Gene 824 (2022) 146441

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Cloning, tissue distribution and mRNA expression of type I collagen alpha 1 gene from Chu's croaker (*Nibea coibor*)

Hua Rong [a,b], Fan Lin [c], Lijun Ning [a], Kun Wu [a], Baojia Chen [a], Jia Zheng [a], Samwel Mchele Limbu [d], Xiaobo Wen [a,*]

[a] College of Marine Sciences, South China Agricultural University, Guangzhou 510642, China
[b] College of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, China
[c] Guangdong Provincial Key Laboratory of Marine Biotechnology, Institute of Marine Sciences, Shantou University, Shantou 515063, China
[d] Department of Aquaculture Technology, School of Aquatic Sciences and Fisheries Technology, University of Dar es Salaam, P. O. Box 60091, Dar es Salaam, Tanzania

ARTICLE INFO

Edited by John Doe

Keywords:
Nibea coibor
COL1A1 gene
cDNA cloning
Collagen synthesis
Chu's croaker swim bladder

ABSTRACT

The demand for collagen has been increasing over years due to its wide application in food, cosmetics and biomedicine industries. The synthesis of collagen protein in fish depends on instructions provided by collagen, type I, alpha 1 (COL1A1) gene. However, cloning, tissue distribution and mRNA expression of COL1A1 gene in a gel-producing Chu's croaker (*Nibea coibor*) is currently unknown. This study cloned the cDNA of COL1A1 gene (GenBank accession number: MK641512) from six *N. coibor* fish. The distribution and mRNA expression pattern of COL1A1 was analyzed in eight tissues of *N. coibor*. The COL1A1 cDNA had a full length of 6130 bp and contained a 4344 bp open reading frame (ORF) encoding a polypeptide of 1448 amino acids. The homology of *N. coibor* COL1A1 amino acid had 98% similarity with *Larimichthys crocea*, indicating conservatism with other members in same family (*Sciaenidae*). The deduced polypeptide contained the same signal peptides, C-propeptide and N-propeptide domains, and triple helix domains, which are the characteristics of type I collagen in vertebrates. The mRNA of COL1A1 gene was expressed significantly higher in the spine of *N. coibor* than in all other tissues ($P < 0.05$), followed by swim bladder, skin and scales. The swim bladder had higher collagen and hydroxyproline contents than other tissues, followed by spine >, scales > and > skin ($P < 0.05$). Our study successfully cloned the COL1A1 gene from *N. coibor* for the first time. The COL1A1 gene contained all the features of collagen pro-α1(I) chain proteins, and shared high homology with other marine teleost. COL1A1 gene in *N. coibor* is highly expressed in spine and swim bladder, consistent with collagen distribution. Our study contributes to better understanding on collagen biosynthesis in *N. coibor* tissues for various industrial uses.

1. Introduction

Collagen is the largest and most abundant protein constituting about 250 to 350 g/kg of the total proteins (Sun et al., 2017). Collagen is widely distributed in most organs in vertebrates (Gelse et al., 2003). It is a huge family with a wide variety of highly complex structures, from molecular structure, supramolecular structure, to tissue distribution and function (Gordon and Hahn, 2009). Collagen has multiple uses including consumption as a health food in some countries such as Singapore, India and Japan (Pal and Suresh, 2016) and Hong Kong and southern China (Tuuli et al., 2016). It is also applied in cosmetics and biomaterials for cell scaffold, wound dressing and soft tissue augmentation (Avila Rodríguez et al., 2018). Moreover, collagen plays biological functions such as antioxidant, antimicrobial activity, anti-hypertensive and inhibitory activity and lipid-lowering effect (Karim and Bhat, 2009). Therefore, collagen is an important protein playing major roles in immunity, mechanical and physiological functions for maintaining health of skin, bones and joint.

Currently, different types of collagen have been cloned, separated and identified. They comprise 28 new types of collagen, which have

Abbreviations: COL1A1, Collagen, type I, alpha 1; RACE, Rapid amplification of cDNA ends; pro-α1(I), Type I procollagen α1; qRT-PCR, Quantitative real-time PCR; VWFC, Von Willebrand factor; NC1, C-terminal non-collagen; RGD, Arginine-glycine-aspartate; NIT, Asparagine-isoleucine-threonine; GPP, Glycine-proline-proline.

* Corresponding author.
E-mail address: wenxbo@stu.edu.cn (X. Wen).

https://doi.org/10.1016/j.gene.2022.146441
Received 15 November 2021; Received in revised form 20 February 2022; Accepted 18 March 2022
Available online 23 March 2022
0378-1119/© 2022 Elsevier B.V. All rights reserved.

been identified through gene cloning and other methods (Banerjee et al., 2017; Bella and Hulmes, 2017). Among the members of the collagen family, type I collagen is the most widely distributed in vertebrates. It has been described from sponges to humans (Exposito et al., 2010) in various tissues (Ricard-Blum, 2011; Rong et al., 2020a). Similarly, different collagen types have also been isolated and identified from fish including *Catla catla* and *Cirrhinus mrigala* (Mahboob, 2015), red drum fish (*Sciaenops ocellatus*) (Chen et al., 2016) and tilapia (*Oreochromis* sp.) (Huang et al., 2016). This information shows that collagen is conserved in all multicellular animals.

Previous studies reported that, type I collagen is widely distributed in dermis, bones, scales, swim bladder and muscle. In these tissues, type I collagen provides the scaffolding and guide cells to migrate, proliferate and differentiate (Stover and Verrelli, 2010; Bhagwat and Dandge, 2018), while a few types II and XI are mostly found in cartilage and spinal cord, and type V exist mainly in muscle (Exposito et al., 2010). The fish type I collagen is one of the earliest known and most clearly understood by scientists (Kimura et al., 1988). It is composed of two α1 and one α2 peptide chains, which are coded by type I collagen α1 (COL1A1) and type I collagen α2 (COL1A2) genes (Rong et al., 2019). The human COL1A1 gene has been studied for a long time (Solomon et al., 1984). It has been revealed that, the human COL1A1 gene is located on chromosome Chr 17q21.31-22.05 with a size of 18 kb and contains 51 exons. Similarity, the human COL1A2 is located on Chr 7q21.3-22.1 has 52 exons with a size of more than 38 kb, and encodes a monomer α peptide chain contains 1464 amino acid residues (Retief et al., 1985). At present, the type I collagen precursor alpha chain genes of many teleost fish have also been reported in rainbow trout (*Oncorhynchus mykiss*) (Saito et al., 2001), zebrafish (*Danio rerio*) (Gistelinck et al., 2018), skate (*Raja kenojei*) (Hwang et al., 2006) and grass carp (*Ctenopharyngodon idellus*) (Yu et al., 2014). These studies highlight that, type I collagen is abundant in many fish species.

The Chu's croaker (*Nibea coibor*) is regarded as an excellent food fish. It contains a large swim bladder, which is rich in collagen used to produce high-grade fish glue (Kite, 2006). The fish glue is considered a tonic with a high market value and demand in Southeast Asia, particularly in Hong Kong and southern China (Tuuli et al., 2016). Consequently, some studies have been carried out on the collagen metabolism and nutritional regulation in *N. coibor* at early stages (Rong et al., 2019; Lin et al., 2020; Rong et al., 2020b). It has been established clearly that, collagen synthesis and deposition in the swim bladder depends on dietary supplementation of proline (Pro) or its derivative, hydroxyproline (Hyp) (Wei et al., 2016; Rong et al., 2019; Lin et al., 2020; Rong et al., 2020a; Rong et al., 2020b). However, cloning, tissue distribution and mRNA analysis of the COL1A1 gene in *N. coibor* has not been studied. This knowledge is important in order to understand the mechanisms for increasing the production of collagen required in the market for the various uses.

Therefore, the present study cloned and analyzed the full-length cDNA of type I procollagen α1 [pro-α1(I)] chain of *N. coibor* for the first time. We further analyzed the COL1A1 gene tissue distribution and expression among eight tissues in *N. coibor*. Moreover, we also determined the hyp and collagen contents and tissue distribution among eight tissues. The results obtained provide scientific foundation for exploring the molecular mechanism of fish collagen synthesis and production for various applications.

2. Materials and methods

2.1. Experimental fish

Six *N. coibor* (body weight 156.20 ± 5.50 g) were randomly selected from the experimental fish used for the study conducted by Lin et al. (2020). The fish used were initially obtained from a local marine fish hatchery (Raoping, Guangdong, China). The obtained fish were cultured alive at the Nan'Ao Marine Biology Station (NAMBS, Shantou University, Shantou, China). The fish were collected and anaesthetized with eugenol (1:10,000, Shanghai Reagent), then slaughtered by decapitating with scissors. The liver, muscle, intestine, swim bladder, skin, scales, fin and spine were collected and divided into two parts. The first part was immediately placed into liquid nitrogen for transporting to the laboratory, where they were frozen at −80 °C for RNA extraction. The second part was stored at −20 °C for collagen content determination.

