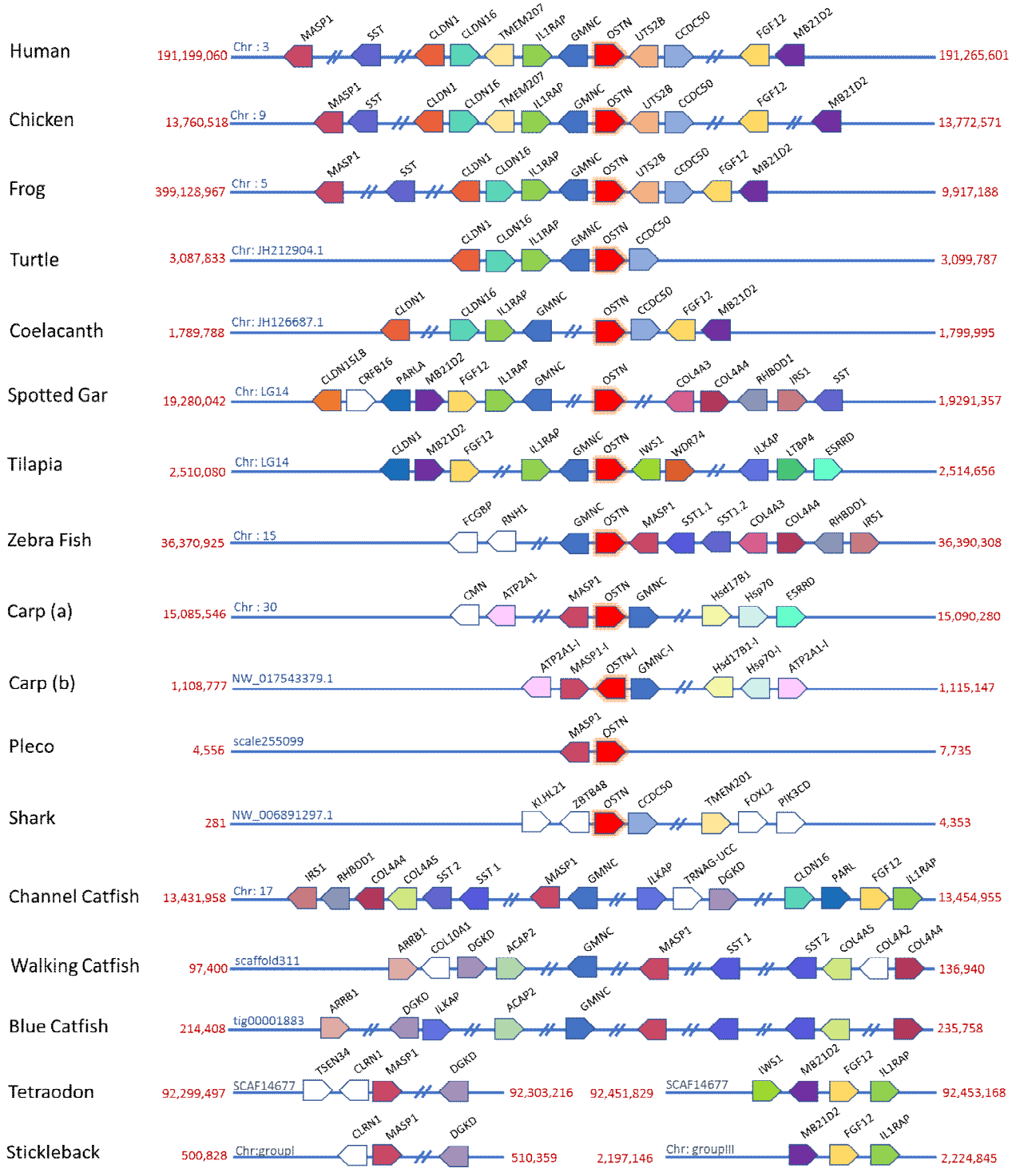

Human
191,199,060
Chr : 3
MASP1 SST CLDN1 CLDN16 TMEM207 IL1RAP GMNC OSTN UTS2B CCDC50 FGF12 MB21D2
191,265,601
Chicken
13,760,518
Chr : 9
MASP1 SST CLDN1 CLDN16 TMEM207 IL1RAP GMNC OSTN UTS2B CCDC50 FGF12 MB21D2
13,772,571
Frog
399,128,967
Chr : 5
MASP1 SST CLDN1 CLDN16 IL1RAP GMNC OSTN UTS2B CCDC50 FGF12 MB21D2
9,917,188
Turtle
3,087,833
Chr: JH212904.1
CLDN1 CLDN16 IL1RAP GMNC OSTN CCDC50
3,099,787
Coelacanth
1,789,788
Chr: JH126687.1
CLDN1 CLDN16 IL1RAP GMNC OSTN CCDC50 FGF12 MB21D2
1,799,995
Spotted Gar
19,280,042
Chr: LG14
CLDN15L8 CRFB16 PARLA MB21D2 FGF12 IL1RAP GMNC OSTN COL4A3 COL4A4 RHBDD1 IRS1 SST
1,9291,357
Tilapia
2,510,080
Chr: LG14
CLDN1 MB21D2 FGF12 IL1RAP GMNC OSTN IWS1 WDR74 ILKAP LTBP4 ESRRD
2,514,656
Zebra Fish
36,370,925
Chr : 15
FCGBP RNH1 GMNC OSTN MASP1 SST1.1 SST1.2 COL4A3 COL4A4 RHBDD1 IRS1
36,390,308
Carp (a)
15,085,546
Chr : 30
CMN ATP2A1 MASP1 OSTN GMNC Hsd17B1 Hsp70 ESRRD
15,090,280
Carp (b)
1,108,777
NW_017543379.1
ATP2A1-l MASP1-l OSTN-l GMNC-l Hsd17B1-l Hsp70-l ATP2A1-l
1,115,147
Pleco
4,556
scale255099
MASP1 OSTN
7,735
Shark
281
NW_006891297.1
KLHL21 ZBTB48 OSTN CCDC50 TMEM201 FOXL2 PIK3CD
4,353
Channel Catfish
13,431,958
Chr: 17
IRS1 RHBDD1 COL4A4 COL4A5 SST 2 SST 1 MASP1 GMNC ILKAP TRNAG-UCC DGKD CLDN16 PARL FGF12 IL1RAP
13,454,955
Walking Catfish
97,400
scaffold311
ARRB1 COL10A1 DGKD ACAP2 GMNC MASP1 SST 1 SST 2 COL4A5 COL4A2 COL4A4
136,940
Blue Catfish
214,408
tig00001883
ARRB1 DGKD ILKAP ACAP2 GMNC
235,758
Tetraodon
92,299,497
SCAF14677
TSEN34 CLRN1 MASP1 DGKD
92,303,216
92,451,829
SCAF14677
IWS1 MB21D2 FGF12 IL1RAP
92,453,168
Stickleback
500,828
Chr:groupI
CLRN1 MASP1 DGKD
510,359
2,197,146
Chr: groupIII
MB21D2 FGF12 IL1RAP
2,224,845

Figure 3. Phylogenetic analysis using ostn amino acid sequences of fish with various types of scales, as well as representative species of amphibians, reptiles, birds, and mammals. The evolutionary history was constructed using the Maximum Likelihood method based on the JTT matrix-based model. The Gamma distribution was used to model evolutionary rate differences among sites (+*G*, parameter). The rate variation model allowed for some sites to be evolutionarily invariable (+*I*). The analysis involved 34 amino acid sequences. Evolutionary analyses were conducted in MEGA6.

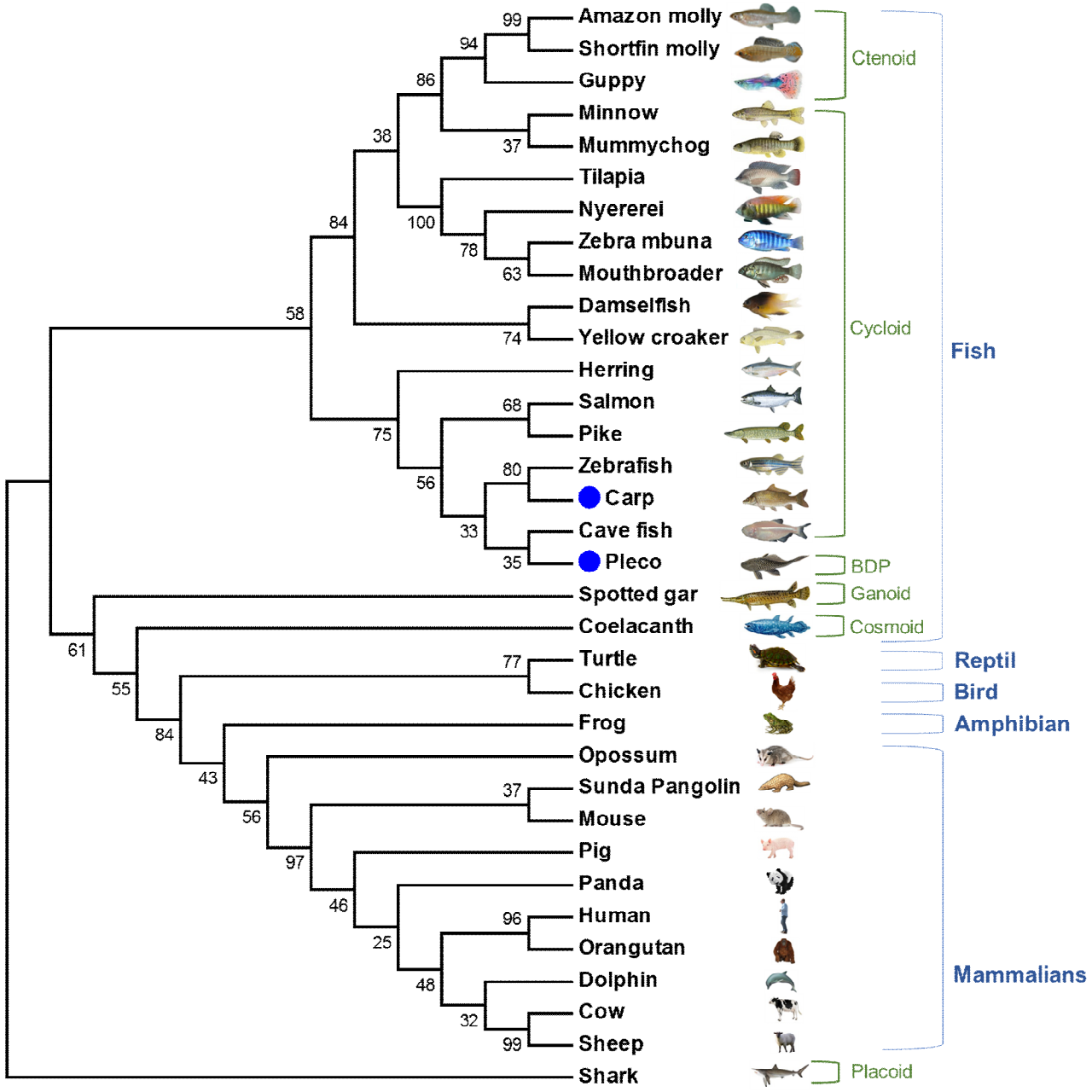
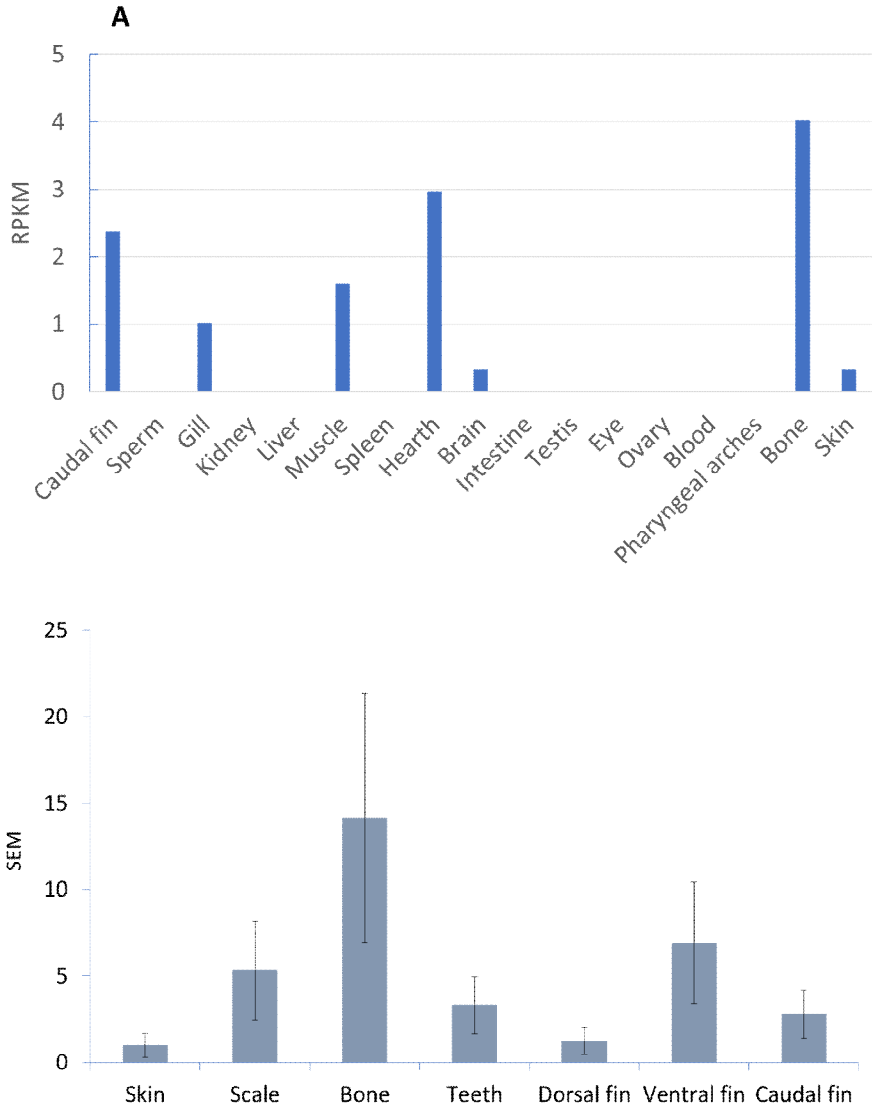
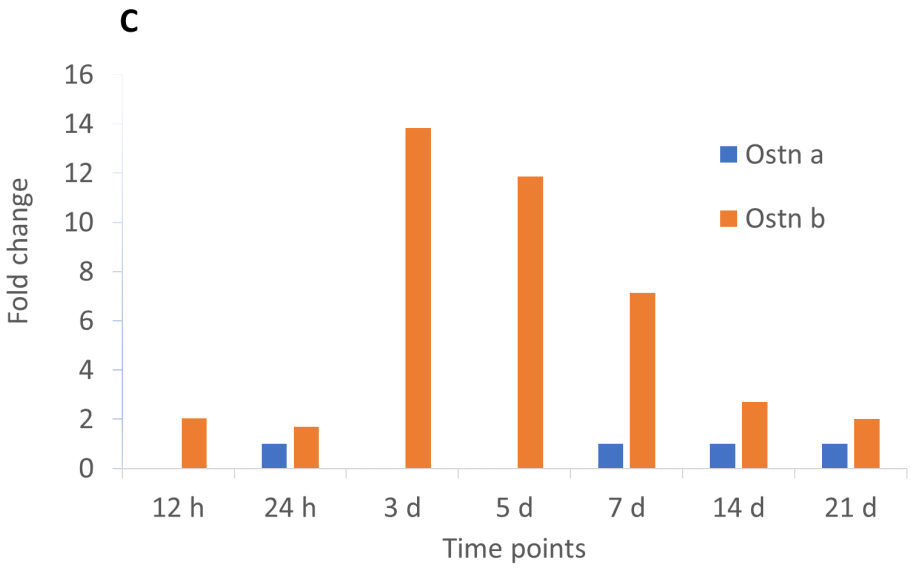


