PLOS ONE

RESEARCH ARTICLE

Comparative transcriptome analysis of albino northern snakehead (*Channa argus*) reveals its various collagen-related DEGs in caudal fin cells

Shixi Chen[1,2©], Ning Li[3©], Fardous Mohammad Safiul Azam[1,4]*, Li Ao[1], Na Li[1], Jianlan Wang[1], Yuanchao Zou[1,2]*, Rui Li[1], Zakaria Hossain Prodhan[1]

1 College of Life Sciences, Neijiang Normal University, Neijiang, China, 2 Conservation and Utilization of Fishes resources in the Upper Reaches of the Yangtze River, Key Laboratory of Sichuan Province, Neijiang, China, 3 Sichuan Yukun Aquatic Technology Co., Tongchuan District, Dazhou City, Sichuan Province, China, 4 Department of Biotechnology and Genetic Engineering, Faculty of Life Sciences, University of Development Alternative, Dhaka, Bangladesh

© These authors contributed equally to this work.
* shojibbiotech@yahoo.com (FMSA); zou3891@163.com (YZ)

Data Availability Statement: The data supporting these findings is available in online repositories. The names of the repository/repositories and

Abstract

The albino northern snakehead (*Channa argus*) is an aquaculture species characterized by heritable albino body color, in contrast to the typical coloration. Additionally, there are gray- and golden-finned individuals, which exhibit distinct coloration in their caudal fins. We performed RNA-seq to profile the transcriptome of caudal fin tissues in albino gray-finned and golden-finned *C. argus*, contrasting these with normal morphs to elucidate the differences between the two groups. A total of 137,130 unigenes were identified in this study. Gene Ontology (GO) analysis showed that the identified DEGs were significantly enriched in cellular components related to cytoplasm. So far, 379 common DEGs have been identified in all three groups. Notably, we observed more DEGs in golden-finned individuals compared to gray-finned individuals. We also revealed that golden-finned individuals were enriched in collagen-related pathways compared with normal individuals. The enriched DEGs of collagen components include collagen I of *COL1A1* and *COL1A2*, collagen II of *COL2A1*, collagen V of *COL5A1* and *COL5A2*, collagen VI of *COL6A1* and *COL6A3*, collagen IX of *COL9A3*, collagen X of *COL10A1*, collagen XI of *COL11A2*, collagen XII of *COL12A1*, collagen XVI of *COL16A1*, collagen XVIII of *COL18A1* and decorin (DCN), all of which play a role in modulating the collagen matrix. In golden-finned albino fish, collagen-related genes were downregulated, suggesting that despite the abundance of collagen types in their caudal fin cells, gene expression was slightly limited. This work provides valuable genetic insights into collagen variation in albino *C. argus*, lays the foundation for research on collagen genes and is crucial for the development and utilization of fish-derived collagen as a biomaterial for tissue engineering and biomedical applications.

1 Introduction

Collagens, the most abundant proteins in the animal kingdom, are best recognized for providing mechanical strength to skin, bones, and other tissues [1]. It plays a key role in retaining the texture of fish muscles [2, 3], and increased collagen content in muscle enhances the firmness of fish fillets [4, 5], which is a vital feature in determining the acceptability of aquatic products. Owing to high water absorption capacity, high biocompatibility, low immunogenicity, biodegradability, high porosity, ease of processing, ability to penetrate a lipid-free interface, limited disease transmission, natural ability to combine with other materials (synthetic polymers), and options for industrial-scale extraction, collagen has limited or no ethical or religious constraints. Because of these properties, collagen has become a key multifunctional protein used in the food, cosmetics, pharmaceutical and biomedical industries [6]. Compared with porcine and bovine collagen is more readily absorbed and has higher bioavailability [7]. Altogether, they have become valuable biomaterials and have received significant attention from biomedical researchers [8]. In its therapeutic aspects, collagen derivatives are used for tissue engineering, cell culture, and drug delivery and have shown an ability to inhibit bacterial growth and promote wound healing [9, 10]. For instance, wound contraction rate and histological examination demonstrated that the prepared collagen effectively promoted wound healing, resulting in improved wound resolution and closure [9, 11]. The evidence reviewed above suggests that fish collagen is a promising option for wound healing and other therapeutic applications. In general, three collagen chains are assembled into a left-handed triple helix to form a protein complex, which is then built into higher-order structures [12, 13]. To date, 28 distinct forms of collagen (collagen I–XXVIII) have been identified [13–15]. In fish, type I collagen is the most frequent type; however, certain species contain type II collagen [10].

The northern snakehead (*Channa argus*) is a Channa species from the Channidae family of Perciformes, found in Asia and Africa [16, 17]. Moreover, in China, the annual production of this species is estimated to be approximately 510,000 tons, with a value of 1.6 billion US dollars [18]. However, it is classified as an invasive species in North America that threatens many other species [19]. This species is the most important snakehead species, with only a few bone spurs cultured in China and Malaysia. It has an edible rate of up to 63% and has gained popularity among people in these countries due to its high nutritional content [20–22]. It is also sold in ethnic live-food fish markets in New York and St. Louis, Missouri [17]. The northern snakehead albino phenotype is a unique heritable albino that is found exclusively in China. It was first recorded in Neijiang City, Sichuan Province, China [23, 24]. The natives of this city believe that albino *C. argus* with golden caudal fins has medicinal value, such as healing wounds and promoting lactation [25]. Considering its unique features and significant uses, several molecular studies have been conducted, including studies on the draft whole genome [18], genome characterization [26], molecular diversity [27], disease and pathogenesis [28], skin color and albinism [29], sex chromosomes [30], feeding behavior [24, 31] and surimi products [32].

Transcriptome sequencing analysis is widely utilized to identify differentially expressed genes (DEGs) and further investigate the molecular mechanisms involved in metabolic pathways [33–35]. Multiple studies have identified DEGs associated with collagen metabolism, such as palatal organ development in bighead carp (*Hypophthalmichthys nobilis*) [36], flesh pigmentation in Atlantic salmon (*Salmo salar*) [37], collagen synthesis genes for connective tissue growth factor (*Nibea coibor*) [38], and collagen deposition in zebrafish [3]. DEG analysis, for example, has shown the up- and downregulation of skin color genes, as well as the identification of five causative genes for albinism in this species via whole transcriptome sequencing [39]. To further elucidate the mechanism of albino features in albino-type (AT)

and bicolor-type (BT) fish, histological and transcriptome analyses were conducted to identify pigment-related genes in *C. argus* [40]. Their study uncovered a comprehensive process in the regulation of genes associated with color and albinism in the skin and spine tissues of *C. argus*. Nevertheless, the distinctive golden-colored caudal fin trait has hitherto not been explored in transcriptome studies. In addition, the hunt for novel collagen sources from fish, particularly fins, is a growing trend in the development of biomaterials for medical and tissue engineering applications [41, 42]. Thus, studying collagen metabolism in *C. argus* caudal fins might lead to the discovery of a novel biomaterial source since albino *C. argus* with golden caudal fins possesses wound healing capabilities that set it apart from the gray-finned albino. To the best of our knowledge, no research has been performed on collagen metabolism in northern snakeheads. This is the first report on albino *C. argus*, which revealed the presence of collagen-rich caudal fin cells.

2 Materials and methods

2.1 Experimental fish

The fish used in this study were healthy, one-year-old *C. argus* collected from the Guiming Aquaculture Family Farm in Chaoyang Town, Neijiang City, Sichuan Province, China. Caudal fin tissues were obtained from nine *C. argus* (three repeats per group), comprising the W group of three black-white *C. argus*, H group of three gray-finned albino *C. argus*, and J group of three golden-finned albino *C. argus* (Fig 1). The tissues from the three replicates were collected in liquid nitrogen to generate three biological replicates, which were then preserved in a -80˚C freezer until needed. The study was approved by the Ethics Committee of Neijiang Normal University (protocol code SK1729 and approved on 2021-10-05). All methods were performed following relevant guidelines and regulations. The study was conducted under the ARRIVE guidelines [43].

2.2 RNA extraction, library construction and sequencing

Caudal fin tissue from each sample was crushed into a powder in liquid nitrogen, and total RNA was extracted using TRIzol reagent (Invitrogen, USA) according to the manufacturer's instructions. RNA quality and concentration were measured via 1% agarose gel electrophoresis and a Nanodrop (Thermo Fisher Scientific, Waltham, MA). A Fragment Analyzer 5400
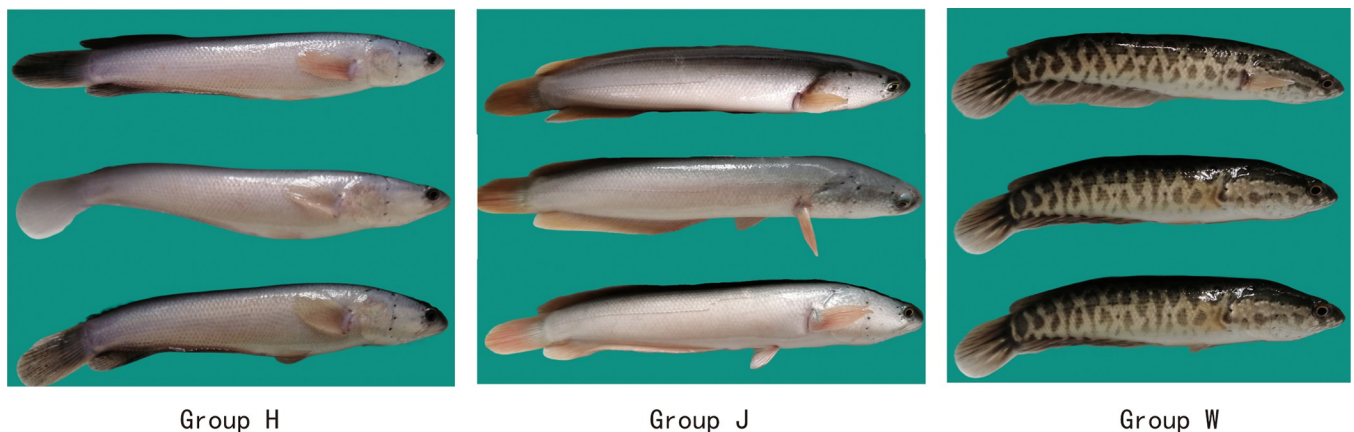
Fig 1. *Chana argus* was used in this study. Three distinct color morphs of *C. argus* were identified: Group H included three gray-finned albinos, Group J included three golden-finned albinos, and Group W included three black-whites.