2.2. Total RNA isolation and reverse transcription

The total RNA from muscle tissue of *N. coibor* was extracted by using Trizol (Invitrogen™, USA) reagent according to the manufacturer's instructions. The RNA sample was treated with DNase I (Takara, Dalian, China) to eliminate the contamination of genomic DNA. The integrity of purified RNA was determined by 1.2% (w/v) agarose gel electrophoresis. The final concentration of RNA was determined by measuring the OD260/280 by using a Nanodrop®ND-2000 spectrophotometer (Thermo Scientific NanoDrop, USA). Reverse transcription of the total RNA was performed by using SMARTer™ rapid amplification of cDNA ends (RACE) cDNA Amplification Kit (Clontech, USA) to create RACE-ready first-strand cDNA.

2.3. Molecular cloning of COL1A1 cDNA

The degenerate primers of *N. coibor* COL1A1 were designed based on sequences of COL1A1 from *Danio rerio* (GenBank ID: NM_199214.1), *Oncorhynchus mykiss* (GenBank ID: NM_001124177.1), *Carassius auratus* (GenBank ID: LC382451.1) and *Hypophthalmichthys molitrix* (GenBank ID: KJ848743.1) by using the Primer Premier 5.0 software (Premier Biosoft International, Palo Alto, CA) (Table 1). The primers were used for amplifying fragments of *N. coibor* putative COL1A1 cDNA. The polymerase chain reaction (PCR) amplification was performed in a total volume mixture of 25 μl, containing 12.5 μl 2 × Taq PCR Master Mix (TianGen, Beijing, China), 1.5 μl of cDNA template, 0.75 μl (0.2 mmol) of each primer pairs (Rs-C1A1F and Rs-C1A1R) and 9.5 μl of double distilled water (ddH_2O). The reaction conditions were as follows: 95 °C for 3 min, 35 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C for 30 s and extension at 72 °C for 1 min, followed by a final extension at 72 °C for 10 min. The bands of desired sizes were excised and purified using a gel Extraction Kit (TianGen) by following manufacturer's guidelines. The purified PCR products were then ligated with pMD19-T Vector (Takara), and transformed into competent *Escherichia coli* cells. Positive recombinant clones were identified by PCR screening with M13R and M13F primers, and subsequently sent to a commercial company (BGI, Shenzhen, China) for sequencing.

The full-length COL1A1 cDNA was amplified through 5′ and 3′ RACE PCR by using the SMARTer™ RACE cDNA Amplification Kit (Clontech) using touchdown PCR (first round PCR) and nested PCR (second round PCR) strategies. The first round PCR included 8 μl of deoxyribose nucleotide triphosphate (dNTPs) (2.5 mM), 5 μl of 5 × PCR buffer, 1.5 μl of C1A1-31 or C1A1-51 primers (10 μM), 5 μl of adapter specific primer (UPM), 0.5 μl of Taq polymerase (5 U μl^{-1}), 1.5 μl of cDNA and 28.5 μl DNA free water. The PCR conditions were as follows: 95 °C for 3 min followed by 35 cycles at 95 °C for 30 s, 64 °C for 30 s, 72 °C for 1 min and a final extension for 10 min at 72 °C. The product of first round PCR was used as template for the subsequent nested PCR, which was performed in a 50 μl reaction volume mixture including 25 μl of 2 × Taq PCR MasterMix (TianGen), 1 μl of C1A1-32 or C1A1-52 primers, 1 μl of universal primer (NUP), 0.5 μl of template, and 22.5 μl DNA free water. The nested PCR conditions consisted of an initial denaturing step at 95 °C for 3 min, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at 62 °C (3′RACE) or 61 °C (5′RACE) for 30 s, extension at 72 °C for 2 min, followed by a final extension at 72 °C for 5 min. The 5′ and 3′ RACE products were subcloned and sequenced as described above.

Table 1
The primer sequence information used for the cDNA cloning and expression analysis of COL1A1.

Purpose	Gene name	Primer number	Primer sequence (5′ to 3′)
RACE-PCR	COL1A1	R1-C1A1F	CCTTCTGGAGAAGTTGGAAAGC
		R1-C1A1R	CGATGAAACCAAGGTCGAATCC
		R2-C1A1F	TGCTGGATCTCCTGGTAAGG
		R2-C1A1R	TTGCCCTCTGCTCTGATCTC
		R3-C1A1F	ATCTGCGAGGACACAACC
		R3-C1A1R	GAATCTCCCTTAGCACCATCC
		R4-C1A1F	TGGTGGCTACGATGACAAGT
		R4-C1A1R	TTCCAACGGGACCCTGAG
		R5-C1A1F	TGGTTCTGATGGCTCTGC
		R5-C1A1R	GCTCTGAGTGGGAGTTACGC
	Universal primer	Oligo-	AAGCAGTGGTATCAACGCAGAGTACNNNNN
		UPM (long)	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT
		UPM (short)	CTAATACGACTCACTATAGGGC
		NUP	AAGCAGTGGTATCAACGCAGAGT
		M13F	GTAAAACGACGGCCAGT
		M13R	ACAGGAAACAGCTATGAC
5′/3′RACE	COL1A1	C1A1-51	GGCTCTCCTCTTGGTCCTCTCTTGC
		C1A1-52	GGGTTCCGTTCTCACCAGGGGCTC
		C1A1-31	AGCGGCGAGTACTGGATTGACCC
		C1A1-32	GGCGGATTCGACCTTGGTTTCA
qRT-PCR	COL1A1	Q-COL1A1F	AGACCTGCGTGACTCCCA
		Q-COL1A1R	AGCCCTCGCTGCCATACT
	β-actin	β-actinF	GGTTACTCCTTCACCACCACAG
		β-actinR	TCCGTCGGGCAGCTCATA

N = undisclosed base in the proprietary SMARTer oligo sequence.

2.4. Characteristics and bioinformatics analyses of COL1A1 cDNA

The full-length cDNA sequence acquired was spliced to obtain the core sequence of COL1A1 cDNA by using DNAMAN software. A homology analysis was conducted by using BLASTX (http://blast.ncbi.nlm.nih.gov/). Furthermore, the amino acid sequence was deduced by using the ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi). The molecular mass and the theoretical isoelectric point (pI) of the collagen protein was predicted by using Compute pI/Mw tool (http://cn.expasy.org/tools/pi_tool.html). The homology alignment of the COL1A1 amino acid sequence was deduced by using ClustalX2 (Borland company, Scotts Valley, CA, USA). The obtained homology was compared with amino acid from other species in GenBank database. A phylogenetic tree was constructed by the neighbor-joining method by using the MEGA 4.0 software (Borland company, Scotts Valley, CA, USA).

2.5. Determination of expression of COL1A1 gene in the different tissues of N. coibor by quantitative real-time PCR (qRT-PCR)

To explore the expression of COL1A1 gene in different tissues, quantitative real-time PCR (qRT-PCR) was performed by using the SYBR® Premix Ex Taq™ II Kit (Takara) in a LightCycler® 480 (Roche Diagnostics, Switzerland) for the liver, muscle, intestine, swim bladder, skin, scales, fin and spine tissues of *N. coibor*. Total RNA from each tissue was extracted by using Trizol (Invitrogen™) as described above. A total amount of 500 ng RNA of each tissue was reverse-transcribed into cDNA by using a TransScript® One-Step gDNA Removal and cDNA Synthesis SuperMix Kit (TranGen) with gDNA Eraser (Takara) according to the manufacturer's instructions. The specific primers were designed by using the Primer Premier 5.0 software (Table 1). The qRT-PCR assays were conducted by using 2 × SYBR® Premix Ex Taq™ II Kit (Takara) in an optimized 20 μl reaction volume mixture, containing 10 μl of SYBR® Premix Ex Taq™ II, 1.5 μl of the three-fold diluted cDNA, 0.75 μl (10 mM) each for forward and reverse primers, and 7 μl of RNAse free water. The qRT-PCR was performed in 96-well optical plates in a Light Cycler®480 (Roche) under the following conditions: 95 °C for 30 s (pre-incubation), followed by 40 cycles of 95 °C for 5 s, 58 °C for 20 s, 72 °C for 15 s (amplification), 95 °C for 5 s (extension) and 60 °C for 1 min (final extension). Each analysis was conducted in duplicate. The relative expression of target gene was normalized by using β-actin. The expression of COL1A1 gene was calculated by using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

2.6. The analysis of Hyp and collagen contents distribution in various tissues

The distribution of collagen in various *N. coibor* tissues was determined by Hyp content method. Hyp is an amino acid that is unique to collagen and is traditionally used to quantify this protein (Kafienah and Sims, 2004). The content of Hyp is 12.5% of the collagen protein content in connective tissue (AOAC, 1995). The Hyp content in liver, muscle, intestine, swim bladder, skin, scales, fin and spine was analyzed by using Diagnostic Reagent Kits (No. A030-3 and No. A030-2) purchased from Nanjing Jiancheng Bioengineering Institute (Nanjing, China) according to the manufacturer's instructions as described recently in our previous study (Rong et al., 2019). The absorbance (OD value) for each sample was measured by using the Infinite® Pro 200 multi-functional micro porous plate detector (Tecan, Switzerland). Finally, the quantity of hydroxyproline was converted into collagen, by multiplying with a factor of 12.5%. The results obtained were expressed as mg g^{-1} protein.