Figure 4. The expression level of *ostn*. The expression level of *ostn* was quantified in several tissues of Wt zebrafish using CLC Genomics Workbench Software (V5.5.2) retrieved from NCBI database and expressed as RPKM (**A**). The expression level of Osteocrin in several calcified zebrafish tissues**.** The analysis using qRT-PCR and expressed as SEM (Standard Expression of The Mean) (**B**). The expression level of Osteocrin during the scale regeneration experiment in common carp**.** The expression level of *ostn* was quantified at eight (8) time points using CLC Genomics Workbench Software (V5.5.2) from RNA dataset and expressed as fold change. The *ostn*-a expressed in random pattern, while *ostn*-b differentially expressed suggesting involvement in scale formation **(C)**





As RNA-Seq data sets existed from carp scale regeneration (Liu 2016), I also conducted a meta-analysis of osteocrin expression using the RNA-Seq datasets. Common carp have two copies of *ostn*, *ostn*-a (149 aa) and *ostn*-b (134 aa). The expression of ostn-a was relatively stable over the course of carp scale regeneration, whereas the expression of ostn-b was significantly up at day 3, day 5, and day 7, up 7-14 folds (Fig. 4C). This suggested that ostn-a may not be seriously involved in the regulation of scale formation, whereas ostn-b appeared to be more directly involved in scale formation.

Phenotypes of *ostn* knockout in zebrafish

Zebrafish (*Danio rerio*) genome has one copy *ostn* span from 36,370,925 - 36,390,308 on chromosome 15. The *ostn* consists of 5 exons and four introns with 957 bp nucleotide sequence encoding 134 amino acids. (Table. 1). We designed three gRNA within 485 bp coding sequence targeting three different positions followed by 'NGG' PAM (Photo Adjacent Motive) sequence (Fig. S3). The gRNA along with mRNA Cas9 were used to inject 200 fries of first stage

embryos. The phenotypes were observed through the entire course of the experiment after injection. Knockout of *ostn* resulted in a number of phenotype defects, including slow scale growth (3 fish), partial scale lost (3 fish), bone deformities (14 fish), imperfect operculum, asymmetrical gill rakers, and pigment dispersion (66 fish) on the heterozygous mutant (*ostn* $^{-/}$). Although osteocrin knockout was not lethal, its mutation had a major effect on the survival of knockout fish. The survival rate of the CRISPR/Cas9 treated group during the knockout experiment was 61% (122 fries) at the end of the experiment, compared to 95% survival rate at the end with the wild type (Wt) group.

Observations were made with both Wt and *ostn*$^{-/}$ at post-embryonic/larval period (<30 dpf/ <11 mm TL) and adult stage (>30 dpf/ >11 mm TL) (Kimmel, 1995; Parichy, 2009). We focused on post-embryonic stages as *ostn* had already been reported as a bone regulator in embryonic stages (Thomas 2003; Moffatt 2007). In the beginning, we observed the post-embryonic periods and found a curved defect on the lateral bone compared to the Wt (7 dpf/ 6 mm TL) (Fig. 5). Observation at 30 dpf (11 mm TL), the Wt fish showed normal scale growth exhibiting melanophore aggregation surrounding the dermal pocket. On the contrary, the mutated *ostn*$^{-/}$ (11 mm TL) showed no signal of first scale forming with no pigments aggregations along with dermal pocket forming suggesting a slower scale growth (Fig. 6). The slow scale growth could be observed as a result of the slow fish growth (in mm TL) itself or because of chromatophores dispersion, so later at 50 dpf (18 mm TL), we observed and found the mutated *ostn*$^{-/}$ fish having less of scale growing on their dorsal compared to 16 mm Wt (Fig. 6). Some of the mutated *ostn*$^{-/}$ also exhibited partial scale lost (Fig. 7) and thin or fragile dermal skeleton (Video S4).

The mutated *ostn*$^{-/}$ fish exhibited lateral bone defects in the adult stage, as well as at early post-embryonic stage (6 mm fry). A number of adult fish had mild and severe lateral bone (spinal cord) defects with a single or double curve at 19-26th of bone segments, counting from the anterior to the posterior site along with the lateral line (Fig. 8). The color of lateral bone in mutated *ostn*$^{-/}$ was pale reddish compare to the Wt fish (Fig. 8). Apparently, the epipleural and epineurals part of the intermuscular bone in *ostn*$^{-/}$ were shorter compare to the Wt. In addition to these severe bone defects, *ostn*$^{-/}$ adult fish also exhibited several mild defects, such as incomplete plate covering the operculum and abnormal gill rakers (Fig. 7), and mild defects on chromatophores (melanophores, xanthophore, iridophore) dispersion (Fig. 9). At the end observation, we also phenotypically screened both Wt fish and whole 122 heterozygous F0 *ostn*$^{-/}$ adult fish (70 dpf/ 28-32 mm TL).

Figure 5. The effect of *ostn* knockout on lateral bone development. Compared to the wild-type (left panel A-C), the *ostn* knockout fish were smaller (5 mm vs. 6 mm at 7 dpf), and carried deformities on the lateral bone (blue arrow, right panel). A and D, lateral view of the spinal cord; B and E, the magnification of A and D, C and F, ventral view of the wild-type fish (C) and the mutant (F).

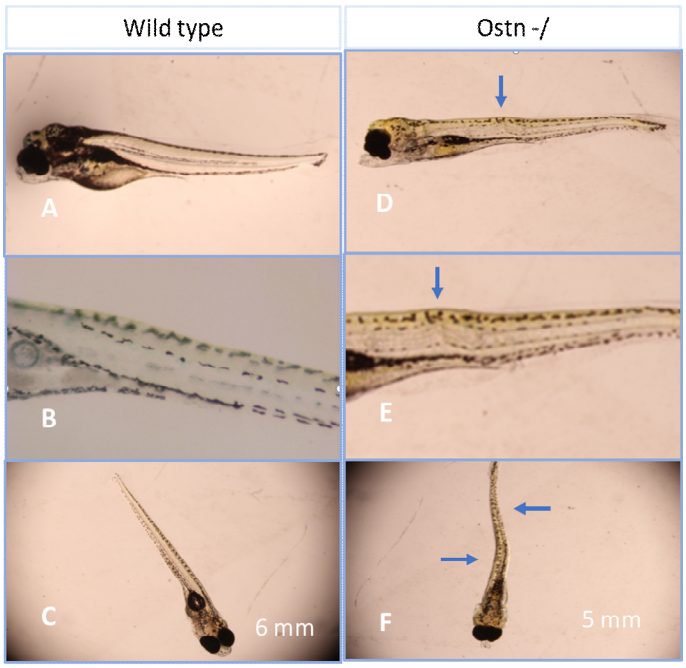


Figure 6. The effect of *ostn* knockout on scale development. Overall, the wild-type fish (left panel) had regular scale development, while in the ostn knockout, the scale development was slowed down. Note the differences at 11 mm long (A and G), and the differences in patterns and intensity of melanophores (compare B with H, for the middle section of fish; and compare C and I for posterior section of the fish). The differences in melanophores are even more prominent at later stages (D, wild type fish, 16 mm long; J, mutant fish 18 mm long). The magnified views of the middle section and posterior section of fish are shown in E and K, and F and L.

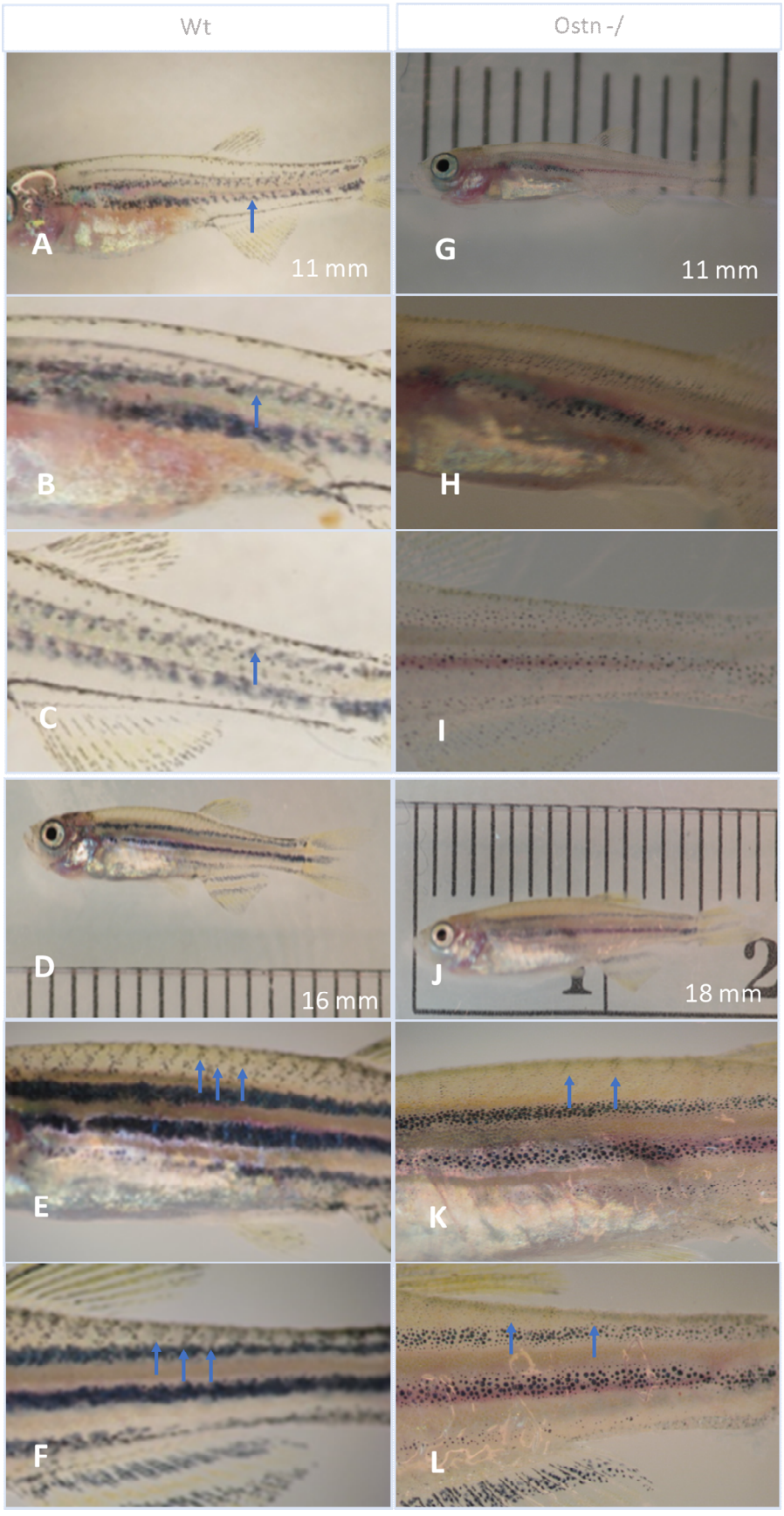

Wt
Ostn -/
A
11 mm
G
11 mm
B
H
C
I
D
16 mm
J
18 mm
E
K
F
L

Figure 7. Effect of ostn knockout on the development of scale, gill rakers, and operculum. The wild type fish is shown on the left panel and the knockout fish on the right panel. A and B showed normal scale numbers in Wt covering entire body; G and H, showed partial loss of scales (indicated by blue arrow). C and D, Wt fish gill rakers, and I and J, *ostn*$^{-/}$ knockout gill rakers. Compared to the wild type operculum (E and F), the development of operculum in *ostn*$^{-/}$ knockout fish (K and L) was interfered such that it was only partially formed, leaving gill exposed rather than covered under the operculum in normal fish. Abbreviations on the figure are: an, anterior; pos, posterior; gr, gill rakers.