(Agilent Technologies, CA, USA) was used to analyze RNA integrity. Index-coded samples were clustered using a cBot Cluster Generation System with a TruSeq PE Cluster Kit v3-cBot-HS (Illumina Inc., San Diego, CA, USA) according to the manufacturer's instructions. Next, we sequenced the library preparations using an Illumina NovaSeq 6000 platform and generated 150 bp paired-end reads.

2.3 Bioinformatics analysis of the nonparticipating transcriptome

The original fluorescence image files obtained from the Illumina platform were transformed into short reads (raw data) by base calling, and these short reads were recorded in FASTQ format, which includes both sequence and sequencing quality data [44]. Raw reads were trimmed to obtain high-quality, clean reads. Trimmomatic software was used to filter the sequencing data [45] and clean reads were assembled via Trinity [46]. The assembly quality of Trinity. fasta, unigene.fa and cluster.fasta was evaluated using the BUSCO software (S1 Fig) [47]. The unigene sequence acquired by Trinity assembly was used as the reference sequence, and the clean reads from each sample were compared to Ref., filtering out the reads with a comparison quality of less than 10, non-comparable reads, and reads compared to multiple regions of the genome. We extracted and merged the read count data to get the read count values for each sample, and the transcripts per kilobase million (TPM) values were computed using RNA-Seq via expectation maximization (RSEM) software with the default threshold. We collected and combined TPM data from each sample to determine gene expression levels [48]. Davidson and Oshlack [49] used the Corset to perform hierarchical transcript clustering. DEGs were defined as having an adjusted P-value <0.05, and a |log2(fold change)| > 1. To detect DEGs in the samples, we performed DESeq2 at different time points compared with the control group [50]. We utilized Cluster Profiler to analyze DEG enrichment in Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [51, 52].

2.4 Phylogenetic tree construction

Each collagen-related transcript identified in *C. argus* was subsequently analyzed using BLASTn to search for homologous genes in the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome; accessed date: April 5, 2024). For the phylogenetic analysis, we selected the top 15 homologous genes from the search results for each collagen gene (S1 Table). Multiple sequence alignments were carried out, and phylogenetic analysis was performed using MEGA11 software to construct a neighbor-joining phylogenetic tree, which was validated by a 1000-iteration bootstrap test [53]. Later, the linked fish species and families of each homologous gene were examined to ascertain the prevalence of each collagen-related gene within these families. The data were then tallied and shown as a stack bar graph.

3 Results

3.1 Transcriptome assembly

The current study categorized the *C. argus* specimens into three groups: H, J, and W. Groups H and J possessed hereditable white body coloring of *C. argus*, whereas group W had both black and white coloration. The disparity in body coloration among the three groups was significant at the caudal fin, with group J differing from the other two by its golden hue. Despite the sequencing of three transcriptome sets, the ensuing research primarily concentrated on the variation between groups W and J due to their pronounced body color features and transcriptome variations.

Table 1. Sequencing data list of *Channa argus*.

mple id	raw base (Gb)	raw sequences	clean reads (Gb)	clean bases	q30 (%)	GC (%)	Assembled sequences	unigene
H1	6.26	41704178	5.83	39432270	99.9	45.5(%)		
H2	6.39	42584530	5.97	40468492	99.9	46.0(%)		
H3	7.46	49740560	6.97	47218788	99.9	46.0(%)		
J1	6.27	41829506	5.87	39783606	99.91	46.0(%)		
J2	6.53	43564594	6.1	41357768	99.9	46.0(%)		
J3	6.37	42467434	5.93	40201484	99.9	46.0(%)		
W1	5.97	39784420	5.56	37660476	99.9	45.0(%)		
W2	6.2	41312750	5.8	39229648	99.9	45.5(%)		
W3	6.48	43183456	6.09	41173154	99.91	45.0(%)		
Total	**57.93**		**54.12**				**174779**	**137130**

https://doi.org/10.1371/journal.pone.0315996.t001

For each fish examined, transcriptome sequencing of *C. argus* produced a raw base of 5.97 to 7.46 Gb; the quantity of clean data was 54.12 Gb, including 174,779 assembled sequences, 137,130 unigenes, 86827 predicted coding sequences (CDSs) and 2,532 clusterSeqs. The GC content of each transcriptome ranged from 45 to 46%. A summary of the caudal fin transcriptome data for *C. argus* is presented in Table 1.

Five functional databases were used to classify these unigenes, and 86,828 NR (63.32%), 105,747 NT (77.11%), 71,814 SwissProt (52.37%), 70,693 KOG (51.55%), and 71,751 UniProt (52.32%) genes were identified; 68,700 of these unigenes were commonly annotated, and 56,615 KEGG (41.29%) and 55,175 GO (40.24%) were annotated (Fig 2A). The stronger the correlation coefficient between the samples, the more similar their expression patterns. The samples for the H, J and W groups were 0.9 ~ 1.0 and were closely correlated (Fig 2B). The sequencing data generated in this study were deposited in the Sequence Read Archive of the National Center for Biotechnology Information (NCBI) (accession number: PRJNA913664).

3.2 Functional annotation of DEG genes and their variations

Gene Ontology (GO) annotation and enrichment analysis of the DEGs in *C. argus* demonstrated that the DEGs were associated with the terms biological process (BP), cellular component (CC), and molecular function (MF), with the most enriched term being the processes related to the cytoplasm of CC (Fig 3).

We identified and clustered them to investigate variations in DEG expression modalities. Expression patterns of the same gene clusters were comparable. DEGs were significantly separated into groups J, H and W as shown in the heatmap. Group J was in a separate cluster compared to groups H and W, indicating a greater disparity between the two groups (Fig 4).

The overlap of DEGs across different comparison combinations is depicted in a Venn diagram (Fig 5A). This study compared the DEGs of groups J or H with those of group W to identify common DEGs, revealing 2098 DEGs in group J, 1603 DEGs in group H, and a total of 379 common DEGs. The GO enrichment analysis indicated that the quantities of upregulated genes, downregulated genes, and total genes within each of the three GO categories (cellular component, biological process, and molecular function) were counted individually and are displayed as bar graphs (Fig 5B and 5C). The results revealed that both groups H and J exhibited more downregulated DEGs compared to group W.

The heatmap and Venn diagram revealed that the DEGs in group J compared to group W exceeded those in group J compared to group H. Consequently, the following analysis was conducted between groups J and W.
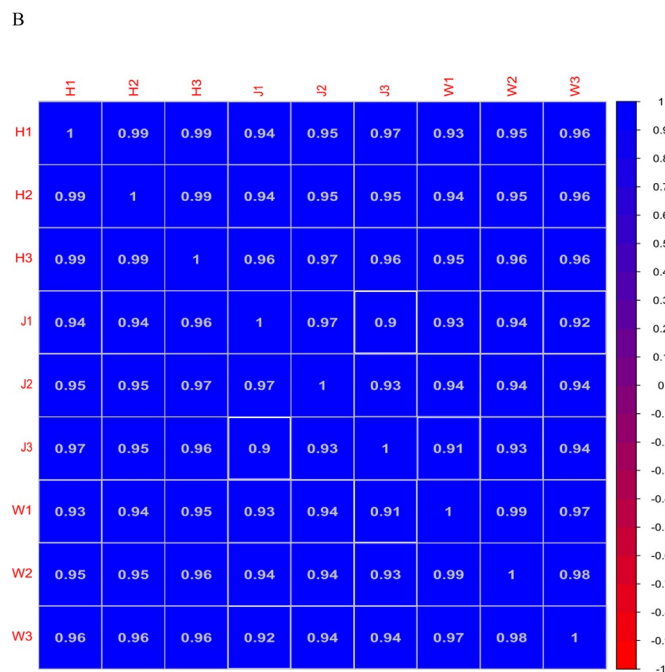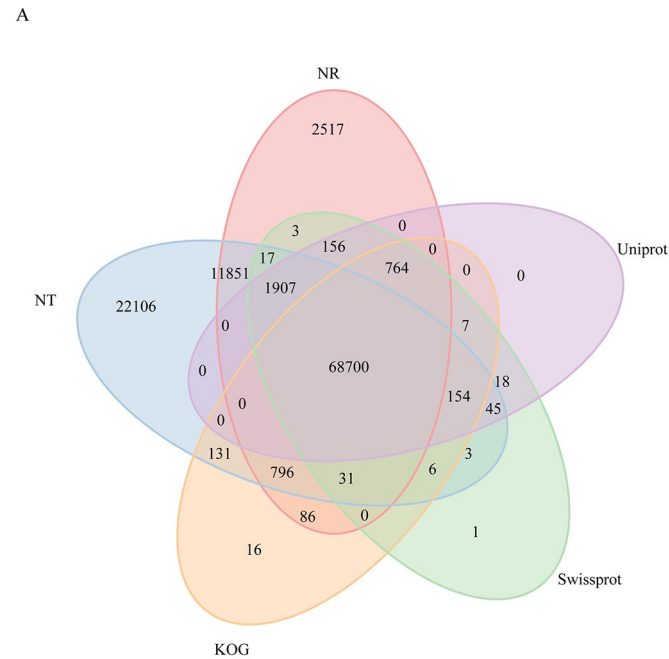
Fig 2. A. Venn diagram based on 5 databases: NR, NT, SwissProt, KOG and UniProt; numbers in the overlapping regions refer to those unigenes that were expressed in more than one database, and 68700 shared unigenes matched with all five databases; B. Correlations of the samples between the H (H1, H2, H3), J (J1, J2, J3) and W (W1, W2, W3) groups. The change in the correlation coefficient is indicated by the change in color from blue (1.0) to red (-1.0), and a deeper color indicates a higher correlation between samples.