2.7. Statistical analyses

The data obtained were tested for normality and homogeneity of variances by using Shapiro-Wilk and Levene's tests, respectively. The mean mRNA expression of COL1A1 gene and collagen distribution among the different tissues were tested for statistical differences by using One-way analysis of variance (ANOVA). Specific significant differences among means for the various tissues were further compared by using Duncan's multiple range tests. Results with $P < 0.05$ were considered significant different. Results are presented as means ± standard error of means (SEM). All statistical analyses were performed by using Statistical Package for Social Science (SPSS) version 20 for Windows (SPSS, IBM, Armonk, NY, USA).

3. Results

3.1. The sequence analysis of COL1A1 cDNA

The study successfully cloned the COL1A1 gene (GenBank accession

number: MK641512) from the *N. coibor* muscle by using RACE technology. The complete cDNA sequence of this gene and the deduced amino acids are shown in Fig. 1. The sequence analysis showed that the full length of the cDNA was 6130 bp, including 1632 bp 3′untranslated region, 154 bp 5′untranslated region and 4344 bp ORF, encoding a total of 1447 amino acids. The protein structure and hydrophobicity analyses showed that, the predicted molecular weight of the COL1A1 gene protein was 136.596 kDa, and the isoelectric point was 5.44 (Supplementary Fig. 1). The total number of negatively charged residues (Asp + Glu) were 143, while the total number of positively charged residues (Arg + Lys) were 127. The calculated lipid index was 39.09, the overall mean hydrophilicity value was −0.688, and the instability index (II) was 24.58, which indicated that the protein was stable.

The COL1A1 protein contained a highly conserved C-terminal non-collagen domain (NC1) unique to the large family of extracellular matrix glycoproteins, and approximately 1000 amino acid residues. The



Fig. 1. Nucleotide and deduced amino acid sequences of COL1A1 from *N. coibor*.



Fig. 1. (*continued*).

long collagen domain had multiple repeated glycine-proline-proline (GPP) tripeptide characteristics. In addition, the COL1A1 protein contained a type of repeated von Willebrand factor (VWFC) domain, which contained 70 amino acids and covered 10 conserved cysteines (Cx(2,3)-C-{CG}-Cx(6,14)-Cx(3,4)-Cx(2,10) -Cx(9,16)-CCx(2,4)-C) (Supplementary Fig. 2). Among them, cysteine residues formed disulfide bonds between/intra-chain. The COL1A1 protein contained a triple helix domain, which included 23 GPP repeat sites, 4 arginine-glycine-aspartic acid (RGD) cell adhesion sites and two internal coupling points (GIKGHR, GMKGHR). The COL1A1 protein also formed eight C-propeptide disulfide bonds within or between chains, a glycosyl binding site (NIT) and 1 phosphorylation (KRGP) site.

A single underlined letter indicates the VWFC domain signature site with consensus pattern: C-x(2,3)-C-{CG}-C-x(6,14)-C-x(3,4)-C-x(2,10)-C-x(9,16)-C-C-x(2,4)-C. arrows indicate cleavage sites of N-propeptides and C-propeptides, highlighted Gly-Pro-Pro repeats in the triple helical domain are labelled in brace, the putative intermolecular crosslinking sites are labelled using wavy lines and bold text, the arginine-glycine-aspartic acid (RGD) sites that represent the potential cell adhesion sites are marked in bold red, asparagine-isoleucine-threonine (NIT) with dotted sequence indicates the attachment site of putative carbohydrate, Fibrillar collagen C-terminal non-collagenous (NC1) domain profile (MATRIX) are double underlined, cysteine residues in C-propetides are highlighted, the bold and italic letters indicate the phosphorylation sites of cAMP and cGMP-dependent protein kinase, and the asterisk indicates the stop codon.

3.2. Multiple sequences alignment and phylogenetic analysis

The COL1A1 polypeptide sequences of *N. coibor* were compared with those of other fish species by using ClustalX2 software. The deduced polypeptide sequences displayed high percentage of identity with other fish species such as *Larimichthys crocea* (98.00%), *Astatotilapia calliptera* (91.09%), *Seriola dumerili* (93.02%), *Oreochromis niloticus* (92.19%), *Amphiprion ocellaris* (92.67%), *Parambassis ranga* (91.98%), *Mastacembelus armatus* (91.78%), *Anabas testudineus* (92.26%), *Labrus bergylta* (90.67%), and *Cynoglossus semilaevis* (90.32%), to name a few with more than 90% similarity (Table 2).

To reveal the relationship between *N. coibor* COL1A1 amino acid with other vertebrate members of the collagen family, multiple sequence comparison of amino acids sequences was performed by using the ClustalX2 software (Fig. 2). The results showed that the *N. coibor* COL1A1 amino acid sequence had the highest similarity to those of *L. crocea*. Furthermore, multi-sequence comparison showed that the entire fibrillar collagen NC1 domain had a high degree of similarity with other vertebrates.

The phylogenetic analysis of the COL1A1 amino acid sequence was performed by using MEGA version 4.0 through the neighbour-joining method (Fig. 3). The phylogenetic analysis showed that the COL1A1 evolutionary tree formed two distinct branches. The first branch was mainly for bony fishes, including *Trachinotus ovatus*, *Carassius auratus*, *Danio rerio*, *Ctenopharyngodon idella*, *Oncorhynchus mykiss*, *Anguilla japonica*, *Oryzias latipes*, *Oreochromis niloticus*, *Larimichthys crocea* and *Nibea coibor*. The second branch was for higher vertebral animals, such as *Gallus gallus*, *Mus musculus*, *Rattus norvegicus*, *Equus asinus*, *Sus scrofa*,

Table 2
COL1A1 polypeptide sequence identity of *N. coibor* with 30 other fishes.

Matched species	Accession no.	Identity
Larimichthys crocea	XP 019117249.1	98.00%
Labrus bergylta	XP_020511562.1	90.67%
Parambassis ranga	XP_028267212.1	91.98%
Cynoglossus semilaevis	XP_008328838.1	90.32%
Seriola dumerili	XP_022603698.1	93.02%
Kryptolebias marmoratus	XP_017267795.1	88.53%
Anabas testudineus	XP_026202622.1	92.26%
Oryzias latipes	NP_001116390.2	89.70%
Monopterus albus	XP_020441049.1	87.28%
Amphiprion ocellaris	XP_023154805.1	92.67%
Astatotilapia calliptera	XP_026021750.1	91.09%
Xiphophorus maculatus	XP_023204469.1	87.77%
Carassius auratus	BAG72200.1	87.78%
Oreochromis niloticus	NP_001266373.1	92.19%
Danio rerio	NP_954684.1	86.87%
Mastacembelus armatus	XP_026180754.1	91.78%
Hypophthalmichthys molitrix	AIL02135.1	87.71%
Ctenopharyngodon idella	ADK35755.1	88.26%
Electrophorus electricus	XP_026853844.1	86.05%
Oryzias melastigma	XP_024116552.1	89.77%
Oncorhynchus mykiss	NP_001117649.1	83.85%
Astyanax mexicanus	XP_007232286.2	87.56%
Oncorhynchus tshawytscha	XP_024288724.1	84.20%
Oncorhynchus kisutch	XP_020348842.1	84.13%
Pangasianodon hypophthalmus	XP_026772561.1	87.37%
Fundulus heteroclitus	XP_012716616.1	86.87%
Trachinotus ovatus	ATI25516.1	81.66%
Oncorhynchus tshawytscha	XP_024245808.1	83.92%
Acipenser schrenckii	BAR72358.1	80.16%

Canis lupus familiaris, *Felis catus*, *Bos taurus* and *Homo sapiens*. In addition, the amphibian African toad *Xenopus tropicalis* was classified as the second branch (higher vertebrates) between the two branches. The *Nibea coibor* were closely related to *Larimichthys crocea*, which belong to the same family of *Sciaenidae*.