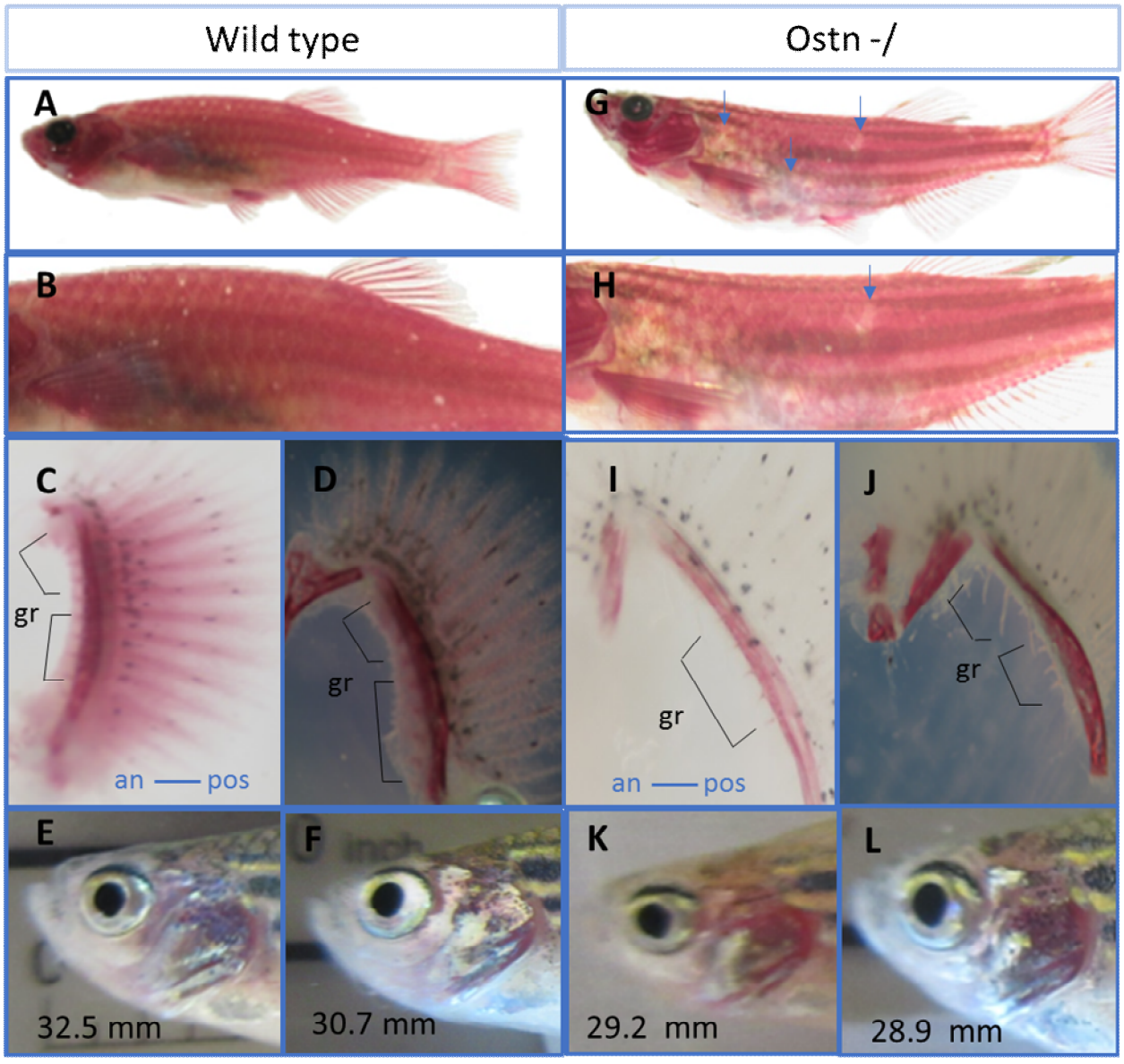


Figure 8. The effect of ostn knockout on the development of lateral bone (spinal cord). Note that the spinal cord was curved with mutants as compared with the wild type (compare E with A, ventral view at 70 dpf). B and F, lateral view when the muscle was digested, and the bones were stained with alizarin red, noting the curved spinal cord (blue arrows in F), which are shown in magnified view in C and G, with the spinal cord segments labeled from anterior to posterior order. D and H, significantly reduced calcification of bones as reflected in reduced staining of the main spinal cord as well as the inter-muscular bones. The abbreviations used in the figure are: en, epineurals; ep, epineurals; an, anterior; and pos, posterior.

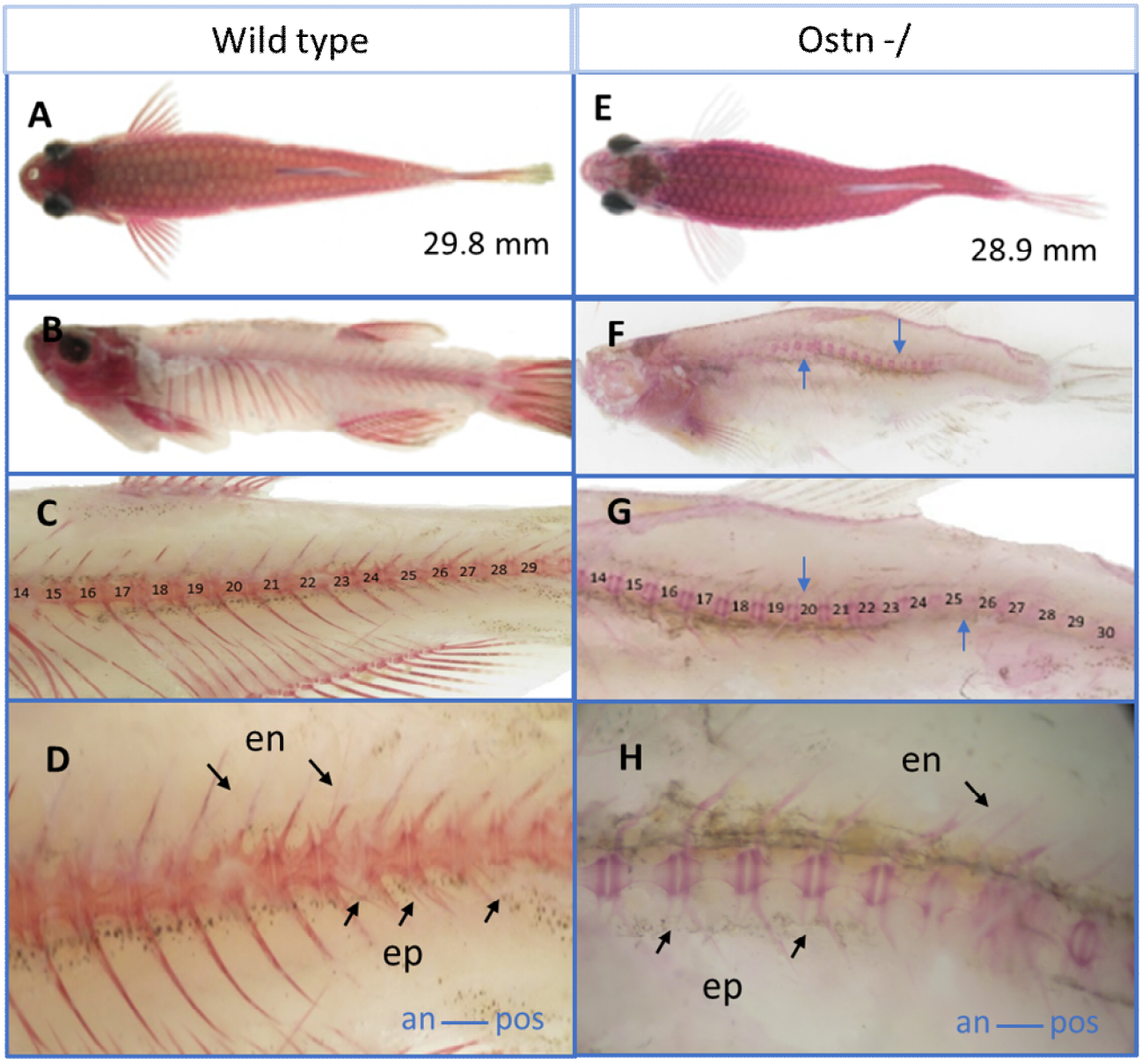
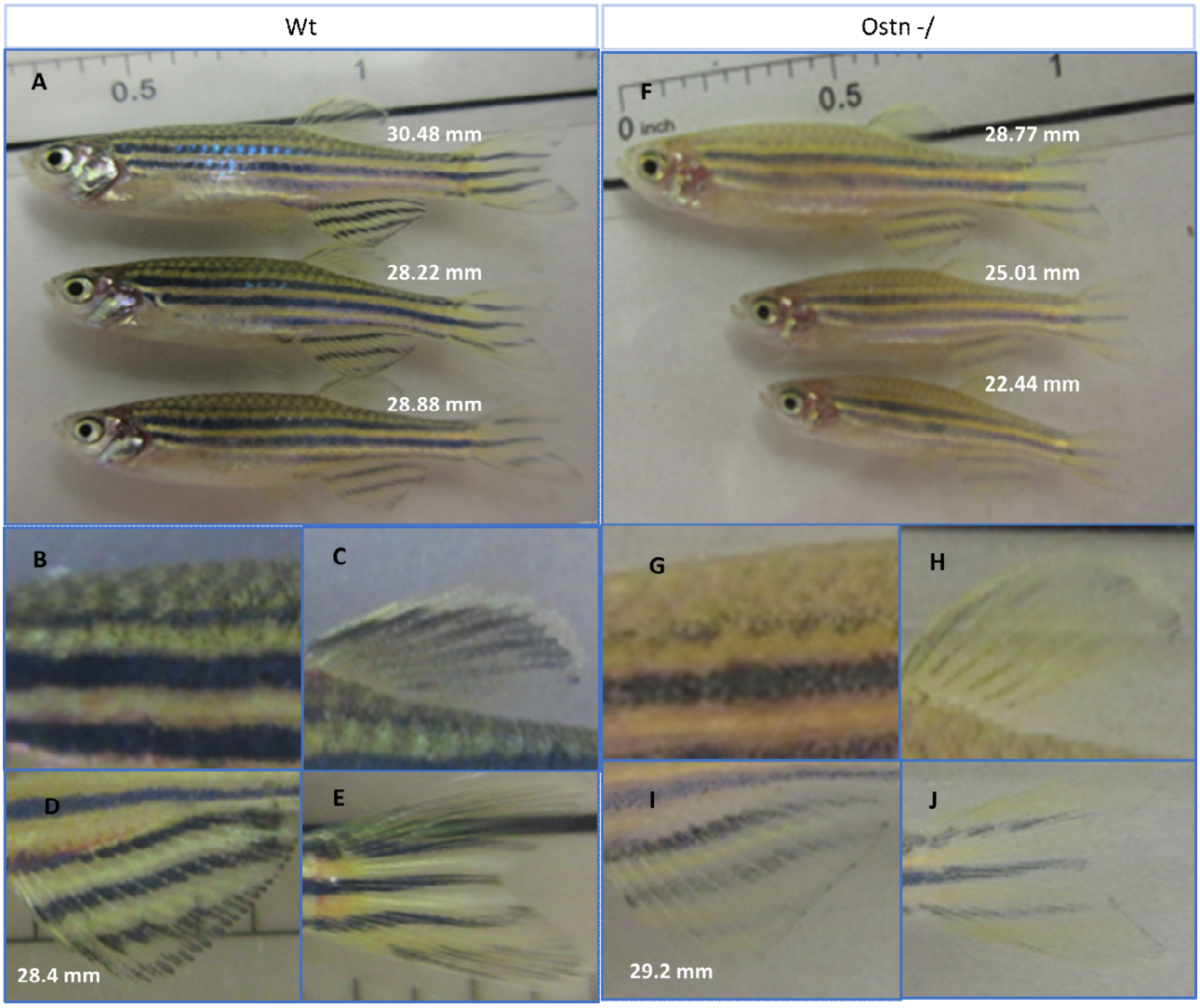
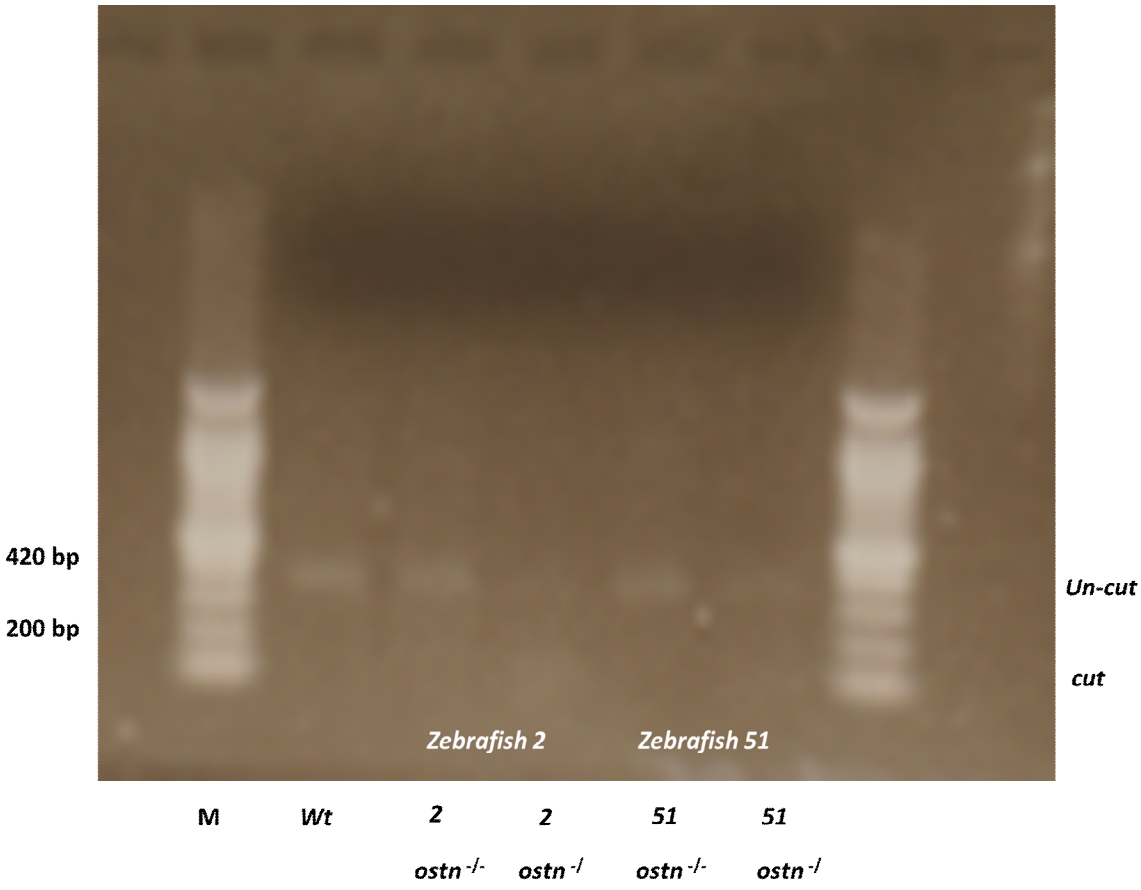