https://doi.org/10.1371/journal.pone.0315996.g002

3.3 Distinct collagen-related pathways in white *C. argus* with golden fins

Gene Ontology (GO) enrichment analysis confirmed that DEGs from group J, compared with W, were significantly enriched in collagen-related pathways (p-value < 0.05). These pathways
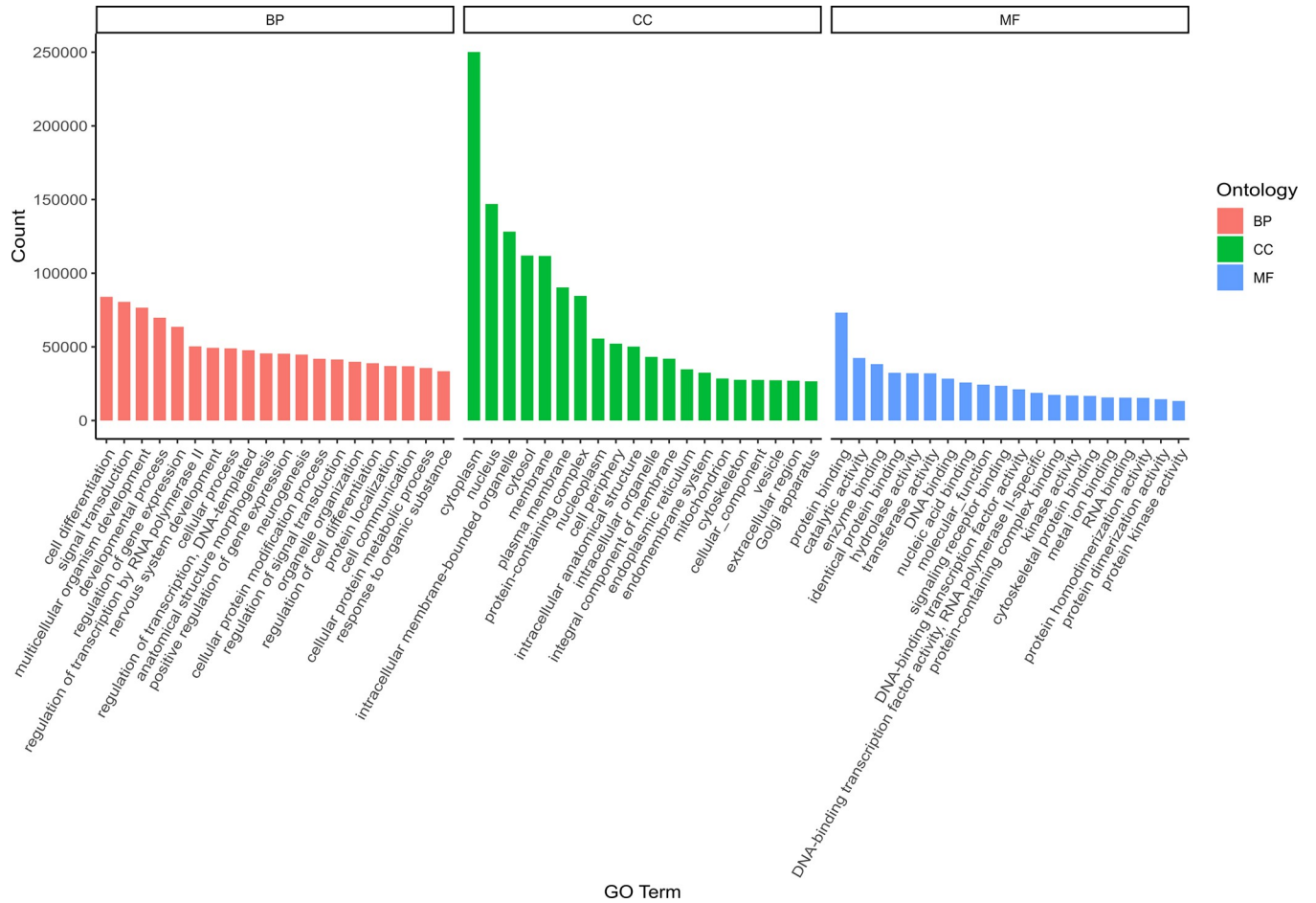
Fig 3. Genes were classified according to three major GO categories: Biological process (BP), cellular component (CC), and molecular function (MF).
The y-axis indicates the number of genes in each category.

https://doi.org/10.1371/journal.pone.0315996.g003

encompassed collagen trimer, fibrillar collagen, banded collagen fibril, a complex of collagen trimers, collagen-activated tyrosine kinase receptor signaling pathway, collagen-activated signaling pathway and collagen type I trimer (Fig 6). Additional pathways included cellular responses to vitamins, synthesis of molecular mediators involved in the inflammatory response, protein heterodimerization, binding platelet-derived growth factor, tooth mineralization, positive regulation of long-term synaptic depression, cellular responses to vitamin E, CDP-alcohol phosphatidyl transferase activity, CDP-diacylglycerol-serine O-phosphatidyl transferase activity, tooth eruption, cardiac neuron differentiation, cardiac neuron development and negative regulation of monocyte chemotactic protein-1 production.

The directed acyclic graph illustrates the outcomes of the differential gene GO enrichment analysis, with branches representing inclusion relationships and progressively specific functional ranges defined from top to bottom (Fig 7). Colored shades denote the degree of enrichment, with darker colors indicating higher enrichment. The J-W CC top GO-directed acyclic network showed distinctive enrichment of the cellular component (1145/52530). The collagen components included a collagen-containing extracellular matrix (60/711), a complex of collagen trimers (16/69), collagen trimers (28/169), fibrillar collagen trimers (16/39), banded collagen fibrils (16/39), collagen type I trimers (7/10), and collagen type V trimers (5/17).
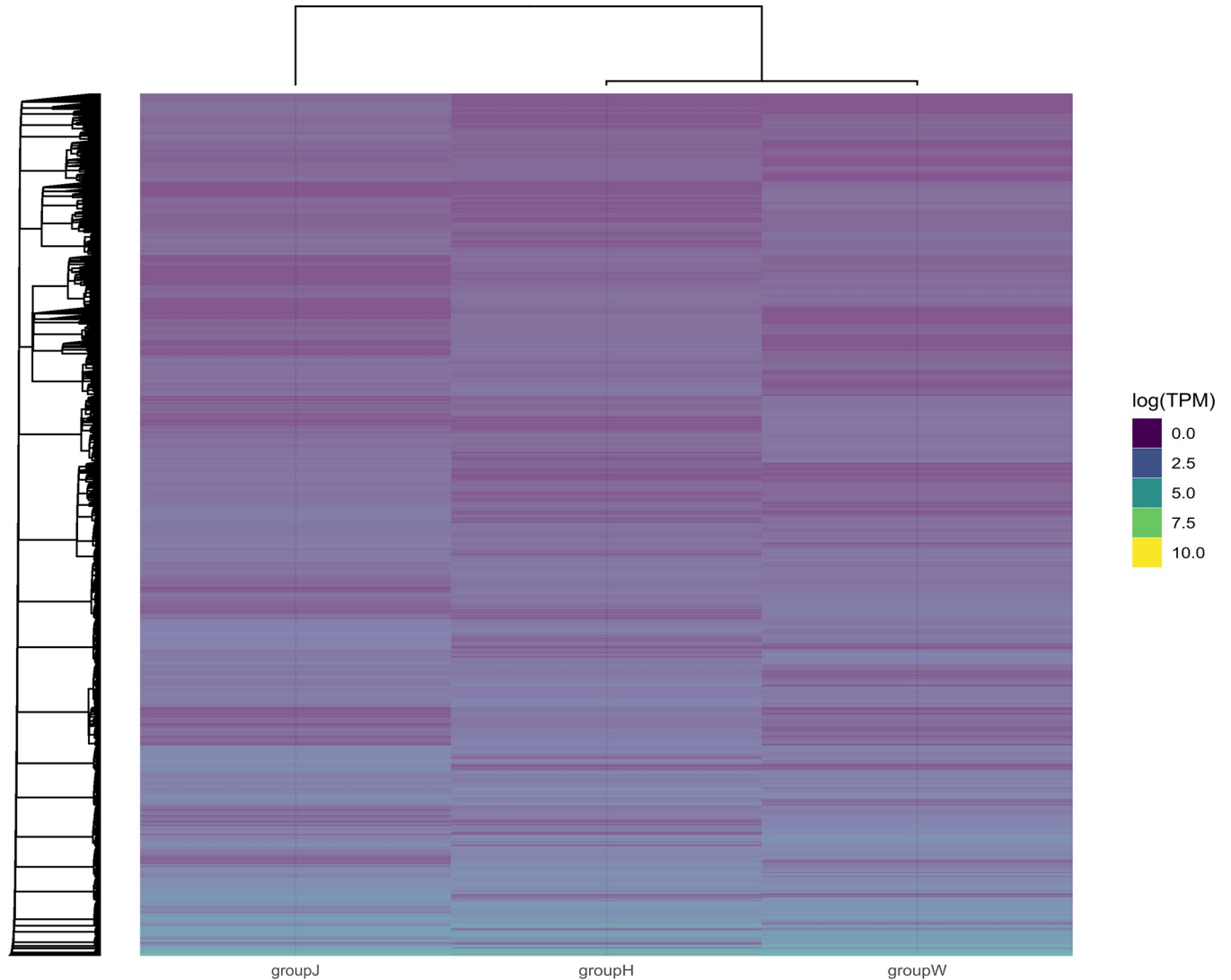
Fig 4. Heatmap of differentially expressed gene clustering. The colors can only be compared horizontally with the same clustering of genes for expression values (TPM) in the high (blue) and low (yellow) colors. Axes: Bottom, variable groups; top, hierarchical clusters for closeness between groups; left, hierarchical clustering among the expressed genes.

https://doi.org/10.1371/journal.pone.0315996.g004

The DEGs associated with collagen component terms included collagen I of *COL1A1* and *COL1A2*, collagen II of *COL2A1*, collagen V of *COL5A1* and *COL5A2*, collagen VI of *COL6A1* and *COL6A3*, collagen IX of *COL9A3*, collagen X of *COL10A1*, collagen XI of *COL11A2*, collagen XII of *COL12A1*, collagen XVI of *COL16A1*, collagen XVIII of *COL18A1* and decorin (DCN) for modulation of the collagen matrix. The DEGs of the J group exhibited a greater abundance of collagen component terms than those in the W group.