Numbers represent the frequencies with which the tree topology presented was replicated after 1000 bootstrap iterations. The scale bar is at the bottom of the figure. The horizontal branch length is proportional to amino acid substitution rate per site. The shorter the distance, the closer the evolutionary relationship. On the contrary, the longer the distance, the further the evolutionary relationship.

Collagens are a huge family of proteins. Currently, 28 types of collagens have been discovered, and each type of collagen has many subtypes. This study constructed a phylogenetic tree based on the published amino acids sequence of COLsAs from the zebrafish collagen family and the COL1A1 amino acid sequence of *N. coibor*. The *N. coibor* COL1A1 was closer to the zebrafish COL1A1a, followed by COL1A1b (Fig. 4). Therefore, these results confirmed scientifically that the full-length cDNAs we cloned was COL1A1.

3.3. The mRNA expression of COL1A1 gene in the different tissues of *N. coibor*

The results on the mRNA expression of the *N. coibor* COL1A1 gene in various tissues including liver, muscle, intestine, swim bladder, skin, scales, fin and spine are shown in Fig. 5. The expression pattern of the *N. coibor* COL1A1 gene exhibited a widespread distribution in all tissues tested. The mRNA expression of the *N. coibor* COL1A1 gene was significantly higher in spine than all other tissues ($P < 0.05$). Similarly, the swim bladder had higher mRNA expression of the *N. coibor* COL1A1 gene than liver, intestine, skin, scale, muscle and fin ($P < 0.05$). The skin also had significantly higher mRNA expression of the *N. coibor* COL1A1 gene than liver, intestine, scale, muscle and fin ($P < 0.05$). The scales had significantly higher mRNA expression of the *N. coibor* COL1A1 gene than in the liver and intestine ($P < 0.05$). However, the mRNA expression of the *N. coibor* COL1A1 gene was statistically similar among muscle, scale and fine tissues ($P > 0.05$). Equally, the fin, liver, intestine and muscle had statistically comparable mRNA expression of the *N. coibor* COL1A1 gene ($P > 0.05$).

3.4. The distribution of Hyp and collagen contents in various tissues

The results on distribution of Hyp and collagen contents in liver, muscle, intestine, swim bladder, skin, scales, fin and spine are shown in Table 3. As expected, the Hyp and collagen contents exhibited a similar trend and differed among all the measured tissues ($P < 0.05$). The swim bladder had higher Hyp and collagen contents than all other tissues ($P < 0.05$). Similarly, the spine had significantly higher Hyp and collagen contents than muscle, liver, scale, skin, fin and intestine ($P < 0.05$). Likewise, the fish scales had significantly higher Hyp and collagen contents than muscle, liver, skin, fin and intestine ($P < 0.05$). Equally, the fish skin had significantly higher Hyp and collagen contents than muscle, liver, fin and intestine ($P < 0.05$). In similar trend, the fins had significantly higher Hyp and collagen contents than muscle, liver and intestine ($P < 0.05$). The fish muscles also had significantly higher Hyp and collagen contents than liver and intestine ($P < 0.05$). Lastly, the fish intestine had significantly higher Hyp and collagen contents than liver ($P < 0.05$).

4. Discussion

The present study reports the sequence of COL1A1 cDNA cloned from *N. coibor* and its gene tissue distribution. The complete COL1A1 cDNA sequence of *N. coibor* encoded a polypeptide of 1448 amino acids, similar to those reported by previous scholars. For example, the COL1A1 protein sequence for skate had 1463 amino acids (Hwang et al., 2006), while that of grass carp had 1448 amino acids (Yu et al., 2014). The deduced 1448 amino acids possessed fibrillar collagen that had the same signal peptides, C-propeptide and N-propeptide domains, and triple helix domains as in many vertebrates (Dubois et al., 2002; Stover and Verrelli, 2010; Bou-Gharios et al., 2020). The amino acids contained COL1A1 N-propeptide VWFC domain signal site (31–79), triple helix domain (132–893) and C-terminal non-collagen (NC1) domain (921–147). Similarly, the propeptide chain of vertebrate fibrous collagen (I, II and III) α1 and (V) α2, contains a VWFC domain of N-propeptide, a triple-helical domain of about 1000 amino acids and a sugar of C-propeptide basal fiber collagen (COLF1 domain) (Slatter and Farndale, 2015; Pozzolini et al., 2020). These results indicate that, despite collagen being a large protein family with 28 members at the moment, they have common characteristics. The molecular features shared by all members of this family include the highly conserved NC1 domain and the triple-helical domain of about 1000 amino acid residues. The most conserved part of fibrous collagen from invertebrates to vertebrates is the NC1 domain (I-III, V, XI, XXIV and XXVII types) (Exposito et al., 2010; Ricard-Blum, 2011). The existence of the same signal peptides, C-propeptide and N-propeptide domains, and triple helix domains appears to be conserved in collagen families for various functions.

It has been documented that N-propeptide regulates the diameter and morphology of collagen fibres, thereby affecting its structure and function, while the triple helix structure plays an important role in the regulation of mRNA transcription (Brodsky and Persikov, 2005; Koivula et al., 2010). Moreover, the total Gly-Pro-Pro content is the dominating factor influencing the change of collagen thermal stability (Engel and Bächinger, 2005). In the triple helical structure of COL1A1 from *N. coibor*, 23 Gly-Pro-Pro repeat peptides were observed, similar to the triple-helical domain of the skate pro-α1(I) chain (Hwang et al., 2006). These are common structures in many vertebrate species including human (D'Alessio et al., 1988) and *Xenopus* type I collagen alpha 1 (COL1A1) (Goto et al., 2000), suggesting that the COL1A1 repeat peptides provide a thermal stability to collagen. In addition, two staggered coupling sites (GIKGHR, GMKGHR) and four arginine-glycine-aspartate (RGD) cell adhesion sites were observed in the triple helical structure of COL1A1 from *N. coibor*. These staggered coupling domains are necessary