Figure 9. The effect of ostn knockout on growth and pigmentation of the skin. The mutants were smaller [compare wild type (A) on the left panel with mutant fish (B) on the right panel]. Pigmentation was much reduced in mutant fish (G-J, right lower panel) than in wild-type fish (B-E, left lower panel), showing various parts of the fish.



Genotyping analysis was conduced using gel electrophoreses of the PCR products with T7 endonuclease 1 assay (Guschin, 2010). It appeared that many fish exhibited heterozygous patterns of knockouts, suggesting that only one copy of the gene was knocked out, or the gene(s) were knocked out at later stages of embryo development (Fig. 10).

Figure 10. Gel electrophorese result from Wt and *ostn* $^{-/}$ cDNA extraction. Genomic PCR from zebrafish was generated and tested for gene knockout-induced mutation by T7 endo I assay. The Zebrafish with bone and partial scale lost defects) and chromatophore dispersion show mismatches at the target site. The ban shows a cut at the heterozygous mutant.



V. DISCUSSION

Osteocrin is a secreted protein which was initially identified in skeletal muscle and bone of mice (Subbotina, 2015; Moffatt 2007). It encodes 14.7 kDa, and 14.8 kDa secreted protein in human and zebrafish, respectively. Although its gene structure and encoded proteins were known, its functions were not well understood. In this study, we found a correlation of the presence of an osteocrin gene with the presence of scales among teleost fishes, suggesting that it may be involved in scale formation. Additional evidence for its involvement in scale formation came from expression analysis. Osteocrin is expressed at highest levels in the bone, but at lower levels in the fin and scales. Nonetheless, it is expressed at much higher levels in calcified tissues than in soft tissues, again suggesting its roles in osteogenesis, fin and scale formation. To gain insights into the functions of osteocrin, especially to determine if it is involved in scale formation, we conducted gene knockout experiments using the CRISPR gene knockout technique. We found *ostn* knockout has multiple effects including reduced rate of survival, reduced growth, bone deformities, partial scale loss, a significant reduction in chromatophores, abnormal gill rakers, and abnormal operculum. All of these defects were observed at various levels of severity, suggesting the pleiotropic functions of osteocrin.

ostn role in scale patterning and evolution

Fish scale plays a substantial role in serving as a physical barrier to protect fish from various injuries or infections. One highly interesting observation was that teleost models in this study were clustered based on their scale type using osteocrin to construct the phylogenetic trees. The scale types included placoid, ganoid, cosmoid, and elasmoid (cycloid, ctenoid). The phylogeny places the placoid from the shark as the ancestor among all vertebrates. The phylogeny also

described the ganoid and cosmoid scale type as the ancestor among the reptiles, bird, amphibian, and mammalians.

It was very interesting to find the correlation of osteocrin gene and scale status in teleosts. Those with scales have osteocrin gene, and those scaleless fishes lack the osteocrin gene, providing an arguable conclusion that osteocrin got to be involved in scale formation. Apparently, when present, it is more certain the osteocrin gene is present in the genome. However, when absent, it is possible that the gene was simply not sequenced even though it may still exist in the genome. In order to provide additional evidence, we conducted the syntenic analysis. Conserved syntenic blocks were identified from various organisms ranging from the old Sarcopterygians coelacanth, teleosts, and vertebrate genomes, suggesting a high level of evolutionary conservation in the genomic neighborhood containing the osteocrin gene. The findings that the osteocrin gene was found in organisms in the broad evolutionary spectrum suggested that the gene was missing in the scaleless fishes such as most of the catfishes in the order of Siluriformes, the green pufferfish, and stickleback. The osteocrin gene from pleco had just one NP motif in the osteocrin gene, compared to two NP motifs in the osteocrin gene of many other teleost species. Whether this is related to the bony dermal plate type of scales in pleco is not known at present, but it appeared that the osteocrin gene structure may have a high level of parallel relationship with the types of scales. Additional evidence of osteocrin involvement in scale formation came from expression analysis. On the one hand, osteocrin was expressed at higher levels in calcified tissues including the scales; on the other hand, ostn-b was highly differentially expressed during carp scale regeneration. Ostn-b expression was upregulated 7-14 times at the early stages of scale regeneration, day 3 to day 14, suggesting its involvement in scale formation.

The pleiotropic effect of osteocrin knockout

Osteocrin has NPs motif which is concluded to belong to NP family even though it lacks cysteine residues which conserved among the NP family member, such as an atrial natriuretic peptide (ANP), brain natriuretic peptide (BNP), and type-C natriuretic peptide (CNP) (Potter 2006; Chiba 2017). CNP affect osteoblast proliferation through NPR2 (GC-B) (Suda, 1998; Potter 2006) and while depleted or clearance by inactivator NPRC may lead dwarfism (Tamura, 2004) or enhance retardation and skeletal defects (Peake, 2014). O*stn* blocks the CNP clearance by NPRC which increase the CNP access to NPR2 thus enhance cGMP activity in the cartilage homeostasis and endochondral bone formation (Moffatt 2007; Chiba). Parallel to this experiment, the *ostn* knockout may stimulate the clearance activity of NPRC towards CNP which causing an anomaly in most of the calcified tissues, including slow scale growth, partial scale loss, bone curves, shorter intermuscular bone, abnormal gill rakers, incomplete operculum, and chromatophores dispersion.

In post-embryonic stage, the zebrafish as in most of the teleost showed late dermal skeleton tissue formation including the scale. The elasmoid scale in zebrafish emerge at the first time at 9 mm length (30 dpf) indicate by aggregation of melanophores forming a pocket in dermis layer, part of skin integumentary system (Elliott 2011). One possibility causing the partial loss and slow growth of the scale in post-embryonic and adult fish is the imperfect calcification process in those scale as part of the integument. The expression level of *ostn* in most calcified tissues but not in soft tissues except for brain and hearth could support this argument. Osteocrin is defined as a peptide secreted from the hearth which contributes to cranial osteogenesis (Chiba 2017) suggesting an expression also occurred in hearth and brain tissues. The calcification

process itself involved in all part of the dermal skeleton, including scale, bone, gill rakers, and skin which grow throughout the life in most teleost as the consequences of development or regeneration (Huysseune, 2007; Witten, 2009; Akimenko, 2003). The fish scale is structured by collagen fibril and salt calcium (Sionkowska A. 2014). The calcification starts at the fibrillary plate of scales which contains a high concentration of calcium (Onozato, 1979; Iguchi 2012). The abnormal calcification process resulted in thin and incomplete scale growth as seen in mutated *ostn* $^{-/}$ fish which showed pale and partial loss scale. The fish scale also provides a barrier between the soft fibril skin with the environment. The loss and thickness of scale could affect the break of skin epidermis easily as shown in mutated *ostn*$^{-/}$ (Video S4).

The knockout of *ostn* resulting several defects not only in the scale but also in other skeletal tissues, such as lateral bone (spinal cord), intermuscular bone, gill rakers, operculum, and chromatophores. Osteocrin has been reported as a gene contributed in human bone formation (Bord 2005) and zebrafish bone development through endocrinal work (Chiba 2017) suggesting a possibility of involvement in calcified tissues. Concomitant with this study, we show the knockout of *ostn* had multiple pleiotropic effects on the bone which also could be observed through abnormal swimming mode (Video S2). The bone mineralization involves the alkaline phosphatase and calcium in orderly compose hydroxyapatite crystals which are closely associated with collagen (type I and type II) to form the bone matrix (Kawasaki 2009). The bone matrix dominated with calcium could be stained using Alizarin red. Interestingly, the bone matrix in mutated zebrafish was also not well stained. The reddish color in the lateral intermuscular bone segment in *ostn* $^{-/}$ fish exhibited shorter ephineural and epipleural along with a disperse, soluble, or pale color in compared to the Wt. On the contrary, the Wt exhibited well-stained bone segments showed a chimeric reddish color. These results, parallel with the

expression of *ostn* in a normal calcified tissue, may provide substantial evidence of its role in a bone matrix formation.

Gills are an evagination part of the body which covered by a bony plate, the operculum (Maina 2002; Kimmel 2008). The gills structure are composed of bony or cartilaginous in gill arch and gill rakers (Engeman, 2009; Glazer, 2014). The gill rakers are part of gills which functionally filtrating tiny feed (Akimenko, 2007; Sanderson, 2016). We found Wt exhibited normal operculum and gill rakers, while the *ostn* $^{-/}$ fish exhibited abnormal operculum, partial loss, and asymmetrical gill rakers. The abnormalities in gill rakers are suggesting two possibilities, an abnormal ossification or a random growth as adaptation process to the feeding habit (Friedland, 2006; Kumari, 2009). In addition, the operculum is part of bone series found in bony fish (Kimmel, 2008) suggesting the abnormalities may be in correlation with abnormal ossification. The abnormal ossification resulted in an incomplete and thin operculum covering the gill pouch.

Chromatophores are derived from neural crest during embryonic development along with the craniofacial cartilage and neurons of the peripheral nervous system (Kelsh, 1996; Le Douarin, 1999; Arduini, 2008). The pigmentations in zebrafish are observed at 24 h postfertilization spread over the skin (Kimmel, 1995). Interestingly, in mutated zebrafish we found that the melanophore stripes and inner stripes were not aggregated as well as in normal Wt, suggesting the occurrence of a pigment dispersion. The pigments dispersion in the mutant occurred on the body surface, fin, and tail, in which parallel with the expression of *ostn* in most ossified tissue. The ossified tissues are known as a pool of calcium. The calcium and cyclic nucleotides, including cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) have been reported in the pigment dispersion and aggregation (Ribeiro,

2009). The cyclic AMP has been reported to perform opposite effect to cyclic GMP on the Ca^{2+} current in single heart cells (Hartzell, 1986). The calcium itself reciprocally affected cAMP resulting in several pigment aggregation and migration in fish (Kotz, 1994). The cation Ca^{2+} has a role in signal transduction resulting melanophore aggregation in *Labeo* (Patil, 1993), and a high calcium content also detected in melanosomes of *Xenopus laevis* and *Poecilia reticulate* (Muller, 1991). On the contrary, the high cAMP and low intracellular Ca^{2+} made erythrophores dispersion in squirrel fish suggesting an opposite work between cAMP and Ca^{2+} (Kotz 1994). In addition, the cGMP also has been reported participate in chromatophore aggregation in freshwater shrimp (Ribeiro, 2009). The disruption of *ostn* maybe effect the cGMP, calcium, and cAMP activity resulting chromatophore dispersion in zebrafish, but a physiological pathway needs to be tested.