The analysis of log2-fold changes in expression levels among the collagen-related genes included six transcripts for *COL1A1*, one for *COL1A2*, two for *COL2A1*, one for *COL5A1*, four for *COL5A2*, one for *COL6A1*, four for *COL6A3*, one for *COL9A3*, one for *COL10A1*, one for *COL11A12*, two for *COL12A1*, one for *COL16A1* and one for *COL18A1* (Fig 8). In comparing the log2-fold changes in expression levels between the W and J groups, the J group was found to have significantly downregulated for collagen-related genes. *COL2A1*, *COL10A1*, *COL11A2*, and *COL18A1* exhibited a downward trend exceeding threefold in the J group.
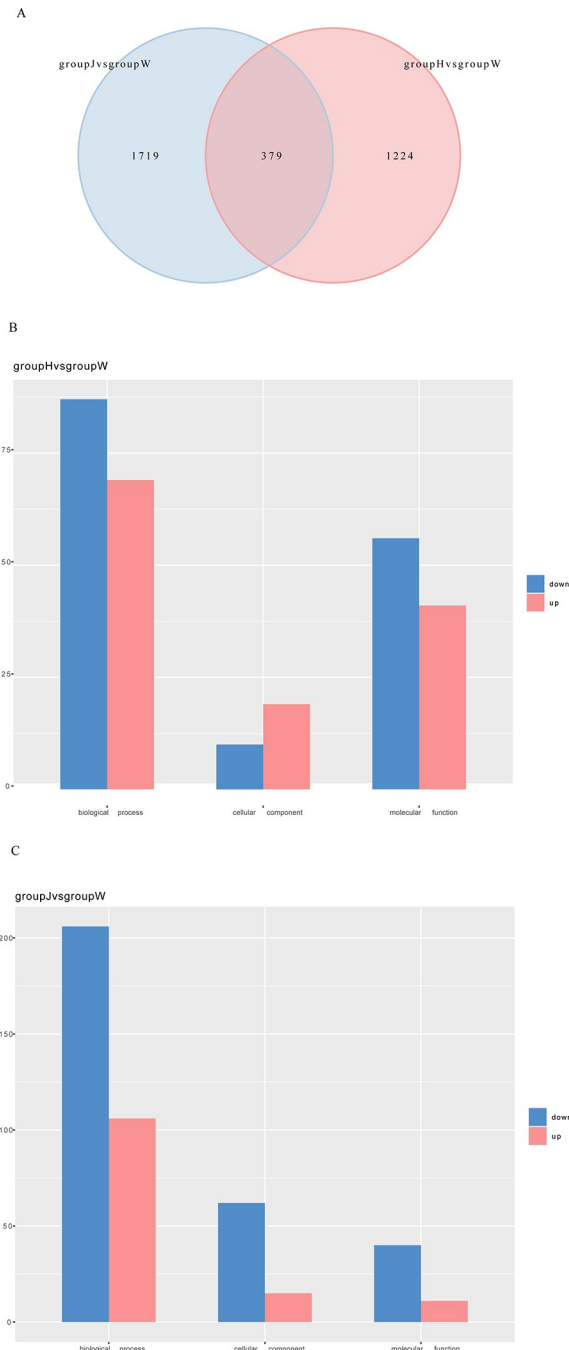
Fig 5. Differential genetic Venn diagram of groups J and H compared with group W and GO enrichment histogram of differentially expressed genes. A. Venn diagram between group J vs. group W and group H vs. group W. The number of genes expressed uniquely within each group is shown and overlapping regions indicate the number of genes expressed in both groups. B. Comparison of GO enrichment analysis for up-and down-regulated genes between groups H and W. C. Comparison of GO enrichment analysis for up- and downregulated genes between groups J and W. X-axis: number of genes; Y-axis: GO categories.
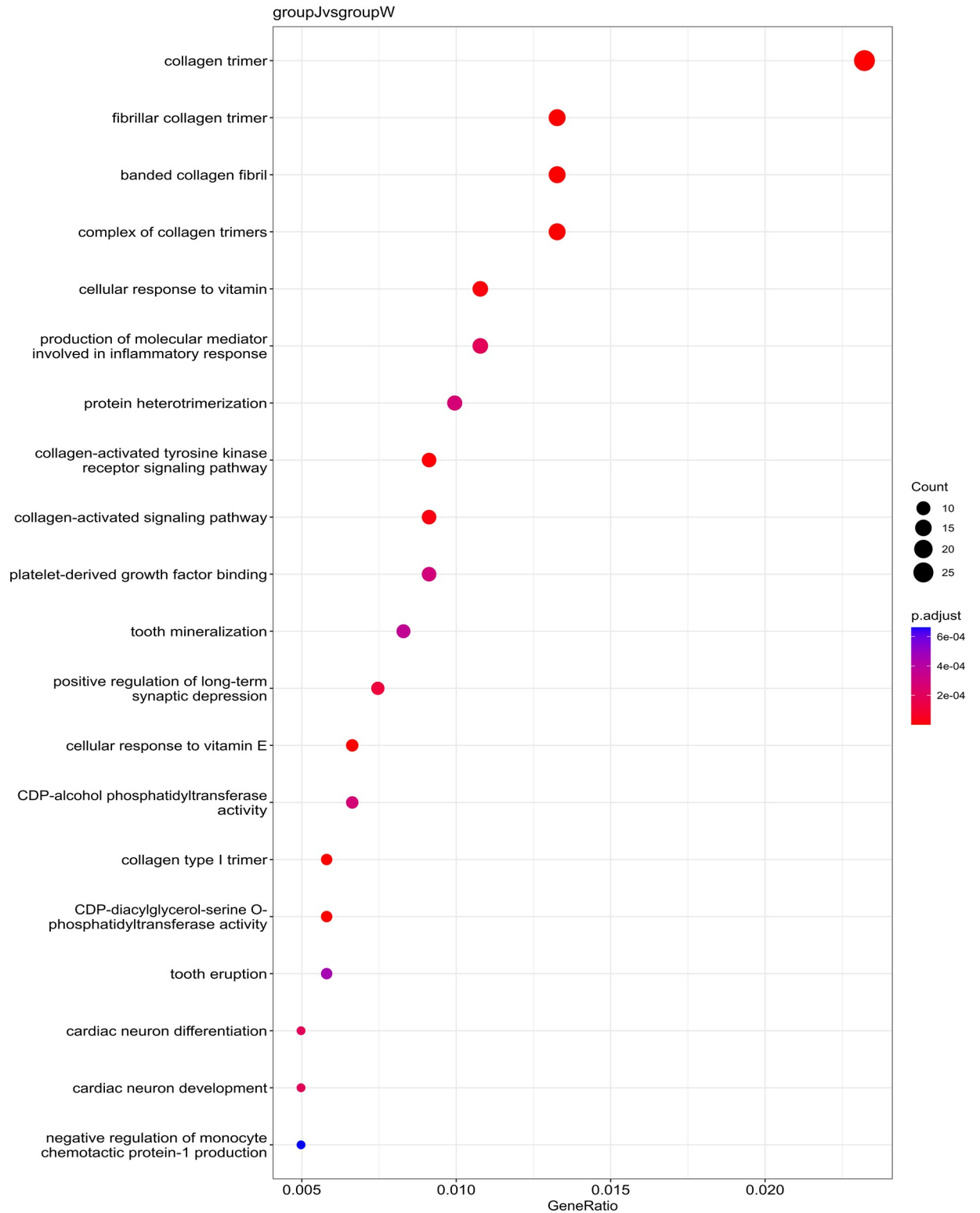
Fig 6. Bubble diagram of the GO enrichment analysis results for group J compared with those for group W. The 20 most significantly enriched GO pathway entries are presented in this figure. The value of p-adjust ranges from 0 to 1, and the closer it is to zero, the more significant the enrichment. GeneRatio is the ratio of the number of differentially expressed genes in pathway entry to the total number of genes involved in pathway entry for all annotated genes.

https://doi.org/10.1371/journal.pone.0315996.g006
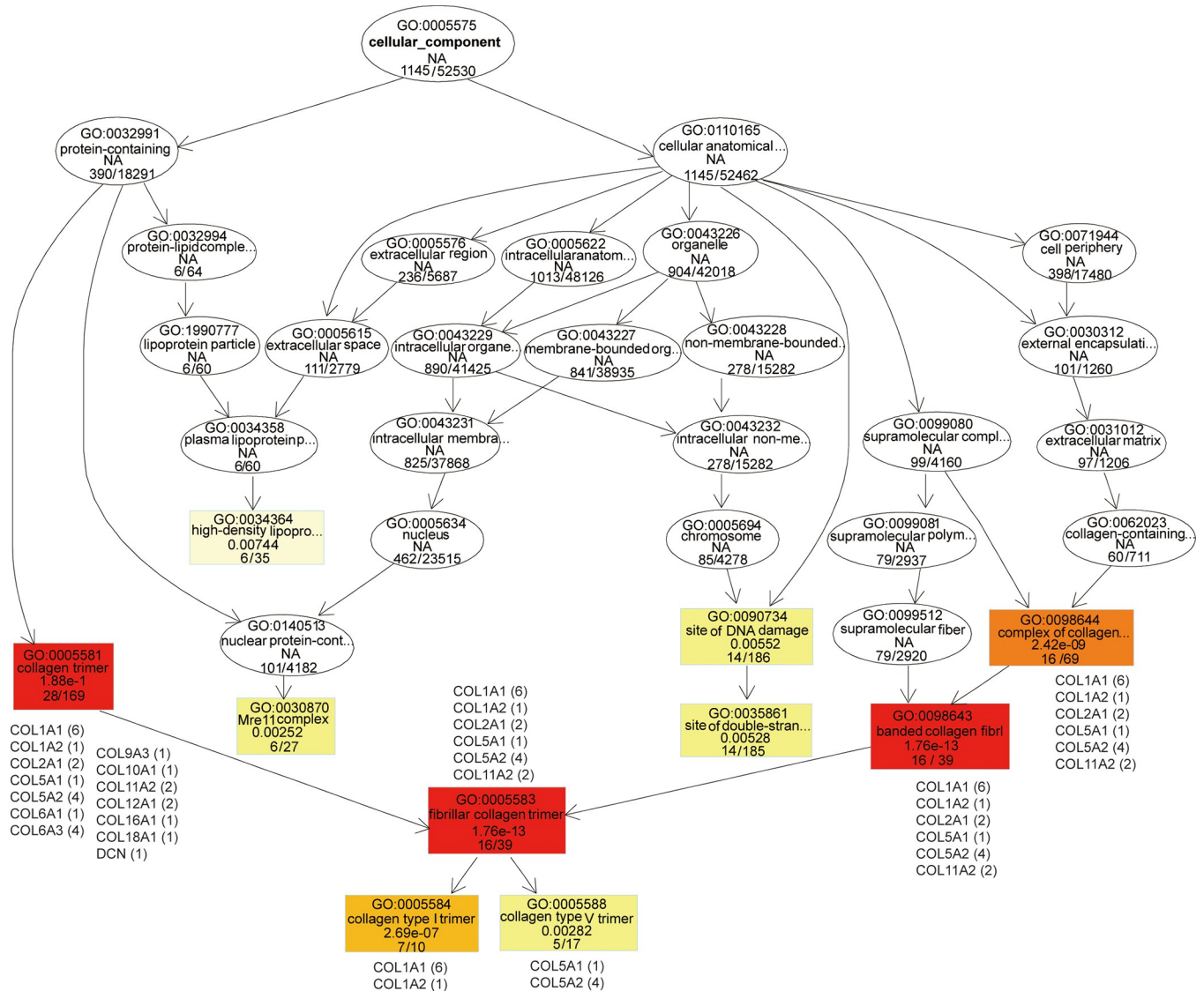
Fig 7. Directed acyclic graph (DAG) of topGO in group J compared to group W for classification of CCs (cell components). The branches represent inclusion relationships and the functional descriptions defined from top to bottom become increasingly specific. Each node represents a GO term. The specific DEGs involved in enrichment are shown alongside functional descriptions. The most important terms are represented by red nodes, whereas the less significant terms are represented by elliptical nodes. Colors represent the degree of enrichment, with darker colors representing higher enrichment.