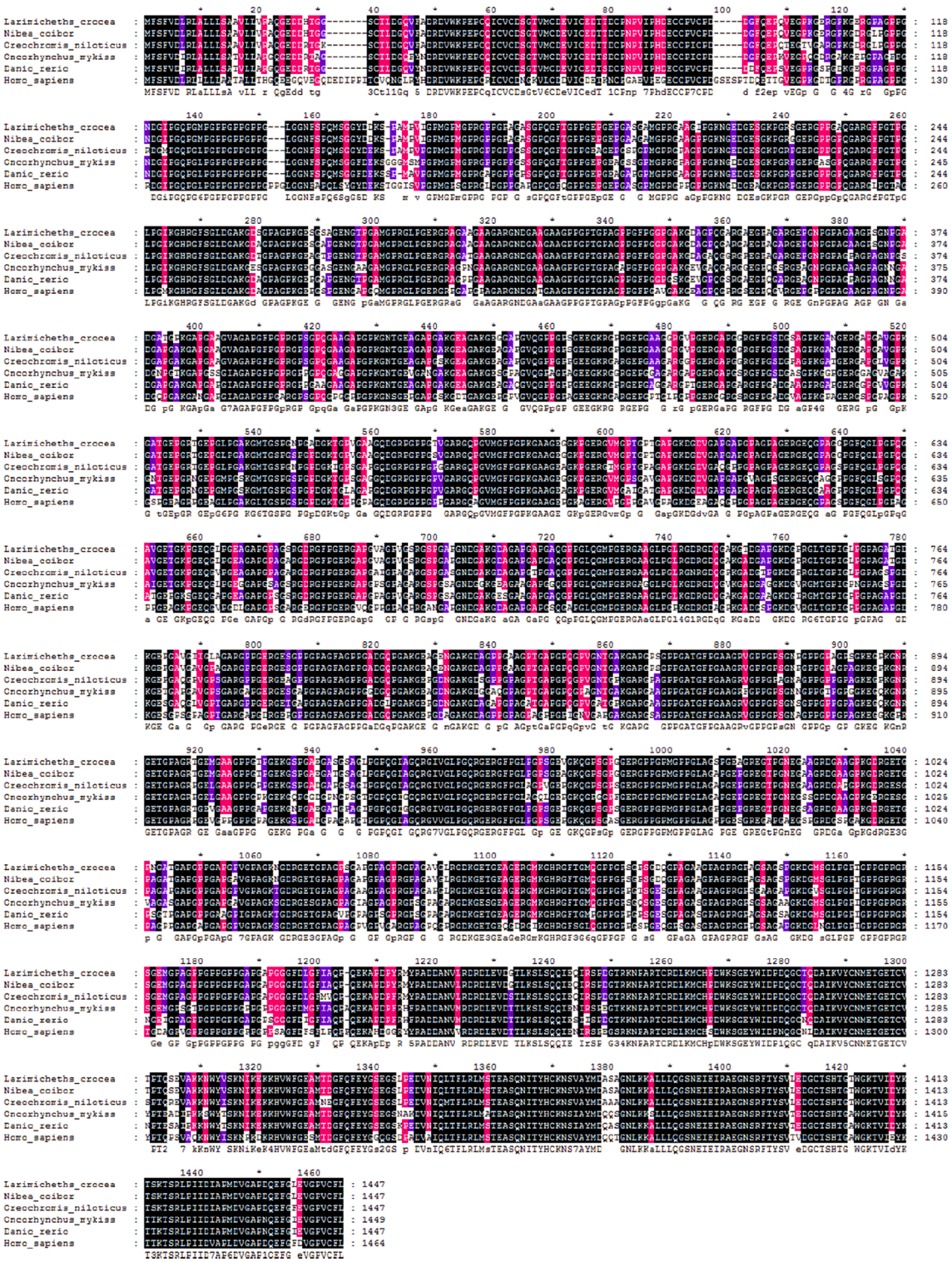


Fig. 2. Multiple alignments of *N. coibor* COL1A1 amino acid sequence with other organisms. The high, medium and low conserved amino acids residues are shown in black, red and blue shadows, respectively. The capital letters below the contrast sequence represent completely conserved amino acid residues, the small letters represent moderately conserved amino acid residues, and the black dotted lines represent intervals. The sequence of *N. coibor* (MK641513) COL1A1 amino acid was compared against those of *L. crocea* (XP_027142671.1), *Oncorhynchus mykiss* (NP_001117679.1), *Oreochromis niloticus* (NP_001269826.1), *Danio rerio* (NP_892012.2) and *Homo sapiens* (NP_892012.2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for the physical and chemical properties of collagen fibres and the stability of structure (Shoulders and Raines, 2009; Motooka et al., 2012). The Arg-Gly-Asp sequence is a potential cell binding site of the triple helix domain, which contains the C-terminal propeptide involved in chain selection and chain association during intracellular formation of procollagen (Leininger et al., 1992). The Arg-Gly-Asp sequence and C-terminal propeptide play a significant role in maintaining the protein in soluble form during transport to the extracellular matrix (Brodsky and Persikov, 2005) for protein metabolism.

There were eight disulfide residues (-cys-) and one tripeptide

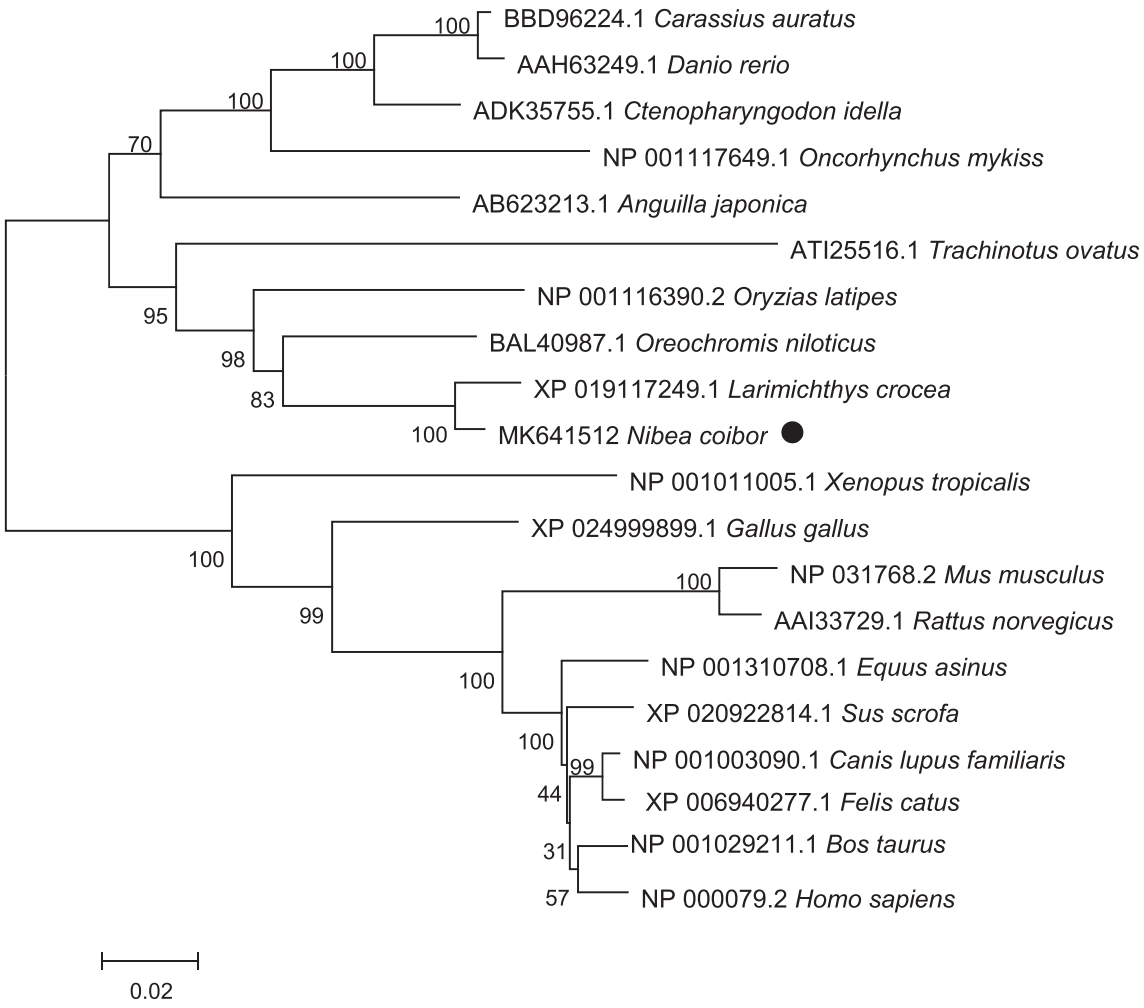


Fig. 3. Phylogenetic analysis of COL1A1 from *N. coibor* (●).

sequence Asn-Ile-Thr (1017–1020) found in the C-terminal propeptide regions of COL1A1 in *N. coibor*. The number of disulfide residues observed in *N. coibor* are similar to those found in rainbow trout (Saito et al., 2001), zebrafish (Dubois et al., 2002) and grass carp (Yu et al., 2014). The Cys residues in C-terminal propeptides are important because they participate in intrachain and interchain disulfide bridges. The interchain disulfide bridges play essential role in the interaction of the procollagen α chains (Barnes et al., 2019). The single tripeptide sequence Asn-Ile-Thr is a potential carbohydrate attachment site, which was also found in the C-terminal propeptide (1364–1366) of the COL1A1 protein sequence of skate (Hwang et al., 2006). These results indicate that the cloned sequences were those of the proα chain of COL1. The collagens in *N. coibor* shared basic biofunctions of fibrillar collagen, such as mRNA transcription, cell-adhesion and cell binding, common in many vertebrate species including mammals. In addition, the characteristic features of the triple helical domain and C-terminal domain are highly conserved in all species suggesting preserved protein functions.

The phylogenetic tree based on the deduced amino acid sequence indicated a clear separation of COL1A1 in all examined vertebrate species. The phylogenetic analysis illustrated that the newly cloned COL1A1 from *N. coibor* grouped together with teleost fish, among which it formed a monophyletic group with *L. crocea*, suggesting they have a close relationship. This not surprising because taxonomically, the two species belong to the same *Sciaenidae* family. Furthermore, their taxonomic relatedness was confirmed by the amino acid sequence similarity of *N. coibor* COL1A1 and that of *L. crocea* (98%) among all the animals compared in the alignment. We further performed a phylogenetic analysis on the Pro-α1 (I) chain of *N. coibor* and 21 different types of collagen precursor peptide chains published in zebrafish. We found that the *N. coibor* COL1A1 was most closely related to the COL1A1A of zebrafish, followed by COL1A1B. This confirmed that the cloned gene belongs to COL1A1 homologous gene. We also found that the molecular relationship between type I collagen α1 and type II collagen α1 was closer than that between type I collagen α1 and type I collagen α2. These results indicate that the naming of collagen types and subtypes is not based on molecular evolution, but may be related to the similarity in protein structure.