VI. CONCLUSION

Through genome comparative study method, the completeness of channel catfish genome could be used to reveal the missing of scales by comparing with the bony dermal plates catfish genome. Furthermore, I validated the predicted genes along with other scaled and scaleless vertebrates. Here, for the first time, the *ostn* recently was found conserved among scaled vertebrates and missing from the scaleless fish genome. Through phylogeny tree, the teleost models fell into theirs scale type, suggesting the involvement of *ostn* in scale type patterning. Along with scale regeneration experiment in common carp, *ostn* expressed specifically in each time describing the development of fibril collagen in scale. The CRISPR/Cas9 system in zebrafish revealed the function of *ostn*. During the post-embryonic period, *ostn* disruption resulted in bone defects which later in adult mutant could be proved by the alteration of some lateral bone segments. At adult stages, mutated *ostn* fish showed slow scale growth compare to the Wt, resulted in a loss and fragile scales. The disruption of *ostn* on adult fish also resulted some defects on gill rakers, operculum, and chromatophores dispersion suggesting an evidence of *ostn* involvement among calcified tissue. This study could support the evolutionary and comparative study of scaled and scaleless fish in associated with calcified tissues in human and among vertebrates.

Referrences

Akimenko, M. A., A. Smith. 2007. 'Paired fin repair and regeneration.' in B. K. Hall (ed.), *Fins into Limbs: Evolution, Development, and Transformation* (The University of Chicago Press: Chicago).

Akimenko, M. A., M. Mari-Beffa, J. Becerra, and J. Geraudie. 2003. 'Old questions, new tools, and some answers to the mystery of fin regeneration', *Dev Dyn*, 226: 190-201.

Alibardi, L. 2004. 'Dermo-epidermal interactions in reptilian scales: speculations on the evolution of scales, feathers, and hairs', *J Exp Zool B Mol Dev Evol*, 302: 365-83.

Arduini, B. L., G. R. Gallagher, and P. D. Henion. 2008. 'Zebrafish Endzone Regulates Neural Crest-Derived Chromatophore Differentiation and Morphology', *PLOS ONE*, 3.

Auer, T. O. 2014. 'CRISPR/Cas9-mediated conversion of eGFP- into Gal4-transgenic lines in zebrafish', *Nat Protoc*, 9.

Bereiter-Hahn J., L. Zylberberg. 1993. 'Regeneration of teleost fish scale', *Comparative biochemistry and physiology*, 105A:625-641.

Bord, S., D. C. Ireland, P. Moffatt, G. P. Thomas, and J. E. Compston. 2005. 'Characterization of osteocrin expression in human bone', *J Histochem Cytochem*, 53: 1181-7.

Brazeau, M. D., M. Friedman. 2014. 'The characters of Palaeozoic jawed vertebrates', *Zool J Linn Soc*, 170: 779-821.

Brown, T.A. 2002. 'Genomes, 2nd Edition', *Oxford: Wiley-Liss*: 520.

Chai, H. J., J. H. Li, H. N. Huang, T. L. Li, Y. L. Chan, C. Y. Shiau, and C. J. Wu. 2010. 'Effects of Sizes and Conformations of Fish-Scale Collagen Peptides on Facial Skin Qualities and Transdermal Penetration Efficiency', *J Biomed Biotechnol*, 2010.

Chang, C., P. Wu, R. E. Baker, P. K. Maini, L. Alibardi, and C. M. Chuong. 2009. 'Reptile scale paradigm: Evo-Devo, pattern formation and regeneration', *Int J Dev Biol*, 53: 813-26.

Chiba, A., H. W. Takano, K. Terai, H. Fukui,, M. Uemura T. Miyazaki, H. Hashimoto, , and S. Fukuhara M. Hibi, N. Mochizuki,. 2017. 'Osteocrin, a peptide secreted from the heart and other tissues, contributes to cranial osteogenesis and chondrogenesis in zebrafish', *Development*, 144: 334-44.

Chuong, C. M., R. Chodankar, R. B. Widelitz, and T. X. Jiang. 2000. 'Evo-Devo of feathers and scales: building complex epithelial appendages', *Curr Opin Genet Dev*, 10: 449-56.

Costa, R. A., J. C. R. Cardoso, and D. M. Power. 2017. 'Evolution of the angiopoietin-like gene family in teleosts and their role in skin regeneration', *BMC Evol Biol*, 17.

D'Agostino, Y. 2016. 'A Rapid and Cheap Methodology for CRISPR/Cas9 Zebrafish Mutant Screening', *Mol Biotechnol*, 58.

D'Agostino, Y., A. Locascio, F. Ristoratore, P. Sordino, A. Spagnuolo, M. Borra, and S. D'Aniello. 2016. 'A Rapid and Cheap Methodology for CRISPR/Cas9 Zebrafish Mutant Screening', *Mol Biotechnol*, 58: 73-8.

Dhouailly, D. 2009. 'A new scenario for the evolutionary origin of hair, feather, and avian scales', *J Anat*, 214: 587-606.

Ding, Y., H. Li, L. L. Chen, and K. Xie. 2016. 'Recent Advances in Genome Editing Using CRISPR/Cas9', *Front Plant Sci*, 7.

Doudna, J.A., E. J. Sontheimer. 2014. 'Methods in Enzymlogy, The Use of CRISPR/Cas9, ZFNs, and TALENs in Generating Site-Specific Genome Alterations', 546: 572.

Dunham, R. 2006. 'History of catfish breeding and its application in the United States: Lessons to be learned? ', *The Israeli Journal of Aquaculture*, 58: 251-56.

Dunham, R., R.O. Smitherman. 1984. 'Ancestry and breeding of catfish in the United States', *Auburn: Alabama Agri. Exp. Station Circular.*: 93.

Ellegren, Hans. 2008. 'Comparative genomics and the study of evolution by natural selection', *Molecular Ecology*, 17: 4586-96.

Elliott, D. G. 2011. 'THE SKIN | Functional Morphology of the Integumentary System in Fishes A2 - Farrell, Anthony P.' in, *Encyclopedia of Fish Physiology* (Academic Press: San Diego).

Engeman, J. M., N. Aspinwall, and P. M. Mabee. 2009. 'Development of the pharyngeal arch skeleton in Catostomus commersonii (Teleostei: Cypriniformes)', *J Morphol*, 270: 291-305.

Finn, R. D., T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork, A. J. Bridge, H. Y. Chang, Z. Dosztanyi, S. El-Gebali, M. Fraser, J. Gough, D. Haft, G. L. Holliday, H. Huang, X. Huang, I. Letunic, R. Lopez, S. Lu, A. Marchler-Bauer, H. Mi, J. Mistry, D. A. Natale, M. Necci, G. Nuka, C. A. Orengo, Y. Park, S. Pesseat, D. Piovesan, S. C. Potter, N. D.

Rawlings, N. Redaschi, L. Richardson, C. Rivoire, A. Sangrador-Vegas, C. Sigrist, I. Sillitoe, B. Smithers, S. Squizzato, G. Sutton, N. Thanki, P. D. Thomas, S. C. Tosatto, C. H. Wu, I. Xenarios, L. S. Yeh, S. Y. Young, and A. L. Mitchell. 2017. 'InterPro in 2017-beyond protein family and domain annotations', *Nucleic Acids Res*, 45: D190-d99.

Friedland, K. D., D. W. Ahrenholz, J. W. Smith, M. Manning, and J. Ryan. 2006. 'Sieving functional morphology of the gill raker feeding apparatus of atlantic menhaden', *J Exp Zool A Comp Exp Biol*, 305: 974-85.

Geng, X., J. Sha, S. Liu, L. Bao, J. Zhang, R. Wang, J. Yao, C. Li, J. Feng, F. Sun, L. Sun, , Y. Zhang C. Jiang, A. Chen, R. Dunham, , and Z. Liu D. Zhi. 2015. 'A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTLs for columnaris disease resistance', *BMC Genomics*, 16: 196.

Gil-Duran, S., D. Arola, and E. A. Ossa. 2016. 'Effect of chemical composition and microstructure on the mechanical behavior of fish scales from Megalops Atlanticus', *J Mech Behav Biomed Mater*, 56: 134-45.

Glazer, A. M., P. A. Cleves, P. A. Erickson, A. Y. Lam, and C. T. Miller. 2014. 'Parallel developmental genetic features underlie stickleback gill raker evolution', *EvoDevo*, 5: 19.

Gomez, C., W. Chua, A. Miremadi, S. Quist, J. Headon D , and M. Watt F 2013. 'The Interfollicular Epidermis of Adult Mouse Tail Comprises Two Distinct Cell Lineages that Are Differentially Regulated by Wnt, Edaradd, and Lrig1', *Stem Cell Reports*, 1: 19-27.

Gumucio, D. L., H. Heilstedt-Williamson, T. A. Gray, S. A. Tarlé, D. A. Shelton, D. A. Tagle, J. L. Slightom, M. Goodman, and F. S. Collins. 1992. 'Phylogenetic footprinting reveals a nuclear protein which binds to silencer sequences in the human gamma and epsilon globin genes', *Mol Cell Biol*, 12: 4919-29.

Guschin, D. Y., A. J. Waite, G. E. Katibah, J. C. Miller, M. C. Holmes, and E. J. Rebar. 2010. 'A rapid and general assay for monitoring endogenous gene modification', *Methods Mol Biol*, 649: 247-56.

Hardison, R. C. 2003. 'Comparative Genomics', *PLoS Biol*, 1.

Harris, M. P., N. Rohner, H. Schwarz, S. Perathoner, P. Konstantinidis, and C. Nusslein-Volhard. 2008. 'Zebrafish eda and edar mutants reveal conserved and ancestral roles of ectodysplasin signaling in vertebrates', *PLoS Genet*, 4: e1000206.

Hartzell, H. C., and R. Fischmeister. 1986. 'Opposite effects of cyclic GMP and cyclic AMP on Ca2+ current in single heart cells', *Nature*, 323: 273-5.

Hsu, P. D., E. S. Lander, and F. Zhang. 2014. 'Development and Applications of CRISPR-Cas9 for Genome Engineering', *Cell*, 157: 1262-78.

Hu, B., G. Xie, C. C. Lo, S. R. Starkenburg, and P. S. Chain. 2011. 'Pathogen comparative genomics in the next-generation sequencing era: genome alignments, pangenomics and metagenomics', *Brief Funct Genomics*, 10: 322-33.

Huysseune, A., B. K. Hall, and P. E. Witten. 2007. 'Establishment, maintenance and modifications of the lower jaw dentition of wild Atlantic salmon (Salmo salar L.) throughout its life cycle', *J Anat*, 211: 471-84.

Hwang, W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai, J. D. Sander, R. T. Peterson, J. R. Yeh, and J. K. Joung. 2013. 'Efficient genome editing in zebrafish using a CRISPR-Cas system', *Nat Biotechnol*, 31: 227-9.

Iger, Y., M. Abraham. 1990. 'The process of skin healing in experimentally wounded carp', *Journal of Fish Biology*, 36: 421-37.

Iguchi, M. 2012. 'Effects of etidronate on calcification of scales and ribs in the goldfish, Carassius auratus', *Fish Physiol Biochem*, 38: 483-91.

Irion, U., J. Krauss, C. Nüsslein-Volhard. 2014. 'Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system', *Development*, 141: 4827-30.

Jao, L. E., S. R. Wente, and W. Chen. 2013. 'Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system', *Proc Natl Acad Sci U S A*, 110: 13904-9.

Jiang, Y., X. Gao, S. Liu, Y. Zhang, H. Liu, F. Sun, L. Bao, G. Waldbieser, and Z. Liu. 2013. 'Whole genome comparative analysis of channel catfish (Ictalurus punctatus) with four model fish species', *BMC Genomics*, 14: 780.

Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. 'The rapid generation of mutation data matrices from protein sequences', *Comput Appl Biosci*, 8: 275-82.

Katoh, K., D. M. Standley. 2013. 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability', *Mol Biol Evol*, 30: 772-80.

Kawasaki, K. 2009. 'The SCPP gene repertoire in bony vertebrates and graded differences in mineralized tissues', *Dev Genes Evol*, 219: 147-57.

Kelsh, R. N., M. Brand, Y. J. Jiang, C. P. Heisenberg, S. Lin, P. Haffter, J. Odenthal, M. C. Mullins, F. J. van Eeden, M. Furutani-Seiki, M. Granato, M. Hammerschmidt, D. A. Kane, R. M. Warga, D. Beuchle, L. Vogelsang, and C. Nusslein-Volhard. 1996. 'Zebrafish pigmentation mutations and the processes of neural crest development', *Development*, 123: 369-89.