https://doi.org/10.1371/journal.pone.0315996.g007

3.4 Homologous genes reveal diverse collagen-rich fish families

A homology search in the NCBI database using BLASTn was conducted for each of the 13 collagen-related genes to further predict and understand the relationships among various fish families. A phylogenetic tree was subsequently constructed, revealing two distinct species clusters for each gene ([Fig 9](#)). Analysis of the 13 trees indicated that *C. argus* grouped with homologous genes from other species, implying that these sets are notable sources of collagen ([Fig 9A–9M](#)). Collagen-specific genes, including *COL1A1*, *COL1A2*, *COL6A1*, *COL6A3*, *COL9A3*, *COL11A1*, *COL16A1* and *COL18A1*, showed a close association between *C. argus* and *Lates calcarifer* ([Fig 9A](#)), *Micropterus salmoides* ([Fig 9B](#)), *Scatophagus argus* ([Fig 9F](#)), *Chelmon rostratus* ([Fig 9G](#)), *Seriola lalandi* dorsalis ([Fig 9H](#)), *Dicentrarchus labrax* ([Fig 9J](#)), *Lates calcarifer*
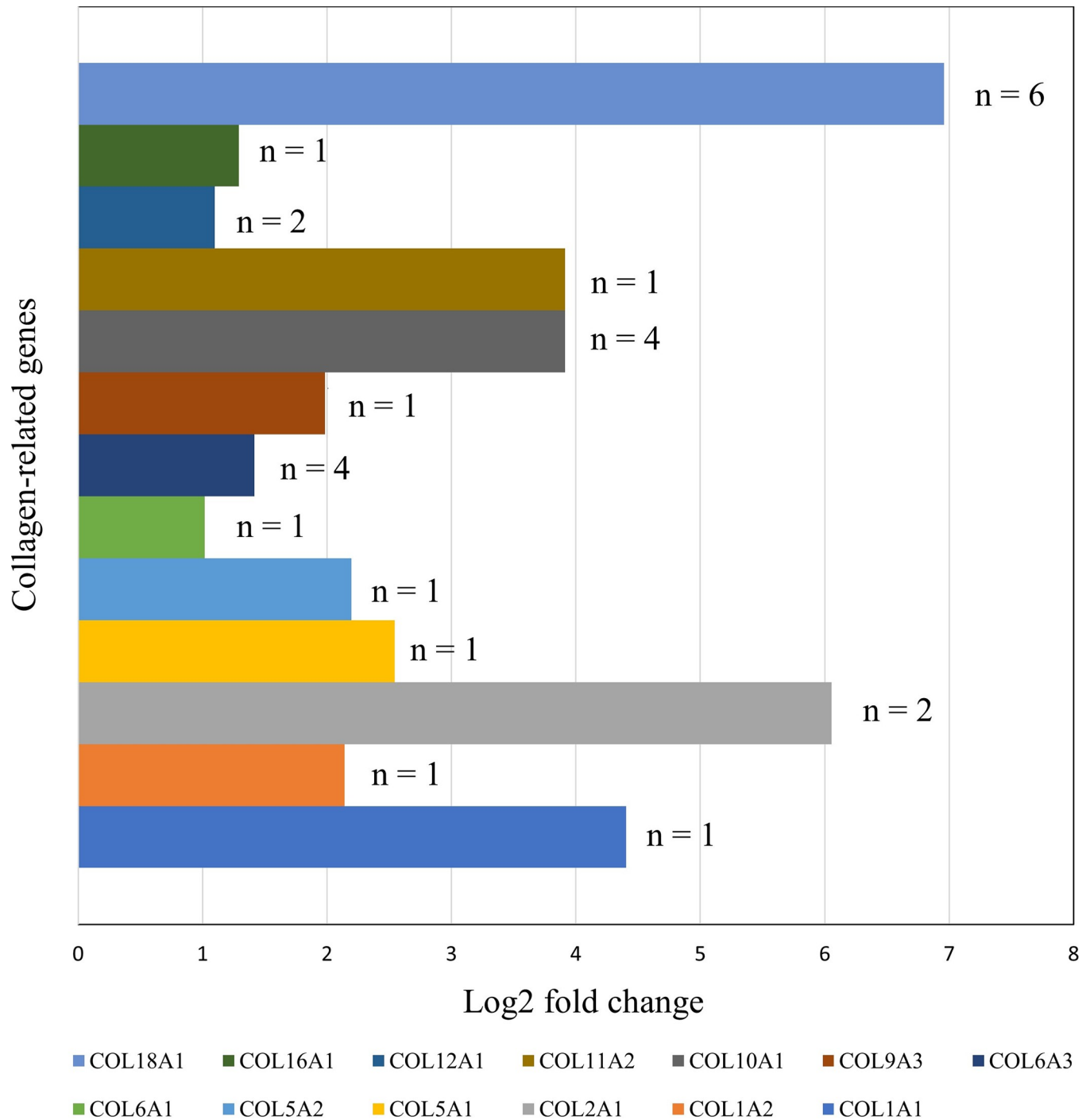
Fig 8. Downregulation of collagen genes in golden-finned *C argus*. Axis: x-axis, gene expression level based on log2-fold change value; y-axis, gene names.

(Fig 9L), and *Solea senegalensis* (Fig 9M). Nonetheless, *C. argus* exhibited a distant genetic link with species concerning the genes *COL2A1* (Fig 9C), *COL5A1* (Fig 9D), *COL5A2* (Fig 9E), *COL10A1* (Fig 9I), and *COL12A1* (Fig 9K).

Interestingly, these homologous genes were found across twenty-four fish families derived from 13 phylogenetic trees, each displaying a unique collagen gene variant (S2 Fig). Fig 9
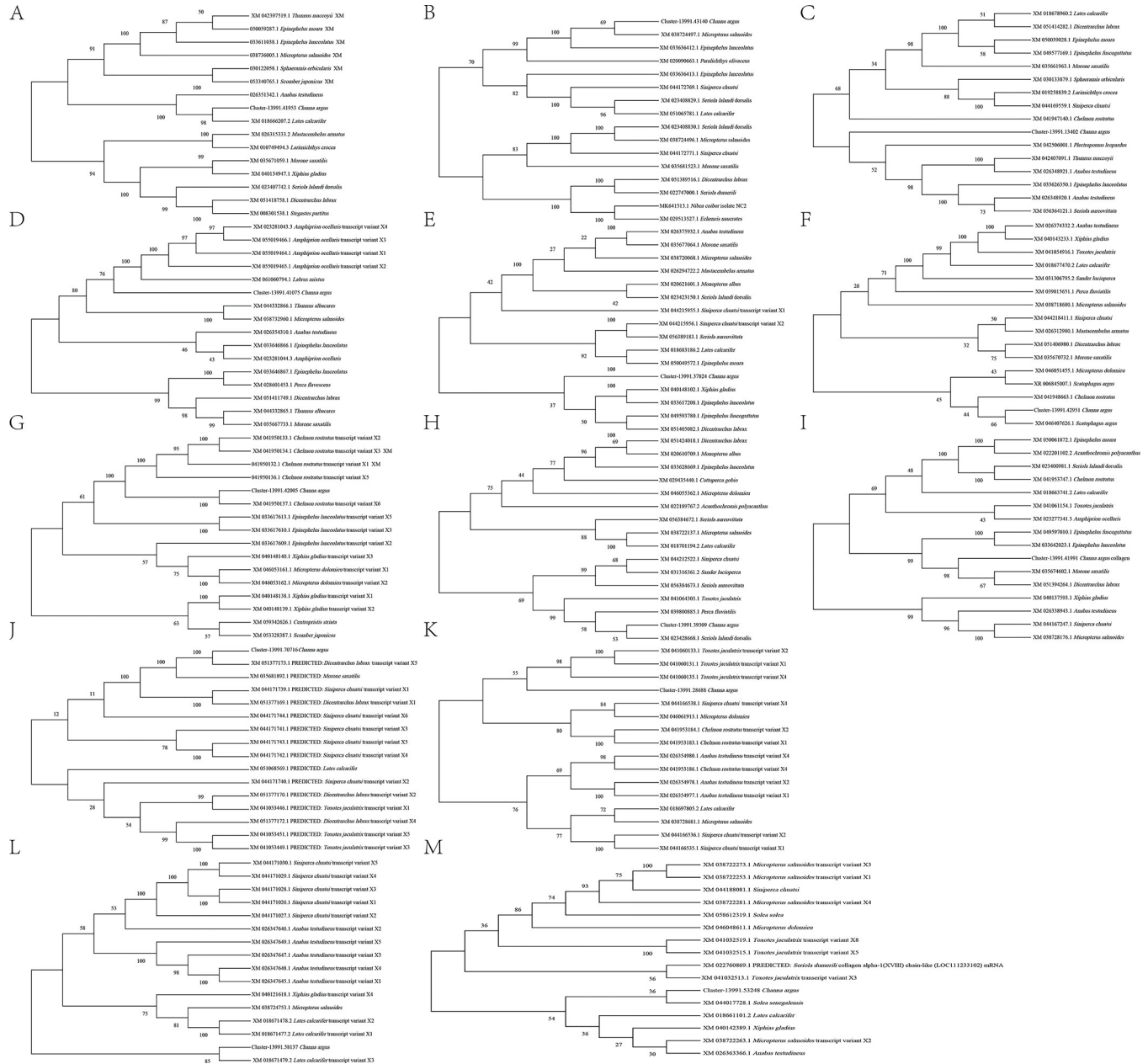
Fig 9. (A-M). Phylogenetic studies of collagen-related genes. These correspond to A. COL1A1, B. COL1A2, C. COL2A1, D. COL5A1, E. COL5A2, F. COL6A1, G. COL6A3, H. COL9A3, I. COL10A1, J. COL11A1, K. COL11A2, L. COL16A1 and M. COL18A1.