Our study indicated that the *N. coibor* COL1A1 gene was mainly expressed in the spine, swim bladder and skin, compared to scale, muscle, fin, liver and intestine. Similarly, the COL1A1 had the highest expression level in skin and scales of golden carp (Kondo and Watabe, 2004). Moreover, Yu et al. (2014) also showed that the expression of COL1A1 gene in skin and fins was higher than in grass carp muscles and intestine. Our results also showed wide expression of COL1A1 gene in the eight sampled tissues similar to results in *Raja kenojei* in which COL1A1 gene was expressed in most tissues such as muscle, skin, cartilage, gill, orbit, brain, liver, intestine, heart, stomach, spleen, pancreas, egg, ovary, fallopian tube and shell gland, except whole blood (Hwang et al., 2006). The expression of COL1A1 gene in different tissues is related to collagen biosynthesis, degree of pro-hydroxylation and cross-linkage of collagen molecules in various tissues (Turecek et al., 2008; Varga et al., 2010). The wide distribution of COL1A1 gene in various tissues suggest the important functions played by collagen in most of them.

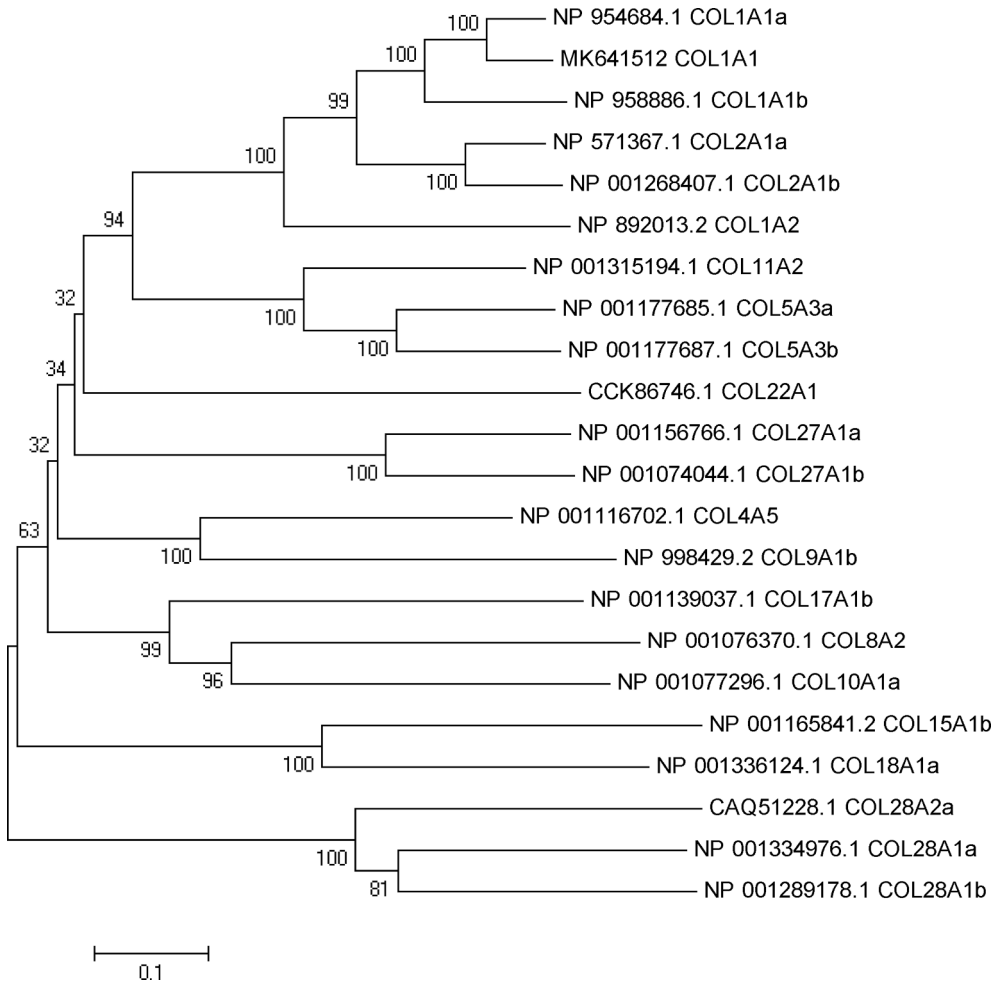


Fig. 4. Phylogenetic analysis of COL1A1 from *N. coibor* (●) and COLsAs from *Danio rerio.*

We also found that, the differential expression of COL1A1 gene in various tissues was related to the distribution of collagen content in the tissues. For example, the collagen content in the swim bladder was the highest, followed by the spine, fish scales and skin, while less collagen concentration was found in the liver, intestine, muscle and fin. The close relationship between collagen content tissue distribution and COL1A1 gene expression might be due to the directional deposition of collagen in the different tissues. Collagen is the main component of the substantive organ-basement membrane representing more than 80% of the organic matter in bones (Viguet-Carrin et al., 2006; Garnero, 2015). Collagen plays important roles in bones and connective tissue by forming and maintaining the integrity of the skeleton (Ferreira et al., 2012). Accordingly, the collagen content in connective tissues such as swim bladder and skin are also high. For example, collagen in swim bladder accounts for about 30% of the wet weight, and more than 70% of the total fish swim bladder protein (Wen et al., 2016; Lin et al., 2020). Moreover, the existence of different types of collagen in various tissues may be another reason for the close relationship between collagen content tissue distribution and COL1A1 gene expression. It has been reported that collagen existing in the swim bladder and digestive tract is solely type I; skin, scale, muscle, and fin contain much type I and less type II; vertebra and snout cartilage have extremely high amounts of type II collagen and moderate type I collagen; and notochord comprise almost solely type II (Hulmes, 2002). In addition, different regulatory mechanisms of collagen deposition may also exist between tissues, as aspect which requires further studies to investigate. The existence of high mRNA expression of COL1A1 gene and collagen content in swim bladder suggest that, this tissue is a target for collagen synthesis and production required for various uses in order to reduce the increasing demand of this valuable product.

5. Conclusion

The present study successfully cloned and showed the tissue distribution of the COL1A1 gene from *N. coibor* for the first time. The deducted protein contained all the features of collagen pro-α1(I) chains proteins, and shared high homology with other marine teleost. The *N. coibor* COL1A1 gene was broadly expressed in most tissues with the highest level in spine and swim bladder, followed by the scales, skin and muscle, and the lowest level in liver, intestine and fin, consistent with the collagen tissue distribution. These results contribute to better understanding of the COL1A1 gene synthesis in *N. coibor* tissues for valuable collagen production required in the market for various uses.

6. Author statement

Xiaobo Wen provided funding for research. Hua Rong and Fan Lin designed the study. Baojia Chen and Jia Zheng conducted the research. Lijun Ning collected and extracted the data. Kun Wu performed the statistical analysis. Hua Rong wrote the draft of the manuscript. Samwel Mchele Limbu and Xiaobo Wen critically wrote and revised the manuscript. All authors read and approved the final version of manuscript for submission.

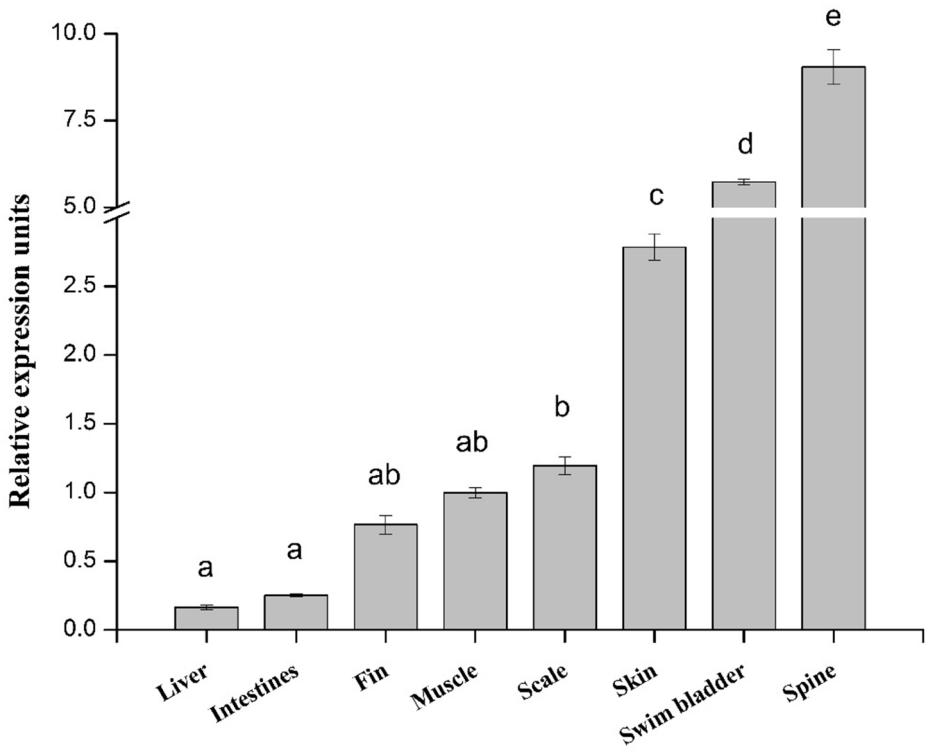


Fig. 5. The mRNA expression of COL1A1 gene in different tissues of *N. coibor* relative to the muscle. Bars with different letters indicate values with significant difference ($n = 3$) ($P < 0.05$).