Kimmel, C. B. 1995. 'Stages of embryonic development of the zebrafish', *Dev Dyn*, 203.

Kimmel, C. B., W. E. Aguirre, B. Ullmann, M. Currey, and W. A. Cresko. 2008. 'Allometric change accompanies opercular shape evolution in Alaskan threespine sticklebacks', *Behaviour*, 145: 669-91.

Koonin, E. V, and M. Y. Galperin. 2003. *Sequence - Evolution - Function: Computational Approaches in Comparative Genomics* (Kluwer Academic: Boston).

Kotz, K. J., and M. A. McNiven. 1994. 'Intracellular calcium and cAMP regulate directional pigment movements in teleost erythrophores', *J Cell Biol*, 124: 463-74.

Kumar, S., and A. Filipski. 2007. 'Multiple sequence alignment: in pursuit of homologous DNA positions', *Genome Res*, 17: 127-35.

Kumari, U., M. Yashpal, S. Mittal, and A. K. Mittal. 2009. 'Surface ultrastructure of gill arches and gill rakers in relation to feeding of an Indian major carp, Cirrhinus mrigala', *Tissue Cell*, 41: 318-25.

Lanctot, C., P. Moffatt, and G. Thomas. 2007. "Method of use of specific natriuretic peptide receptor c ligands, transgenic non-human mammals expressing specific natriuretic peptide receptor c antagonists and cells thereof." In.: Google Patents.

Le Douarin N., and C. Kalcheim. 1999. 'The neural crest', *Developmental and Cell Biology Series*: pp xxiii, 445, [18] of plates.

Le Guellec, D., G. Morvan-Dubois, and J. Y. Sire. 2004. 'Skin development in bony fish with particular emphasis on collagen deposition in the dermis of the zebrafish (Danio rerio)', *Int J Dev Biol*, 48: 217-31.

Liang, G., H. Zhang, D. Lou, and D. Yu. 2016. 'Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing', *Sci Rep*, 6.

Lieschke, G. J., and P. D. Currie. 2007. 'Animal models of human disease: zebrafish swim into view', *Nat Rev Genet*, 8.

Lim, K. C., P. E. Lim, V. C. Chong, and K. H. Loh. 2015. 'Molecular and Morphological Analyses Reveal Phylogenetic Relationships of Stingrays Focusing on the Family Dasyatidae (Myliobatiformes)', *PLOS ONE*, 10.

Liu, S., Y. Zhang, Z. Zhou, G. Waldbieser, F. Sun, J. Lu, J. Zhang, Y. Jiang, H. Zhang, X. Wang, K. V. Rajendran, L. Khoo, H. Kucuktas, E. Peatman, and Z. Liu. 2012. 'Efficient assembly and annotation of the transcriptome of catfish by RNA-Seq analysis of a doubled haploid homozygote', *BMC Genomics*, 13: 595.

Liu, Y., X. Huo, X. F. Pang, Z. H. Zong, X. Meng, and G. L. Liu. 2008. 'Musclin inhibits insulin activation of Akt/protein kinase B in rat skeletal muscle', *J Int Med Res*, 36: 496-504.

Liu, Z. 2011. 'Development of genomic resources in support of sequencing, assembly, and annotation of the catfish genome', *Comparative biochemistry and physiology. Part D, Genomics & proteomics*, 6: 11-17.

Liu, Z., S. Liu, J. Yao, L. Bao, J. Zhang, Y. Li, C. Jiang, L. Sun, R. Wang, Y. Zhang, T. Zhou, Q. Zeng, Q. Fu, S. Gao, N. Li, S. Koren, Y. Jiang, A. Zimin, P. Xu, A. M. Phillippy, X. Geng, L. Song, F. Sun, C. Li, X. Wang, A. Chen, Y. Jin, Z. Yuan, Y. Yang, S. Tan, E. Peatman, J. Lu, Z. Qin, R. Dunham, Z. Li, T. Sonstegard, J. Feng, R. G. Danzmann, S. Schroeder, B. Scheffler, M. V. Duke, L. Ballard, H. Kucuktas, L. Kaltenboeck, H. Liu, J. Armbruster, Y. Xie, M. L. Kirby, Y. Tian, M. E. Flanagan, W. Mu, and G. C. Waldbieser. 2016. 'The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts', *Nat Commun*, 7: 11757.

Livak, K. J., and T. D. Schmittgen. 2001. 'Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method', *Methods*, 25: 402-8.

Louis, A., N. T. T. Nguyen, M. Muffato, , and H. R. Crollius. 2015. 'Genomicus update 2015: KaryoView and MatrixView provide a genome-wide perspective to multispecies comparative genomics', *Nucleic Acids Research*, 43: D682-D89.

Maina, J. 2002. 'Structure, function and evolution of the gas exchangers: comparative perspectives', *J Anat*, 201: 281-304.

Mali, P., K. M. Esvelt, and G. M. Church. 2013. 'Cas9 as a versatile tool for engineering biology', *Nat Methods*, 10: 957-63.

Maruyama, T., S. K. Dougan, M. C. Truttmann, A. M. Bilate, J. R. Ingram, and H. L. Ploegh. 2015. 'Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining', *Nat Biotechnol,* 33: 538-42.

Meeker, N. D. 2007. 'Method for isolation of PCR-ready genomic DNA from zebrafish tissues', *Biotechniques,* 43.

Meyer, W., M. Liumsiricharoen, A. Suprasert, L. G. Fleischer, and M. Hewicker-Trautwein. 2013. 'Immunohistochemical demonstration of keratins in the epidermal layers of the Malayan pangolin (Manis javanica), with remarks on the evolution of the integumental scale armour', *Eur J Histochem,* 57: e27.

Moffatt, P., G. Thomas, K. Sellin, M. C. Bessette, F. Lafreniere, O. Akhouayri, R. St-Arnaud, and C. Lanctot. 2007. 'Osteocrin is a specific ligand of the natriuretic Peptide clearance receptor that modulates bone growth', *J Biol Chem,* 282: 36454-62.

Moreno-Mateos, M. A. 2015. 'CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo', *Nat Methods,* 12.

Moyle, P. B., J. J. J. Cech. 2004. 'Fishes: an introduction to ichthyology 5th edn': 726.

Muller, T., and J. Bereiter-Hahn. 1991. 'Demonstration of calcium in dermal melanocytes of Xenopus laevis and Poecilia reticulata with electron energy-loss spectroscopy and electron spectroscopic imaging', *J Microsc,* 162: 141-6.

Nishizawa, H., M. Matsuda, Y. Yamada, K. Kawai, E. Suzuki, M. Makishima, T. Kitamura, and I. Shimomura. 2004. 'Musclin, a novel skeletal muscle-derived secretory factor', *J Biol Chem*, 279: 19391-5.

Niu, L., W. Zhang, D. H. Pashley, L. Breschi, J. Mao, J. Chen, and F. R. Tay. 2014. 'Biomimetic remineralization of dentin', *Dent Mater*, 30.

Ohira Y., M. Shimizu, K. Ura, Y. Takagi. 2007. 'Scale regeneration and calcification in goldfish Carassius auratus: quantitative and morphological processes.', *Fisheries Science*, 73:46-54.

Onozato, H., and N. Watabe. 1979. 'Studies on fish scale formation and resorption. III. Fine structure and calcification of the fibrillary plates of the scales in Carassius auratus (Cypriniformes: Cyprinidae)', *Cell Tissue Res,* 201: 409-22.

Oshima, N., N. Nakamaru, S. Araki, and M. Sugimoto. 2001. 'Comparative analyses of the pigment-aggregating and -dispersing actions of MCH on fish chromatophores', *Comp Biochem Physiol C Toxicol Pharmacol*, 129: 75-84.

Ota, S. 2014. 'Multiple genome modifications by the CRISPR/Cas9 system in zebrafish', *Genes Cells*, 19.

Paix, A., A. Folkmann, D. Rasoloson, and G. Seydoux. 2015. 'High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes', *Genetics*, 201: 47-54.

Palmer, L. C., C. J. Newcomb, S. R. Kaltz, E. D. Spoerke, and S. I. Stupp. 2008. 'Biomimetic Systems for Hydroxyapatite Mineralization Inspired By Bone and Enamel', *Chem Rev*, 108: 4754-83.

Parichy, D. M., M. R. Elizondo, M. G. Mills, T. N. Gordon, and R. E. Engeszer. 2009. 'Normal Table of Post-Embryonic Zebrafish Development: Staging by Externally Visible Anatomy of the Living Fish', *Dev Dyn*, 238: 2975-3015.

Patil, Shashi; Jain, A.K. 1993. 'Role of calcium in melanosome aggregation within Labeo melanophores ', *J. Biosci*, Vol. 18, Number 1: pp 83–91.

Peake, N. J., A. J. Hobbs, B. Pingguan-Murphy, D. M. Salter, F. Berenbaum, and T. T. Chowdhury. 2014. 'Role of C-type natriuretic peptide signalling in maintaining cartilage and bone function', *Osteoarthritis Cartilage*, 22: 1800-7.

Potter, L. R., S. Abbey-Hosch, and D. M. Dickey. 2006. 'Natriuretic peptides, their receptors, and cyclic guanosine monophosphate-dependent signaling functions', *Endocr Rev*, 27: 47-72.

Pyne, M. E., M. R. Bruder, M. Moo-Young, D. A. Chung, and C. P. Chou. 2016. 'Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in Clostridium', *Sci Rep*, 6.

Qu, Q., T. Haitina, M. Zhu, and P. E. Ahlberg. 2015. 'New genomic and fossil data illuminate the origin of enamel', *Nature*, 526: 108-11.

Quilhac, A., and J. Y. Sire. 1999. 'Spreading, proliferation, and differentiation of the epidermis after wounding a cichlid fish, Hemichromis bimaculatus', *Anat Rec*, 254: 435-51.

Ramlee, M. K. 2015. 'High-throughput genotyping of CRISPR/Cas9-mediated mutants using fluorescent PCR-capillary gel electrophoresis', *Sci Rep*, 26.

Ribeiro, M. R., and J. C. McNamara. 2009. 'Cyclic guanosine monophosphate signaling cascade mediates pigment aggregation in freshwater shrimp chromatophores', *Biol Bull*, 216: 138-48.

Rohner, N., M. Bercsenyi, L. Orban, M. E. Kolanczyk, D. Linke, M. Brand, C. Nusslein-Volhard, and M. P. Harris. 2009. 'Duplication of fgfr1 permits Fgf signaling to serve as a target for selection during domestication', *Curr Biol*, 19: 1642-7.

Rozen, S., and H. Skaletsky. 2000. 'Primer3 on the WWW for general users and for biologist programmers', *Methods Mol Biol*, 132: 365-86.

Sander, J. D., and J. K. Joung. 2014. 'CRISPR-Cas systems for genome editing, regulation and targeting', *Nat Biotechnol*, 32: 347-55.

Sanderson, S. Laurie, Erin Roberts, Jillian Lineburg, and Hannah Brooks. 2016. 'Fish mouths as engineering structures for vortical cross-step filtration', 7: 11092.

Sasagawa, Ichiro, Mikio Ishiyama, Hiroyuki Yokosuka, and Masato Mikami. 2013. 'Teeth and ganoid scales in Polypterus and Lepisosteus, the basic actinopterygian fish: An approach to understand the origin of the tooth enamel', *Journal of Oral Biosciences*, 55: 76-84.

Scharer, R. M., W. F. Patterson, 3rd, J. K. Carlson, and G. R. Poulakis. 2012. 'Age and growth of endangered smalltooth sawfish (Pristis pectinata) verified with LA-ICP-MS analysis of vertebrae', *PLOS ONE*, 7: e47850.

Schneider, C. A., W. S. Rasband, K. W. Eliceiri. . 2012. 'NIH Image to ImageJ: 25 years of image analysis', *Nat Meth*, 9: 671-75.

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J. D. Thompson, and D. G. Higgins. 2011. 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Mol Syst Biol*, 7: 539.

Sionkowska A., J. Kozlowska. 2014. 'Fish Scales as a Biocomposite of Collagen and Calcium Salts', *Key Engineering Materials*, Vol. 587: pp. 185-90.

Sire, J., F. Allizard, O. Babiar, J. Bourguignon, and A. Quilhac. 1997. 'Scale development in zebrafish (Danio rerio)', *J Anat*, 190: 545-61.

Sire, J. Y., A. N. N. Huysseune. 2003. 'Formation of dermal skeletal and dental tissues in fish: a comparative and evolutionary approach', *Biological Reviews*, 78: 219-49.

Sire, J. Y., and M. A. Akimenko. 2004. 'Scale development in fish: a review, with description of sonic hedgehog (shh) expression in the zebrafish (Danio rerio)', *Int J Dev Biol*, 48: 233-47.

Solovyev, V., P. Kosarev, I. Seledsov, and D. Vorobyev. 2006. 'Automatic annotation of eukaryotic genes, pseudogenes, and promoters', *Genome Biol*, 7: S10.