illustrates the diversity of collagen-related genes in different species among these families including eleven in Centrarchidae and Latidae; ten in Sinipercidae; nine in Anabantidae and Moronidae; eight in Serranidae; seven in Carangidae and Xiphiidae; six in Toxotidae; five in Chaetodontidae; four in Scombridae; three in Mastacembelidae, Percidae, Pomacentridae and Sciaenidae; two in Apogonidae and Synbranchidae; and one in Bovichtidae, Echeneidae, Labridae, Paralichthyidae, Pomacentridae, Scatophagidae and Soleidae. Distinct type of collagen genes were prominent in various fish families, including COL2A1 in Serranidae (n = 4), COL5A1 in Pomacentridae (n = 5), COL6A3 in Chaetodontidae (n = 5) and Serranidae

(n = 4), COL11A2 in Moronidae (n = 5) and Sinipercidae (n = 6), COL16A1 in Anabantidae (n = 5) and Sinipercidae (n = 5), and COL18A1 in Centrarchidae (n = 5).

4 Discussion

Numerous species of albino fish exist, including *Silurus glanis*, *Corydoras aeneus*, *Danio rerio*, *Hydrolagus colliei*, *Oryzias latipes*, *Macropodus opercularis*, *Rhamdia quelen*, *Ictalurus punctatus*, *Senegalese sole*, *Scophthalmus maximus*, *Paralichthys olivaceus*, *Triplophysa rosa*, *Carassius auratus*, *Apostichopus japonicus* and *C. argus* [54–63]. Recent genomic research on albino fish, including whole-genome sequencing of snakeheads [64] and RNA sequencing, has been performed to analyze the transcriptome associated with skin pigmentation in this species [39]. Our work presents two phenotypes of heritable albino *C. argus*, designated as groups J and H, and compares their DEGs with those of normal *C. argus* (group W). The body coloration of golden-finned albino *C. argus* in group J was markedly different from that in the normal group. Our study concentrated on three specific color variations in snakehead caudal fins and compared their DEGs, which were similar to those of *Aulonocara baenschi*, an albino cichlid fish [65]. The identified and clustered DEGs provided additional evidence of a significant difference between groups J and W in the present investigation. In comparison with the W group, the DEGs in group J showed mostly enriched cytoplasmic processes inside the cellular component (CC), particularly for collagen components. Likewise, the DEGs in the albino yellow catfish (*Pelteobagrus fulvidraco*) [66] and *Oryzias* species [67] accounted for the highest quantity of genes related to cellular processes and cells and binding cellular components.

Additionally, collagen-related genes were found to be downregulated in the present study, representing a significant discovery in understanding the relationship between albinism and collagen. A recent study reported that albino golden snakehead fish possesses five causative genes associated with albinism, namely *TYR* (tyrosinase), *SOX10* (SRY-box transcription factor 10), *S-100* (S100 calcium-binding protein A1), and *NLRC3* (NLR family CARD domain containing 3), which are significantly downregulated in golden albino *C. argus* [39]. The receptor for advanced glycation end products (RAGE) shares its receptors with the S100 family proteins, collagen (type I and IV) and several other ligands. S100A11 signaling via RAGE activates p38 MAPK kinase, thereby enhancing the synthesis of human collagen X [68]. In the present study on albino snakeheads, the downregulation of *COL10A1* may be linked to S100 proteins, leading to reduced expression of collagen genes, necessitating further investigation for confirmation.

Structural collagen genes are tightly associated with skeletogenesis and are expressed in osteoblasts of fish (gars and zebrafish) [69]. Our study uncovered their complex regulatory mechanism. The DEGs related to the collagen component included *COL1A1* and *COL1A2* for collagen I, *COL2A1* for collagen II, *COL5A1* and *COL5A2* for collagen V, *COL6A1* and *COL6A3* for collagen VI, *COL9A3* for collagen IX, *COL10A1* for collagen X, *COL11A2* for collagen XI, *COL12A1* for collagen XII, *COL16A1* for collagen XVI, and *COL18A1* for collagen XVIII, along with decorin (DCN), which were highly enriched in the caudal fins of golden-finned albino *C. argus*. The set of DEGs identified in this study has not been previously documented in other fish species.

Fish collagen derived from diverse sources may contain different components. Collagens from finfish and shellfish are classified as major and minor types I and V, respectively [70]. Types I and V are ideal for biomedical applications because of their superior biocompatibility and low immunogenicity [71]. Fibril-forming collagens are categorized into Types I, II, III and V because a considerable percentage of homologous sequences are not species-dependent [72–74]. Collagen type I trimers consist of two α1 chains and one α2 chain, resulting in triple

helices that form banded fibrils [75–77]. This collagen type constitutes the primary protein component of the bone extracellular matrix, accounting for up to 90% of the organic matrix [78, 79]. These are necessary for bone tissue engineering [80]. It is an essential component of skin tissue and has been utilized for several purposes, including anti-inflammatory properties, skin enhancement, connective tissue strengthening and bone health [81, 82]. In most species, genes encoding for collagen type I trimers are *COL1A1* and *COL1A2*. The data reported in this study may facilitate further investigation of skin and skeletal disorders associated with abnormal type I collagen synthesis. However, in zebrafish, three type I collagen genes have been identified: *COL1A1a*, *COL1A1b*, and *COL1A2*, which encode α1(I), α2(I) and α3(I) chains, respectively [76].

The collagen type V trimer is a minor, low-abundance, yet crucial structural protein present in the extracellular matrix of diverse tissues that governs the morphology and tensile strength of collagen fibrils [83]. It is composed of three α1(V) chains and one α2(V) chain, referred to as the α1(V)2α2(V) heterotrimer. These chains form fibrils that regulate their diameter [70]. Imamura et al. [84] reported that *COL5A1*, *COL5A2*, and *COL5A3* encode collagen type V trimers. This heterotrimer is essential for dermal fibrillogenesis and aids in the formation of a functional skin matrix [85]. The main chain of type V collagen is *COL5A1*, which is expressed in the connective tissue surrounding ray fibroblasts and is partially connected to skeletal precursor cells [86]. It is found in several tissues, including skin, bones, muscles, and blood vessels. Nucleation of type I collagen fibrils relies on type V collagen fibrillar type I collagen assembly in the presence of fibronectin and integrins [87]. Collagen type V trimers promote cell migration and proliferation and participate in several critical biological processes [88–90]. In a joint effort, primary fibril-forming collagens (types I, II and III) together with a minor presence of collagen types V and XI form connective tissues, with types V and XI acting to regulate fibril-forming collagens [91].

Collagen types II, VI, VII, X, XI, XII, XVI, XVIII, and DCN are important proteins that confer structural support to many tissues, including skin, bone, cartilage, and tendons [92]. Collagen II is the key component of cartilage, required for maintaining its tensile strength and elasticity [11, 93]. It is also found in the eyes and other connective tissues [94]. Collagen VI connects the extracellular matrix to cells in several tissues such as muscles, cartilage, and skin [95]. Collagen VII is located in the basement membrane of epithelial cells and binds to the underlying tissues [77]. Collagen XI is present in the cartilage, bone, and cornea of the eye, and contributes to their structural stability. [96]. Furthermore, collagen coexists with type II collagen in cartilage tissues [97]. Collagen X plays a significant role in bone growth and mineralization as well as cartilage formation, particularly in jawed vertebrates [98, 99]. Tendons, ligaments, and other connective tissues contain collagen XII, which facilitates the anchoring of cells to the extracellular matrix [100]. Collagen XVI involves the formation of fibrils inside connective tissues and may also be found in the skin and other organs [101]. Collagen XVIII plays a role in angiogenesis and is present in the cornea of the eye [102]. DCN, or decorin, is a proteoglycan that modulates collagen fiber assembly and functionality across several tissues, such as skin, tendons, and cartilage [103].