Table 3
The total Hyp and collagen contents in different tissues of *N. coibor*.

	Hyp (g kg^{-1} wet basis)	Collagen (g kg^{-1} wet basis)
Liver	0.15 ± 0.01 ^{h}	1.18 ± 0.09^{a}
Intestine	0.60 ± 0.01 ^{g}	4.81 ± 0.04^{b}
Muscle	1.42 ± 0.02^{f}	11.35 ± 0.15^{c}
Fin	1.99 ± 0.04^{e}	15.94 ± 0.31^{d}
Skin	4.12 ± 0.11^{d}	32.97 ± 0.90^{e}
Scale	6.01 ± 0.23^{c}	48.11 ± 1.80^{f}
Spine	10.91 ± 0.22^{b}	87.26 ± 1.78 ^{g}
Swim bladder	24.48 ± 0.05^{a}	195.87 ± 0.39 ^{h}

Values represent mean ± SEM ($n = 3$). Values in the same column with different superscripts are significantly different ($P < 0.05$).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was financially supported by the Key Program of Marine Economy Development (Six Marine Industries) Special Foundation of Department of Natural Resources of Guangdong Province (GDNRC [2020]076) project, the Team Project of Department of Education of Guangdong Province (2018KCXTD012) and Guangdong Province Basic and Applied Basic Research Fund—Provincial and Municipal Joint Fund (2021A1515110179). We are grateful to all laboratory members for technical advice and valuable help during the feeding trial, sample collection and data analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2022.146441.

References

AOAC, 1995. Association of Official Analytical Chemists (AOAC). Official Methods of Analysis of Official Analytical Chemists International, 16th ed. Association of Official Analytical Chemists, Arlington, VA.

Avila Rodríguez, M.I., Rodríguez Barroso, L.G., Sánchez, M.L., 2018. Collagen: a review on its sources and potential cosmetic applications. J. Cosmet. Dermatol. 17, 20–26. https://doi.org/10.1111/jocd.12450.

Banerjee, S., Gill, S.S., Jain, P.K., Sirohi, A., 2017. Isolation, cloning, and characterization of a cuticle collagen gene, Mi-col-5, in *Meloidogyne incognita*. 3. Biotech 7, 64. https://doi.org/10.1007/s13205-017-0665-1.

Barnes, A.M., Ashok, A., Makareeva, E.N., Brusel, M., Cabral, W.A., Weis, M., Moali, C., Bettler, E., Eyre, D.R., Cassella, J.P., Leikin, S., Hulmes, D.J.S., Kessler, E., Marini, J. C., 2019. COL1A1 C-propeptide mutations cause ER mislocalization of procollagen and impair C-terminal procollagen processing. Biochim. Biophys. Acta Mol. Basis Dis. 1865, 2210–2223. https://doi.org/10.1016/j.bbadis.2019.04.018.

Bella, J., Hulmes, D.J.S., 2017. Fibrillar collagens. In: Parry, D.A.D., Squire, J.M. (Eds.), Fibrous Proteins: Structures and Mechanisms. Springer International Publishing, Cham, pp. 457–490.

Bhagwat, P.K., Dandge, P.B., 2018. Collagen and collagenolytic proteases: A review. Biocatal. Agric. Biotechnol. 15, 43–55. https://doi.org/10.1016/j.bcab.2018.05.005.

Bou-Gharios, G., Abraham, D., de Crombrugghe, B., 2020. Chapter 13 - Type I collagen structure, synthesis, and regulation. In: Bilezikian, J.P., Martin, T.J., Clemens, T.L., Rosen, C.J. (Eds.), Principles of Bone Biology (Fourth Edition). Academic Press, pp. 295–337.

Brodsky, B., Persikov, A.V., 2005. Molecular Structure of the Collagen Triple Helix, Advances in Protein Chemistry. Academic Press, pp. 301-339.

Chen, S., Chen, H., Xie, Q., Hong, B., Chen, J., Hua, F., Bai, K., He, J., Yi, R., Wu, H., 2016. Rapid isolation of high purity pepsin-soluble type I collagen from scales of red drum fish (*Sciaenops ocellatus*). Food Hydrocoll. 52, 468–477. https://doi.org/10.1016/j.foodhyd.2015.07.027.

D'Alessio, M., Bernard, M., Pretorious, P.J., de Wet, W., Ramirez, F., 1988. Complete nucleotide sequence of the region encompassing the first twenty-five exons of the human proα 1(I) collagen gene (COL1A1). Gene 67, 105–115. https://doi.org/10.1016/0378-1119(88)90013-3.

Dubois, G.M., Haftek, Z., Crozet, C., Garrone, R., Le Guellec, D., 2002. Structure and spatio temporal expression of the full length DNA complementary to RNA coding for