Stanke, M., and B. Morgenstern. 2005. 'AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints', *Nucleic Acids Res*, 33: W465-7.

Subbotina, E., A. Sierra, Z. Zhu, Z. Gao, S. R. Koganti, S. Reyes, E. Stepniak, S. A. Walsh, M. R. Acevedo, C. M. Perez-Terzic, D. M. Hodgson-Zingman, and L. V. Zingman. 2015. 'Musclin is an activity-stimulated myokine that enhances physical endurance', *Proc Natl Acad Sci U S A*, 112: 16042-7.

Suda, M., Y. Ogawa, K. Tanaka, N. Tamura, A. Yasoda, T. Takigawa, M. Uehira, H. Nishimoto, H. Itoh, Y. Saito, K. Shiota, and K. Nakao. 1998. 'Skeletal overgrowth in transgenic mice that overexpress brain natriuretic peptide', *Proc Natl Acad Sci U S A*, 95: 2337-42.

Sullivan, J. P., J. G. Lundberg, and M. Hardman. 2006. 'A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences', *Mol Phylogenet Evol*, 41: 636-62.

Takagi, Y., and K. Ura. 2007. 'Teleost fish scales: a unique biological model for the fabrication of materials for corneal stroma regeneration', *J Nanosci Nanotechnol*, 7: 757-62.

Tamura, N., L. K. Doolittle, R. E. Hammer, J. M. Shelton, J. A. Richardson, and D. L. Garbers. 2004. 'Critical roles of the guanylyl cyclase B receptor in endochondral ossification and development of female reproductive organs', *Proc Natl Acad Sci U S A*, 101: 17300-5.

Thomas, G., P. Moffatt, P. Salois, M. H. Gaumond, R. Gingras, E. Godin, D. Miao, D. Goltzman, and C. Lanctot. 2003. 'Osteocrin, a novel bone-specific secreted protein that modulates the osteoblast phenotype', *J Biol Chem*, 278: 50563-71.

Tingaud-Sequeira, A., J. Forgue, M. Andre, and P. J. Babin. 2006. 'Epidermal transient down-regulation of retinol-binding protein 4 and mirror expression of apolipoprotein Eb and estrogen receptor 2a during zebrafish fin and scale development', *Dev Dyn*, 235: 3071-9.

Tintut, Y., F. Parhami, K. Bostrom, S. M. Jackson, and L. L. Demer. 1998. 'cAMP stimulates osteoblast-like differentiation of calcifying vascular cells. Potential signaling pathway for vascular calcification', *J Biol Chem*, 273: 7547-53.

Touchman, J. 2010. 'Comparative Genomics', *Nature Education Knowledge 3(10):13*.

Varshney, G. K., W. Pei, M. C. LaFave, J. Idol, L. Xu, V. Gallardo, B. Carrington, K. Bishop, M. P. Jones, M. Li, U. Harper, S. C. Huang, A. Prakash, W. Chen, R. Sood, J. Ledin, and S. M. Burgess. 2015. 'High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9', *Genome Res*, 25: 1030-42.

Vartak, S. V., and S. C. Raghavan. 2015. 'Inhibition of nonhomologous end joining to increase the specificity of CRISPR/Cas9 genome editing', *Febs j*, 282: 4289-94.

Veeckman, E., T. Ruttink, and K. Vandepoele. 2016. 'Are We There Yet? Reliably Estimating the Completeness of Plant Genome Sequences', *Plant Cell*, 28: 1759-68.

Vickaryous, M. K., and J. Y. Sire. 2009. 'The integumentary skeleton of tetrapods: origin, evolution, and development', *J Anat*, 214: 441-64.

Vieira, F. A., S. F. Gregorio, S. Ferraresso, M. A. Thorne, R. Costa, M. Milan, L. Bargelloni, M. S. Clark, A. V. Canario, and D. M. Power. 2011. 'Skin healing and scale regeneration in fed and unfed sea bream, Sparus auratus', *BMC Genomics*, 12: 490.

Wake, Marvalee H. 1992. 'Hyman's Comparative Vertebrate Anatomy': 171.

Wei, L., Y. Liu, I. Dubchak, J. Shon, and J. Park. 2002. 'Comparative genomics approaches to study organism similarities and differences', *J Biomed Inform*, 35: 142-50.

Westerfield, M. 2007. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio)* (Univ. of Oregon Press: Eugene).

Wilson, D. E. & Reeder. 2005. 'D. M. Mammal Species of the World: A Taxonomic and Geographic Reference (3rd ed)': 2,142 pp.

Witten, P. E., L. Gil-Martens, A. Huysseune, H. Takle, and K. Hjelde. 2009. 'Towards a classification and an understanding of developmental relationships of vertebral body malformations in Atlantic salmon (Salmo salar L.)', *Aquaculture*, 295.

Wu, P., L. Hou, M. Plikus, M. Hughes, J. Scehnet, S. Suksaweang, R. B. Widelitz, T. X. Jiang, and C. M. Chuong. 2004. 'Evo-Devo of Amniote Integuments and Appendages', *Int J Dev Biol*, 48: 249-70.

Xu, Z., D. Parra, D. Gómez, I. Salinas, Y. A. Zhang, L. von Gersdorff Jørgensen, R. D. Heinecke, K. Buchmann, S. LaPatra, and J. O. Sunyer. 2013. 'Teleost skin, an ancient mucosal surface that elicits gut-like immune responses', *Proc Natl Acad Sci U S A*, 110: 13097-102.

Yu, C. 2014. 'A PCR based protocol for detecting indel mutations induced by TALENs and CRISPR/Cas9 in zebrafish', *PLOS ONE*, 9.

Zhang, Jiaren. 2014. 'Comparative genome analysis of related catfish species', *Dissertation, Auburn University*.

APPENDICES

Table S1. Identification of osteocrin gene in several vertebrate species: latin name, chromosome, location, and accession number.

Species		Chr	Exon	Location	Accession no.
Human	*Homo sapiens*	3	6	191199060 - 191265601	NP_937827.1
Mouse	*Mus musculus*	16	5	27307641 - 27351209	NP_932780.1
Opossum	*Monodelphis domestica*	7	6	255933296 - 255966671	XP_007502518.1
Cow	*Bos taurus*	1	5	76685699 - 76721920	NP_001092405.1
Sheep	*Ovis aries*	1	5	194190931 - 194227927	XP_004003087.1
Dolphin	*Tursiops truncatus*	Un	4	71296553 - 71323544	XP_004325977.2
Sunda pangolin	*Manis javanica*	Un	6	5771-41077	XP_017527677.1
Orangutan	*Pongo abelii*	3	3	195063815 - 195121622	NP_001127550.1
Panda	*Ailuropoda melanoleuca*	Un	4	750071 - 789685	XP_019658894.1
Guinea pig	*Cavia porcellus*	Un	6	55487836 - 55536523	XP_003477138.2
Turtle	*Chelonia mydas*	JH212904.1	3	3087833 - 3099787	XP_007057892.1
Chicken	*Gallus gallus*	9	5	13760518 - 13772571	XP_015146889.1
Frog	*Xenophus tropicalis*	5	4	99128967 - 99171883	XP_017949518.1
Amazon molly	*Poecilia formosa*	Un	5	380544 - 384800	XP_016536480.1
Shortfin molly	*Poecilia mexicana*	Un	5	57978 - 62214	XP_014868589.1
Guppy	*Poecilia reticulata*	Un	6	959382 - 966941	XP_008400274.1
Minnow	*Cyprinodon variegatus*	Un	3	37222 - 38040	XP_015260777.1
Mummichog	*Fundulus heteroclitus*	Un	7	32595 - 41336	XP_012705304.1
Damselfish	*Stegastes partitus*	Un	5	59125 - 65540	XP_008285446.1
Yellow croaker	*Larimichthys crocea*	Un	7	832190 - 858849	XP_019133074.1
Tilapia	*Oreochromis niloticus*	LG14	5	2510080 - 2514656	XP_005474946.1
Zebra Mbuna	*Maylandia zebra*	Un	5	1668237 - 1674003	XP_012778372.1
Moutbrooder	*Haplochromis burtoni*	Un	6	974899 - 982241	XP_014191347.1
Nyererei	*Pundamilia nyererei*	Un	5	2135800 - 2140779	XP_005737519.1
Salmon	*Salmo salar*	ssa20	5	483515 - 523652	XP_014015462.1
Pike	*Esox lucius*	Un	2	15903 - 17501	XP_019899954.1
Herring	*Clupea harengus*	Un	2	99404 - 101154	XP_012673157.1
Zebrafish	*Danio rario*	15	5	36370925 - 36390308	NP_001313338.1
Mexican tetra	*Astyanax mexicanus*	KB871758.1	5	142566 - 161390	XP_007233576.1
Carp	*Cyprinus carpio*	30	4	15085546 - 15090280	XP_018973289.1
		Un	5	1108777-1115147	XP_018954743.1
Pleco	*Pterygoplichthys pardalis*	scale255099	4	4556 - 7735	NP_001313338.1

Elephant shark	*Challorhinchus milii*	Un	5	281 - 4353	XP_007882804.1
Coelacanth	*Latimeria chalumnae*	JH126687.1	5	1789788 - 1799995	XP_005992527.1
Spottted Gar	*Lepisosteus oculatus*	LG14	5	19280042 - 19291357	XP_006637936.1

Tabel S2. Structure of the gene used in syntenic analysis among scaleless fish: Geminin coil-coil domain containing (*gmnc*)

Species	Latin name	Chr	Exon	Location	mRNA (bp)	AA	Domain		Accession no. / ID	Copy	Scale Type
							slc8a1	*gmnc*			
Channel catfish	*Ictalurus punctatus*	17	9	13431958 - 13454955	2003	432		58-431	XM_017490550.1	1	Scaleless
Walking catfish	*Clarias bratachus*	scaffold311_size522883	9	97400 - 136940	1299	432		145-384	XP_017346039.1	1	Scaleless
Blue catfish	*Ictalurus furcatus*	tig00001883	8	214408 - 235758	1395	464		3-363	XP_017346039.1	1	Scaleless

Tabel S3. Structure of the gene used in syntenic analysis among scaleless fish: mannan-binding lectin serine peptidase 1 (*masp1*)

Species	Latin name	Chr	Exon	Location	mRNA (bp)	AA	Domain				Accession no./ ID	Copy	Scale Type
							Sperma-dhesin	*EGF*	SCR domain	Trypsin like serine protease			
Tetraodon	*Tetraodon rubripes*	Un_random	11	92,299,497-92,303,216	2238	746	8-282	125-166	358-442	449-746	ENSTNIT00000014330.1	1	BDP
Stickleback	*Gasterosteus aculeatus*	groupI	12	500,828-510,359	2136	712	3-271	114-156	273-404	424-712	ENSGACT00000006370.1	1	BDP

Tabel S4. Structure of the gene used in syntenic analysis among scaleless fish: fibroblast growth factor 12 (*fgf12*)

Species	Latin name	Chr	Exon	Location	mRNA (bp)	AA	Domain	Accession no.	Copy	Scale Type
							cytokine			
Tetraodon	*Tetraodon rubripes*	15		5,717,166-5,734,869	732	243	66-220	ENSTNIG00000015272.1	1	BDP
Stickleback	*Gasterosteus aculeatus*	groupIII		2,197,146-2,224,845	729	242	65-219	ENSGACG00000013900	1	BDP

**BDP: Bony Dermal Plate*

Tabel S5. The symbol and description of the gene used in the syntenic analysis.