While collagen genes in osteoblasts have been extensively studied in humans, mice, and chicks, little is known about these genes in fish [69]. Consequently, our findings concerning collagen-encoded genes expressed in snakehead caudal fins provide a novel avenue for in-depth research on cartilage- or bone-forming cells, with insights into the extracellular matrix composition in finned fish [86]. Moreover, our phylogenetic analysis identified similar marine resources for exploring the potential of collagen-base biomaterials, as highlighted by Senadheera et al. [104] and Pozzolini et al. [105]. The global marine collagen market was evaluated at USD 685 million and USD 633 million in 2020 and 2022, respectively, with projections

indicating growth to USD 1123 million by 2032 [106]. Collagen is highly sought after in industries such as cosmetics, food and nutraceuticals, wound healing, food packaging, pharmaceuticals and biomedicine, with the skin and other byproducts of cattle, pigs and birds serving as key sources [107]. Conversely, fish and other marine resources are gaining popularity [108], due to the reduced risk of disease transmission to humans from mammals and birds. Therefore, they may serve as alternatives to bovine collagen [41, 109]. In tissue engineering and regenerative medicine, type I collagen is widely utilized as the preferred biomaterial for manufacturing cell-instructive scaffolds [110, 111], which are artificial matrices that act as replacements for temporary extracellular matrices. These scaffolds aid in the delivery of specific signals to cells, hence stimulating cellular processes, such as tissue regeneration and matrix remodeling [97]. Improvement and sustained collagen extraction from marine and freshwater fish necessitates an understanding of the molecular basis of collagen synthesis and the identification of genetic resources [112]. Indeed, our present study on *C. argus* explains collagen regulation and may offer valuable molecular and genomic insights into a substantial volume of fish collagen.

Numerous studies have revealed the beneficial and functional properties of fish-derived collagen [9, 25, 42, 113]. Wound-healing properties have been observed in the common *C. argus* (group W) [114, 115]. In contrast, albino fish with a golden caudal fin (group J), utilized in the current investigation, showed wound healing characteristics [25]; yet, a greater number of genes are involved in collagen pathways compared to group W. Our findings on collagen-related genes may enhance the wound-healing capacity of collagen from this albino fish. Although we discovered that collagen-related genes are downregulated in golden albino *C. argus*, they may enrich collagen in areas exhibiting diverse gene expression patterns, necessitating further comprehensive studies. Sveen et al. similarly noted distinct collagen-related gene expression in the caudal fins of Atlantic salmon (*Salmo salar* L.) [116].

Several researchers have used marine-derived biomaterials [117–119]. Our current research has broadened the scope of further investigations to determine the physicochemical properties of collagen in *C. argus*, including its composition, extraction methods, processing factors, molecular and evolutionary foundations, as well as novel physical, chemical, and enzymatic changes in the collagen structure of this species, to develop suitable biomaterials for tissue engineering and other biomedical applications in near future.

5 Conclusion

This study examined two phenotypes of the heritable albino *C. argus* (gray-finned and golden-finned) and compared their DEGs to those of normal-morph individuals. The transcriptome sequencing of *C. argus* yielded 174,779 assembled sequences and 137,130 unigenes. In the golden-finned albino group, DEGs were mostly enriched in gene ontology related to cytoplasm and collagen, a finding that has not been previously documented in fish species. The heatmap indicates that the DEGs were significantly segregated, with black-white and gray-finned albinos clustering together, whereas golden-finned albinos were distinctly grouped apart. A total of 379 common DEGs were identified across all three groups. Furthermore, GO enrichment analysis revealed that both gray-finned and golden-finned albino individuals showed a greater number of downregulated DEGs compared to black-white individuals. In addition, DEGs were predominantly enriched in collagen-related pathways. The collagen components identified included collagen I of *COL1A1* and *COL1A2*, collagen II of *COL2A1*, collagen V of *COL5A1* and *COL5A2*, collagen VI of *COL6A1* and *COL6A3*, collagen IX of *COL9A3*, collagen X of *COL10A1*, collagen XI of *COL11A2*, collagen XII of *COL12A1*, collagen XVI of *COL16A1*, collagen XVIII of *COL18A1* and decorin (DCN), all of which play a role in

modulating the collagen matrix. Notably, the expression of these genes was reduced in golden albino *C. argus*, demonstrating a correlation between albinism and gene expression. These findings open avenues for further research on the regulation of collagen-related genes in fish. Fish collagen, derived from various sources, contains various components, with collagen types I and V being particularly desirable for biomedical applications. Our findings shed light on the genetic and molecular mechanisms underlying collagen metabolism in fish, which may lead to the development of novel biotechnological applications and provide an intuitive understanding of vertebrate regenerative capabilities.

Supporting information

S1 Table. Search results (Hit-tables) of homologous genes for each collagen-related gene of *C. argus* using blastn tool.
(XLSX)

S1 Fig. Percentage of completeness of the core set of genes in BUSCO for the assemblies obtained in this study. Light blue: complete (C) and single copy genes (S); dark blue: complete (C) and duplicated genes (D); yellow: fragmented genes (F); red: missing genes (M).
(TIF)

S2 Fig. Stackbar graphical representation on diversity of genes present in fish families homologous to collagen-related genes of *C. argus*.
(TIFF)

Acknowledgments

Author Contributions

Conceptualization: Shixi Chen, Fardous Mohammad Safiul Azam, Yuanchao Zou.

Data curation: Shixi Chen, Li Ao, Rui Li.

Formal analysis: Shixi Chen, Na Li, Jianlan Wang.

Funding acquisition: Shixi Chen.

Investigation: Li Ao.

Methodology: Shixi Chen, Fardous Mohammad Safiul Azam.

Resources: Ning Li, Fardous Mohammad Safiul Azam, Yuanchao Zou, Rui Li.

Supervision: Yuanchao Zou.

Validation: Li Ao, Na Li, Jianlan Wang.

Visualization: Fardous Mohammad Safiul Azam, Na Li, Jianlan Wang.

Writing – original draft: Shixi Chen, Fardous Mohammad Safiul Azam.

Writing – review & editing: Shixi Chen, Ning Li, Fardous Mohammad Safiul Azam, Zakaria Hossain Prodhan.

References

1. Dunsmore SE. Extracellular Matrix | Collagens. Encyclopedia of Respiratory Medicine, ed. Laurent G. J., Shapiro S.D. Oxford: Academic Press; 2006. pp. 168–175. https://doi.org/10.1016/b0-12-370879-6/00145-9

2. Ando M, Toyohara H, Sakaguchi M. Three-dimensional structure of collagen fibrillar network of pericellular connective tissue in association with firmness of fish muscle. Nippon Suisan Gakk. 1992; 58 (7):1361–1364. https://doi.org/10.2331/suisan.58.1361

3. Lin F, Ye X, Lin J, Liu X, Yuan Y, Guo H, et al. Comparative transcriptome analysis between muscle and swim bladder reveals key genes regulating collagen deposition in zebrafish. Aquaculture Reports. 2022 Apr 1; 23:101053. https://doi.org/10.1016/j.aqrep.2022.101053

4. Johnston IA. The biological basis of variability in the texture of fish flesh. Improving seafood products for the consumer. ed. Børresen T. Cambridge: Woodhead Publishing Limited; 2008. pp. 465–489. https://doi.org/10.1533/9781845694586.5.465

5. Song D, Yun Y, Mi J, Luo J, Jin M, Nie G, et al. Effects of faba bean on growth performance and fillet texture of Yellow River carp, *Cyprinus carpio haematopterus*. Aquaculture Reports. 2020 Jul 1; 17:100379. https://doi.org/10.1016/j.aqrep.2020.100379

6. Tylingo R, Mania S, Panek A, Piątek R, Pawłowicz R. Isolation and characterization of acid soluble collagen from the skin of African catfish (*Clarias gariepinus*), Salmon (*Salmo salar*) and Baltic cod (*Gadus morhua*). Journal of Biotechnology and Biomaterials. 2016; 6(234):2. https://doi.org/10.4172/2155-952X.1000234

7. Sripriya R, Kumar R. A novel enzymatic method for preparation and characterization of collagen film from swim bladder of fish Rohu (*Labeo rohita*). Food and Nutrition Sciences. 2015; 6(15):1468. https://doi.org/10.4236/fns.2015.615151

8. Jeevithan E, Qingbo Z, Bao B, Wu W. Biomedical and pharmaceutical application of fish collagen and gelatin: a review. Journal of Nutritional Therapeutics. 2013; 2(4):218–27. https://doi.org/10.6000/1929-5634.2013.02.04.6

9. Shalaby M, Agwa M, Saeed H, Khedr SM, Morsy O, El-Demellawy MA. Fish scale collagen preparation, characterization and its application in wound healing. Journal of Polymers and the Environment. 2020 Jan; 28:166–78. https://doi.org/10.1007/s10924-019-01594-w

10. Subhan F, Hussain Z, Tauseef I, Shehzad A, Wahid F. A review on recent advances and applications of fish collagen. Critical Reviews in Food Science and Nutrition. 2021 Mar 26; 61(6):1027–37. https://doi.org/10.1080/10408398.2020.1751585 PMID: 32345036

11. Lai CS, Tu CW, Kuo HC, Sun PP, Tsai ML. Type II collagen from cartilage of *Acipenser baerii* promotes wound healing in human dermal fibroblasts and in mouse skin. Marine Drugs. 2020 Oct 11; 18 (10):511. https://doi.org/10.3390/md18100511 PMID: 33050593

12. Gordon MK, Hahn RA. Collagens. Cell Tissue Research, 2010; 339, 247. https://doi.org/10.1007/s00441-009-0844-4 PMID: 19693541

13. Ricard-Blum S. The collagen family. Cold Spring Harbor perspectives in biology. 2011 Jan 1; 3(1): a004978. https://doi.org/10.1101/cshperspect.a004978 PMID: 21421911

14. Felician FF, Xia C, Qi W, Xu H. Collagen from marine biological sources and medical applications. Chemistry & Biodiversity. 2018 May; 15(5):e1700557. https://doi.org/10.1002/cbdv.201700557 PMID: 29521032

15. Liu C, Liu X, Xue Y, Ding T, Sun J. Hydrolyzed tilapia fish collagen modulates the biological behavior of macrophages under inflammatory conditions. RSC Advances. 2015; 5(39):30727–36. https://doi.org/10.1039/c5ra02355f

16. Cheng QT, Zheng BS. Systematic Search of Fish in China. Beijing: Science Press; 1987. pp. 456.

17. Courtenay WR Jr, Williams JD. Snakeheads (Pisces, Channidae): a biological synopsis and risk assessment. US Geological Survey; 2004; 43:1251. https://doi.org/10.3133/cir1251

18. Xu J, Bian C, Chen K, Liu G, Jiang Y, Luo Q, et al. Giga Science. 2017 Apr; 6(4):gix011. https://doi.org/10.1093/gigascience/gix011 PMID: 28327946