α2 type I collagen of zebrafish. Gene 294, 55–65. https://doi.org/10.1016/S0378-1119(02)00770-9.
Engel, J., Bächinger, H.P., 2005. Structure, stability and folding of the collagen triple helix. In: Brinckmann, J., Notbohm, H., Müller, P.K. (Eds.), Collagen: Primer in Structure, Processing and Assembly. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 7–33.
Exposito, J.-Y., Valcourt, U., Cluzel, C., Lethias, C., 2010. The fibrillar collagen family. Int. J. Mol. Sci. 11, 407–426. https://doi.org/10.3390/ijms11020407.
Ferreira, A.M., Gentile, P., Chiono, V., Ciardelli, G., 2012. Collagen for bone tissue regeneration. Acta Biomater. 8, 3191–3200. https://doi.org/10.1016/j.actbio.2012.06.014.
Garnero, P., 2015. The role of collagen organization on the properties of bone. Calcif. Tissue Int. 97, 229–240. https://doi.org/10.1007/s00223-015-9996-2.
Gelse, K., Pöschl, E., Aigner, T., 2003. Collagens - structure, function, and biosynthesis. Adv. Drug Del. Rev. 55, 1531–1546. https://doi.org/10.1016/j.addr.2003.08.002.
Gistelinck, C., Kwon, R.Y., Malfait, F., Symoens, S., Harris, M.P., Henke, K., Hawkins, M. B., Fisher, S., Sips, P., Guillemyn, B., Bek, J.W., Vermassen, P., De Saffel, H., Witten, P.E., Weis, M., De Paepe, A., Eyre, D.R., Willaert, A., Coucke, P.J., 2018. Zebrafish type I collagen mutants faithfully recapitulate human type I collagenopathies. Proceedings of the National Academy of Sciences 115, E8037. https://doi.org/10.1073/pnas.1722200115.
Gordon, M.K., Hahn, R.A., 2009. Collagens. Cell Tissue Res. 339 (1), 247–257.
Goto, T., Katada, T., Kinoshita, T., Kubota, H.Y., 2000. Expression and characterization of *Xenopus type I collagen alpha 1 (COL1A1)* during embryonic development. Dev. Growth Differ. 42, 249–256. https://doi.org/10.1046/j.1440-169x.2000.00505.x.
Huang, C.-Y., Kuo, J.-M., Wu, S.-J., Tsai, H.-T., 2016. Isolation and characterization of fish scale collagen from tilapia (*Oreochromis* sp.) by a novel extrusion–hydro-extraction process. Food Chem. 190, 997–1006. https://doi.org/10.1016/j.foodchem.2015.06.066.
Hulmes, D.J.S., 2002. Building collagen molecules, fibrils, and suprafibrillar structures. J. Struct. Biol. 137, 2–10. https://doi.org/10.1006/jsbi.2002.4450.
Hwang, J.-H., Yokoyama, Y., Mizuta, S., Yoshinaka, R., 2006. cDNA cloning and characterization of Type I procollagen α1 chain in the skate *Raja kenojei*. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 144, 1–10. https://doi.org/10.1016/j.cbpb.2005.12.026.
Kafienah, W., Sims, T.J., 2004. Biochemical methods for the analysis of tissue-engineered cartilage. In: Hollander, A.P., Hatton, P.V. (Eds.), Biopolymer Methods in Tissue Engineering. Humana Press, Totowa, NJ, pp. 217–229.
Karim, A.A., Bhat, R., 2009. Fish gelatin: properties, challenges, and prospects as an alternative to mammalian gelatins. Food Hydrocoll. 23, 563–576. https://doi.org/10.1016/j.foodhyd.2008.07.002.
Kimura, S., Zhu, X.-P., Matsui, R., Shijoh, M., Takamizawa, S., 1988. Characterization of fish muscle type I collagen. J. Food Sci. 53, 1315–1318. https://doi.org/10.1111/j.1365-2621.1988.tb09266.x.
Kite, M., 2006. Collagen products: glues, gelatine, gut membrane and sausage casings. In: Kite, M., Thomson, R. (Eds.), Conservation of Leather and related materials. Routledge, Oxford, UK, pp. 214–219.
Koivula, M.-K., Ruotsalainen, V., Björkman, M., Nurmenniemi, S., Ikäheimo, R., Savolainen, K., Sorva, A., Risteli, J., 2010. Difference between total and intact assays for N-terminal propeptide of type I procollagen reflects degradation of pN-collagen rather than denaturation of intact propeptide. Ann. Clin. Biochem. Int. J. Lab. Med. 47, 67–71. https://doi.org/10.1258/acb.2009.009110.
Kondo, H., Watabe, S., 2004. Temperature-dependent enhancement of cell proliferation and mRNA expression for type I collagen and HSP70 in primary cultured goldfish cells. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology 138, 221–228. https://doi.org/10.1016/j.cbpb.2004.04.001.
Leininger, E., Ewanowich, C.A., Bhargava, A., Peppler, M.S., Kenimer, J.G., Brennan, M. J., 1992. Comparative roles of the Arg-Gly-Asp sequence present in the *Bordetella pertussis* adhesins pertactin and filamentous hemagglutinin. Infect. Immun. 60, 2380–2385.
Lin, F., Rong, H., Lin, J., Yuan, Y., Yu, J., Yu, C., You, C., Wang, S., Sun, Z., Wen, X., 2020. Enhancement of collagen deposition in swim bladder of Chu's croaker (*Nibea coibor*) by proline: View from in-vitro and in-vivo study. Aquaculture 523, 735175. https://doi.org/10.1016/j.aquaculture.2020.735175.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods 25, 402–408. https://doi.org/10.1006/meth.2001.1262.
Mahboob, S., 2015. Isolation and characterization of collagen from fish waste material-skin, scales and fins of *Catla catla* and *Cirrhinus mrigala*. J. Food Sci. Technol. 52, 4296–4305. https://doi.org/10.1007/s13197-014-1520-6.
Motooka, D., Kawahara, K., Nakamura, S., Doi, M., Nishi, Y., Nishiuchi, Y., Kang, Y.K., Nakazawa, T., Uchiyama, S., Yoshida, T., Ohkubo, T., Kobayashi, Y., 2012. The triple helical structure and stability of collagen model peptide with 4(s)-hydroxyprolyl-pro-gly units. Biopolymers 98, 111–121. https://doi.org/10.1002/bip.21730.
Pal, G.K., Suresh, P.V., 2016. Sustainable valorisation of seafood by-products: Recovery of collagen and development of collagen-based novel functional food ingredients. Innov. Food Sci. Emerg. Technol. 37, 201–215. https://doi.org/10.1016/j.ifset.2016.03.015.
Pozzolini, M., Scarfì, S., Giovine, M., 2020. Marine Collagen and its Biotechnological Applications. In: Kim, S.-K. (Ed.), Encyclopedia of Marine Biotechnology. Wiley, pp. 1007–1030.
Retief, E., Parker, M.I., Retief, A.E., 1985. Regional chromosome mapping of human collagen genes alpha 2(I) and alpha 1(I) (COLIA2 and COLIA1). Hum. Genet. 69, 304–308. https://doi.org/10.1007/BF00291646.
Ricard-Blum, S., 2011. The collagen family. Cold Spring Harb. Perspect. Biol. 3 (1), a004978.
Rong, H., Zhang, Y., Hao, M., Zou, W., Yu, J., Yu, C., Shi, Q., Wen, X., 2019. Effects of dietary hydroxyproline on collagen metabolism, proline 4-hydroxylase activity, and expression of related gene in swim bladder of juvenile *Nibea diacanthus*. Fish Physiol. Biochem. 45, 1779–1790. https://doi.org/10.1007/s10695-019-00676-9.
Rong, H., Lin, F., Limbu, S.M., Lin, Z., Bi, B., Dou, T., Zhao, L., Wen, X., 2020a. Effects of dietary proline on swim bladder collagen synthesis and its possible regulation by the TGF-β/Smad pathway in spotted drum, *Nibea diacanthus*. Aquacult. Nutr. 26, 1792–1805. https://doi.org/10.1111/anu.13130.
Rong, H., Zhang, Y., Hao, M., Lin, F., Zou, W., Zhang, H., Yu, C., Yu, J., Shi, Q., Aweya, J. J., Wen, X., 2020b. Effect of hydroxyproline supplementation on growth performance, body composition, amino acid profiles, blood-biochemistry and collagen synthesis of juvenile chu's croaker (*Nibea coibor*). Aquacult. Res. 51, 1264–1275. https://doi.org/10.1111/are.14477.
Saito, M., Takenouchi, Y., Kunisaki, N., Kimura, S., 2001. Complete primary structure of rainbow trout type I collagen consisting of α1(I)α2(I)α3(I) heterotrimers. Eur. J. Biochem. 268, 2817–2827. https://doi.org/10.1046/j.1432-1327.2001.02160.x.
Shoulders, M.D., Raines, R.T., 2009. Collagen structure and stability. Annu. Rev. Biochem. 78, 929–958. https://doi.org/10.1146/annurev.biochem.77.032207.120833.
Slatter, D.A., Farndale, R.W., 2015. Structural constraints on the evolution of the collagen fibril: convergence on a 1014-residue COL domain. Open Biol. 5 (5), 140220.
Solomon, E., Hiorns, L., Sheer, D., Rowe, D., 1984. Confirmation that the Type I collagen gene on chromosome 17 is COL1A1 (α1(I)), using a human genomic probe. Ann. Hum. Genet. 48, 39–42. https://doi.org/10.1111/j.1469-1809.1984.tb00831.x.
Stover, D.A., Verrelli, B.C., 2010. Comparative vertebrate evolutionary analyses of type I collagen: Potential of COL1a1 gene structure and intron variation for common bone-related diseases. Mol. Biol. Evol. 28, 533–542. https://doi.org/10.1093/molbev/msq221.
Sun, L., Hou, H., Li, B., Zhang, Y., 2017. Characterization of acid- and pepsin-soluble collagen extracted from the skin of Nile tilapia (*Oreochromis niloticus*). Int. J. Biol. Macromol. 99, 8–14. https://doi.org/10.1016/j.ijbiomac.2017.02.057.
Turecek, C., Fratzl-Zelman, N., Rumpler, M., Buchinger, B., Spitzer, S., Zoehrer, R., Durchschlag, E., Klaushofer, K., Paschalis, E.P., Varga, F., 2008. Collagen cross-linking influences osteoblastic differentiation. Calcif. Tissue Int. 82, 392–400. https://doi.org/10.1007/s00223-008-9136-3.
Tuuli, C.D., de Mitcheson, Y.S., Wai-Chuen, N.G., 2016. Molecular identification of croaker dried swim bladders (maw) on sale in Hong Kong using 16S rRNA nucleotide sequences and implications for conservation. Fish. Res. 174, 260–269. https://doi.org/10.1016/j.fishres.2015.10.018.
Varga, F., Rumpler, M., Zoehrer, R., Turecek, C., Spitzer, S., Thaler, R., Paschalis, E.P., Klaushofer, K., 2010. T3 affects expression of collagen I and collagen cross-linking in bone cell cultures. Biochem. Biophys. Res. Commun. 402, 180–185. https://doi.org/10.1016/j.bbrc.2010.08.022.
Viguet-Carrin, S., Garnero, P., Delmas, P.D., 2006. The role of collagen in bone strength. Osteoporosis Int. 17, 319–336. https://doi.org/10.1007/s00198-005-2035-9.
Wei, Z., Ma, J., Pan, X., Mu, H., Li, J., Shentu, J., Zhang, W., Mai, K., 2016. Dietary hydroxyproline improves the growth and muscle quality of large yellow croaker *Larimichthys crocea*. Aquaculture 464, 497–504. https://doi.org/10.1016/j.aquaculture.2016.07.015.
Wen, J., Zeng, L., Chen, Z., Xu, Y., 2016. Comparison of nutritional quality in fish maw product of croaker *Protonibea diacanthus* and perch *Lates niloticus*. J. Ocean Univ. China 15, 726–730. https://doi.org/10.1007/s11802-016-2870-7.
Yu, E.M., Liu, B.H., Wang, G.J., Yu, D.G., Xie, J., Xia, Y., Gong, W.B., Wang, H.H., Li, Z.F., Wei, N., 2014. Molecular cloning of type I collagen cDNA and nutritional regulation of type I collagen mRNA expression in grass carp. J. Anim. Physiol. Anim. Nutr. 98, 755–765. https://doi.org/10.1111/jpn.12132.