Gene symbol	Gene description	Gene symbol	Gene description
Abcg1	ATP binding cassette subfamily G member 1	**Hsp70**	heat shock protein 70
Acap2	ArfGAP with coiled-coil, ankyrin repeat and PH domains 2	**Il1rap**	interleukin 1 receptor accessory protein
adcyap1	adenylate cyclase activating polypeptide	**Ilkap**	ILK associated serine/threonine phosphatase
Arrb1	arrestin beta 1	**Irs1**	insulin receptor substrate 1
Atg16l1	autophagy related 16 like 1	**Iws1**	IWS1, SUPT6H interacting protein
Atg16l2	autophagy related 16 like 2	**Kirrel3l**	kin of IRRE like 3 like
Atp13a4	ATPase type 13A4	**Leprel1**	leprecan-like 1
Atp13a5	ATPase type 13A5	**Ltbp4**	latent transforming growth factor beta binding protein 4
atp2a1	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 1	**Masp1**	mannan binding lectin serine peptidase 1
Ccdc105	coiled-coil domain containing 105	**Mb21D2**	Mab-21 domain containing 2
Ccdc50	coiled-coil domain containing 50	**Nphs1**	NPHS1, nephrin
Ccser2	coiled-coil serine rich protein 2	**Nyap2a**	neuronal tyrosine-phosphorylated phosphoinositide-3-kinase adaptor 2a
Cldn1	claudin 1	**Ostn**	Osteocrin
Cldn15	claudin 15	**P3h2**	prolyl 3-hydroxylase 2
Cldn15lb	claudin 15-like b	**Parl**	presenilin associated rhomboid-like
Cldn16	claudin 16	**Parla**	presenilin associated, rhomboid-like a
Clip3	CAP-Gly domain-containing linker protein 3	**Plekhg2**	pleckstrin homology and RhoGEF domain containing G2
cmn	calymmin	**Rhbdd1**	rhomboid domain containing 1
Col4a3	collagen type IV alpha 3 chain	**Rnf169**	ring finger protein 169
Col4a4	collagen type IV alpha 4 chain	**Rnh1**	ribonuclease/angiogenin inhibitor

			1
Col4a5	collagen type IV alpha 5 chain	**Sagb**	streptolysin associated protein SagB
Col6a4	collagen, type VI, alpha 4	**Slc8a1**	sodium/calcium exchanger 2
CrfB16	cytokine receptor family member B16	**Spcs2**	signal peptidase complex subunit 2
Dgkd	diacylglycerol kinase delta	**SST**	somatostatin
Dll3	delta-like 3	**SST1.1**	somatostatin 1, tandem duplicate 1
Esrrd	estrogen-related receptor delta	**SST1.1**	somatostatin 1, tandem duplicate 2
Exoc3l2b	exocyst complex component 3-like 2b	**tim10**	protein transporter TIM10
F10	coagulation factor 10	**Timm50**	translocase of inner mitochondrial membrane 50
Fam219b	family with sequence similarity 219 member B	**Tmem207**	transmembrane protein 207
Fcgbp	Fc fragment of IgG binding protein	**Tp63**	tumor protein p63
Fgf12	fibroblast growth factor 12	**Tprg1**	tumor protein p63 regulated 1
Gmnc	geminin coiled-coil domain containing	**Trnag-ucc**	transfer RNA glycine (anticodon UCC)
Gpr171	G protein-coupled receptor 171	Uts2b	urotensin 2B
Hrasls	HRAS like suppressor	Wdr74	WD repeat domain 74
Hsd17b1	hydroxysteroid 17-beta dehydrogenase 1	Zan	zonadhesin (gene/pseudogene)

Figure S1. Multi-sequence alignment of Osteocrin. The alignment was generated using amino acid from human, frog, chicken, turtle, zebrafish, pleco, coelacanth, salmon, and carp using Clustal Omega. Putative dibasic cleavage sites are highlighted with blue boxes as indicated (C1 and C2). Two regions are boxes which homolog to the NP-like motif (NM1 And NM2). NPs considered essential in binding to NPR-C receptor which involved in osteogenesis and chondrogenesis. All the vertebrate models exhibit two NPs sites except for the bony dermal pleco which only has one NPs site.

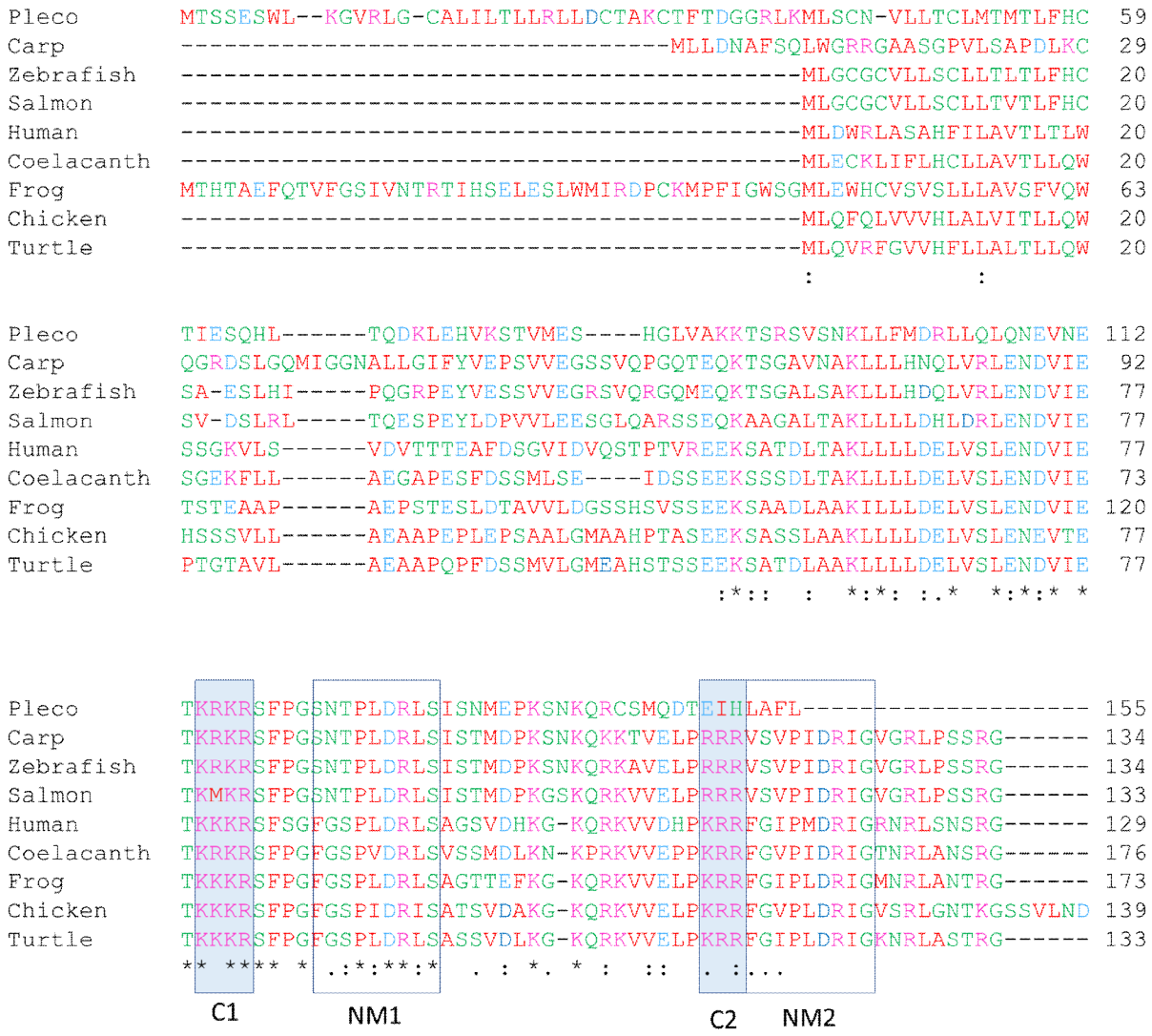


Figure S2. Protein alignment of *ostn* copy number 1 (149aa) and 2 (134aa) in common carp *Cyprinus carpio*. Putative dibasic cleavage sites are highlighted with blue boxes as indicated. Two regions are boxes which homolog to NPs (NM1 And NM2).

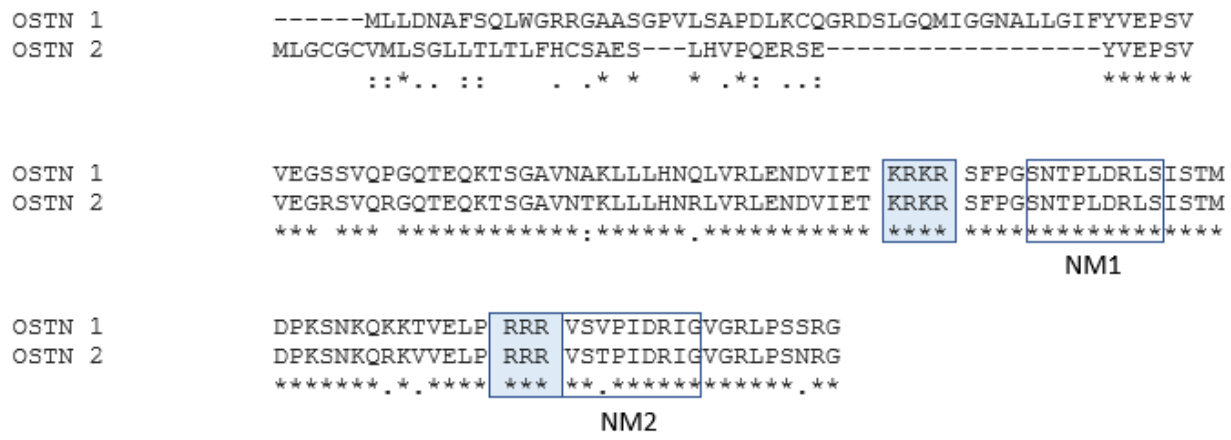


Figure S3. Three gRNA of the osteocrin gene was used along the cas9 in CRISPR/Cas9 system. The target sequences span along 405 bp coding sequence (123 - 527 bp) in exon 3. The remarks were: uppercase/lowercase as exon boundary, red block as gRNA in sense strand, and purple block as gRNA in the anti-sense strand (A). Two primers used for detecting the indels using fluorescence PCR (B).

A.

ATGCTGGGCTGTGGATGTGTGCTGCTCTCTTGTCTGCTGACACTGACCTTGTTC
CACTGCAGTGCCGAAAGCCTTCACATACCGCAAGGAAGACCAGAGtatgtggaa
tcttctgtcgtggagggccgcagtgtccagcgaggccaaatggagcagaaaacctcaggagcactga
gtgctaaactcctcctgcacgaccagttggttcggctggagaacgatgtcattgagaccaaaaggaa
acgaagcttcccctggatccaacacgccactggaccgtctgtcaatcagcaccatggaccctaaaagc
aacaagcagagAAAGGCTGTTGAACTGCCGCGGCGGCGAGTCAGCGTGCCGAT
TGACCGGATTGGTGTAGGCCGCCTTCCCAGCAGCCGAGGATAA

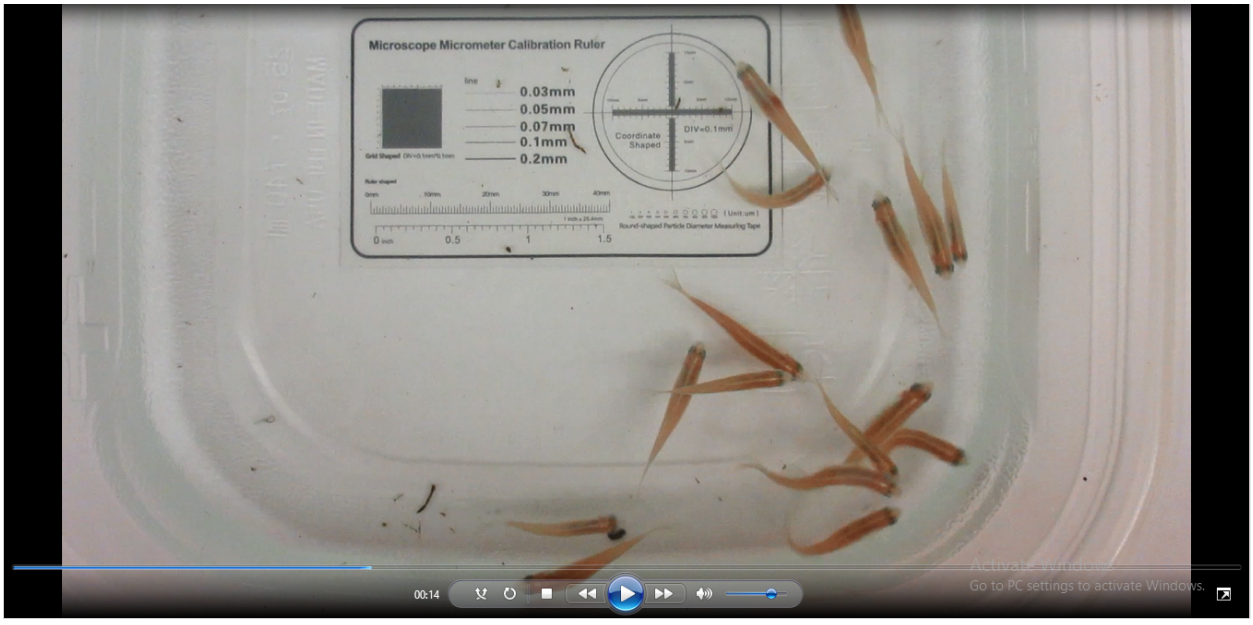
UPPERCASE/lowercase = exon boundary
red = gRNA (sense strand) red = PAM "NGG"
purple = gRNA (anti-sense strand) purple = PAM "CCN"

B.

Ostn_primer1	F) 5'	tttgtgggtttttgactaatttga
	R) 5'	ttccaaattagaagaacagcattt
Ostn_primer2	F) 5'	tgtctaccagctgaaaccatt
	R) 5'	agggtccatggtgctgattg

Video S1. **Ventral view of Wt zebrafish**. The Wt zebrafish exhibited normal Swimming modes and normal lateral bone (spinal cord). The total length showed normal size variant in those Wt adult zebrafish.

Video S2. **Ventral view of mutated (*Ostn* $^{-/}$) zebrafish**. The mutant exhibited random swimming modes and deformities in their lateral bone (spinal cord). The deformity spread at 19-26th of bone segments. The mutant total length showed more in length variant compare to the Wt.



Video S3. **Lateral view of individual Wt zebrafish**. The Wt zebrafish showed normal pigment and thickness.

Video S4. **Lateral view of individual mutated (*Ostn* $^{-/-}$) zebrafish**. The mutant exhibited pigment dispersion, thin scale, and partial scale lost (brighter color in skin and scales). The mutant also showed slower growth compared to the Wt.