19. Jiao Y, Lapointe NW, Angermeier PL, Murphy BR. Hierarchical demographic approaches for assessing invasion dynamics of non-indigenous species: an example using northern snakehead (*Channa argus*). Ecological Modelling. 2009 Jul 17; 220(13–14):1681–9. https://doi.org/10.1016/j.ecolmodel.2009.04.008

20. Duan R, Zhang J, Wu J, Wang J. Preparation and characterization of collagens from the skin of Northern snakehead (*Channa argus*). Journal of Polymers and the Environment. 2018 Mar; 26:867–72. https://doi.org/10.1007/s10924-017-1005-6

21. Chen FD, Yang LF. *Channa argus*: Nutritional Value and the Intensive Processing Development Prospect. Academic Periodical of Farm Products Processing. 2013; 4, 51–3. https://doi.org/10.3969/j.issn.1671-9646(x).2013.04.042

22. Zuraini A, Somchit MN, Solihah MH, Goh YM, Arifah AK, Zakaria MS, et al. Fatty acid and amino acid composition of three local Malaysian *Channa* spp. fish. Food Chemistry. 2006 Aug 1; 97(4):674–8. https://doi.org/10.1016/j.foodchem.2005.04.031

23. Shi BN, Deng QX. The history and formation of fish species in Jialing river. Journal of Southwest China Normal University (Natural Science). 1980; 2:34–44. https://doi.org/10.1378/j.cnki.xsxb.1980.02.004

24. Yang S, Wen ZY, Zou YC, Qin CJ, Wang J, Yuan DY, et al. Molecular cloning, tissue distribution, and effect of fasting and refeeding on the expression of neuropeptide Y in *Channa argus*. General and Comparative Endocrinology. 2018 Apr 1; 259:147–53. https://doi.org/10.1016/j.ygcen.2017.11.017 PMID: 29174870

25. Chen S, Lu M, Xiang H, Xue Y, Huang H, Li L, et al. Research Progress on Nutritional Evaluation of and Processing and Preservation Technologies for *Channa argus*. Meat Research. 2023; 37:40–45.

26. Zhou C, Li Y, Zhou Y, Zou Y, Yuan D, Deng X, et al. Chromosome-Scale Assembly and Characterization of the Albino Northern Snakehead, *Channa argus* var. (Teleostei: Channidae) Genome. Frontiers in Marine Science. 2022 Apr 6; 9:839225. https://doi.org/10.3389/fmars.2022.839225

27. Fan W, Zhang L, Su J, Luo Y, Jiao XL, Huang ZP, et al. Genetic diversity of two color morphs of Northern snakehead (*Channa argus*) unveiled by the mitochondrial DNA D-loop region. Mitochondrial DNA Part B. 2022 Mar 4; 7(3):515–20. https://doi.org/10.1080/23802359.2022.2029601 PMID: 35386959

28. Zhang DX, Kang YH, Song MF, Shu HP, Guo SN, Jia JP, et al. Identity and virulence properties of Aeromonas isolates from healthy Northern snakehead (*Channa argus*) in China. Letters in Applied Microbiology. 2019 Aug 1; 69(2):100–9. https://doi.org/doi.org/10.1111/lam.13172

29. Sun D, Qi X, Wen H, Li C, Li J, Chen J, et al. The genetic basis and potential molecular mechanism of yellow-albino northern snakehead (*Channa argus*). Open Biology. 2023 Feb 15; 13(2):220235. https://doi.org/10.1098/rsob.220235 PMID: 36789536

30. Wang L, Xie N, Shen Y, Ye B, Yue GH, Feng X. Constructing high-density genetic maps and developing sexing markers in northern snakehead (*Channa argus*). Marine Biotechnology. 2019 Jun 15; 21:348–58. https://doi.org/10.1007/s10126-019-09884-z PMID: 30888532

31. Qin CJ, Wen ZY, Wang J, He Y, Yuan DY, Li R. Uncoupling protein 1 in snakehead (*Channa argus*): Cloning, tissue distribution, and its expression in response to fasting and refeeding. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology. 2018 Nov 1; 225:1–6. https://doi.org/10.1016/j.cbpa.2018.06.010 PMID: 29886254

32. Wu F, Zhang G, Huo Y, Xiong S, Du H. Rheology and texture properties of surimi gels of northern snakehead (*Channa Argus*) as Affected by *Angelica Sinensis* (Oliv.) Diels. (Danggui) Powder. Journal of Aquatic Food Product Technology. 2018 Apr 21; 27(4):486–95. https://doi.org/10.1080/10498850.2018.1449151

33. Amparyup P, Charoensapsri W, Samaluka N, Chumtong P, Yocawibun P, Imjongjirak C. Transcriptome analysis identifies immune-related genes and antimicrobial peptides in Siamese fighting fish (*Betta splendens*). Fish & Shellfish Immunology. 2020 Apr 1; 99:403–13. https://doi.org/10.1016/j.fsi.2020.02.030 PMID: 32081810

34. Lu X, Chen HM, Qian XQ, Gui JF. Transcriptome analysis of grass carp (*Ctenopharyngodon idella*) between fast-and slow-growing fish. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics. 2020 Sep 1; 35:100688. https://doi.org/10.1016/j.cbd.2020.100688 PMID: 32454298

35. Wentzel AS, Petit J, van Veen WG, Fink IR, Scheer MH, Piazzon MC, et al. Transcriptome sequencing supports a conservation of macrophage polarization in fish. Scientific Reports. 2020 Aug 10; 10(1):13470. https://doi.org/10.1038/s41598-020-70248-y PMID: 32778701

36. Wang J, Chen G, Yu X, Zhou X, Zhang Y, Wu Y, et al. Transcriptome analyses reveal differentially expressed genes associated with development of the palatal organ in bighead carp (*Hypophthalmichthys nobilis*). Comparative Biochemistry and Physiology Part D: Genomics and Proteomics. 2023 Jun 1; 46:10107. https://doi.org/10.1016/j.cbd.2023.101072 PMID: 36990038

37. Vo TT, Amoroso G, Ventura T, Elizur A. Histological and transcriptomic analysis of muscular atrophy associated with depleted flesh pigmentation in Atlantic salmon (*Salmo salar*) exposed to elevated seawater temperatures. Scientific Reports. 2023 Mar 14; 13(1):4218. https://doi.org/10.1038/s41598-023-31242-2 PMID: 36918611

38. Lin F, Yuan Y, Ye X, Lv J, Liu X, Guo H, et al. Characterization and role of connective tissue growth factor gene in collagen synthesis in swim bladder of chu's croaker (*Nibea coibor*). International Journal of Biological Macromolecules. 2023 Feb 1; 227:1336–45. https://doi.org/10.1016/j.ijbiomac.2022.12.010 PMID: 36473534

39. Mao L, Zhu Y, Yan J, Zhang L, Zhu S, An L, et al. Full-length transcriptome sequencing analysis reveals differential skin color regulation in snakeheads fish *Channa argus*. Aquaculture and Fisheries. 2024. 9, 4: 590–596. https://doi.org/10.1016/j.aaf.2022.11.008

40. Zhou A, Xie S, Feng Y, Sun D, Liu S, Sun Z, et al. Insights into the albinism mechanism for two distinct color morphs of Northern Snakehead, *Channa argus* through histological and transcriptome analyses. Frontiers in Genetics. 2020 Sep 18; 11:830. https://doi.org/10.3389/fgene.2020.00830 PMID: 33193565

41. Xu N, Peng XL, Li HR, Liu JX, Cheng JS, Qi XY, et al. Marine-derived collagen as biomaterials for human health. Frontiers in Nutrition. 2021 Aug 24; 8:702108. https://doi.org/10.3389/fnut.2021.702108 PMID: 34504861

42. Geahchan S, Baharlouei P, Rahman A. Marine collagen: a promising biomaterial for wound healing, skin anti-aging, and bone regeneration. Marine Drugs. 2022 Jan 10; 20(1):61. https://doi.org/10.3390/md20010061 PMID: 35049916

43. Percie du Sert N, Ahluwalia A, Alam S, Avey MT, Baker M, Browne WJ, et al. Reporting animal research: Explanation and elaboration for the ARRIVE guidelines 2.0. PLoS Biol. 2020; 18(7): e3000411. https://doi.org/10.1371/journal.pbio.3000411 PMID: 32663221

44. Thrash A, Arick M II, Peterson DG. Quack: A quality assurance tool for high throughput sequence data. Analytical Biochemistry. 2018 May 1; 548:38–43. https://doi.org/10.1016/j.ab.2018.01.028 PMID: 29410015

45. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug 1; 30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

46. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nature Biotechnology. 2011 Jul; 29 (7):644. https://doi.org/10.1038/nbt.1883

47. Seppey M, Manni M, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness. Gene prediction: methods and protocols. 2019:227–45. https://doi.org/10.1007/978-1-4939-9173-0_14

48. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011 Dec; 12:1–6. https://doi.org/10.1186/1471-2105-12-323

49. Davidson NM, Oshlack A. Corset: enabling differential gene expression analysis for de novoassembled transcriptomes. Genome Biology. 2014 Jul; 15(7):1–4. https://doi.org/10.1186/s13059-014-0410-6 PMID: 25063469

50. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014 Dec; 15(12):1–21. https://doi.org/10.1186/s13059-014-0550-8 PMID: 25516281

51. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Research. 2023 Jan 6; 51(D1):D587–92. https://doi.org/10.1093/nar/gkac963 PMID: 36300620

52. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Research. 2019 Jan 8; 47(D1):D419–26. https://doi.org/10.1093/nar/gky1038 PMID: 30407594

53. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution. 2021, 38:3022–3027. https://doi.org/10.1093/molbev/msab120 PMID: 33892491

54. Bondari K. Performance of albino and normal channel catfish (*Ictalurus punctatus*) in different water temperatures. Aquaculture Research. 1984 Jul; 15(3):131–40. https://doi.org/10.1111/j.1365-2109.1984.tb00844.x

55. Dalton HC, Goodrich HB. Chromatophore Reactions in the Normal and Albino Paradise Fish, *Macropodus opercularis* L. The Biological Bulletin. 1937 Dec; 73(3):535–41. https://doi.org/10.2307/1537612

56. Fabacher DL, Little EE, Ostrander GK. Tolerance of an albino fish to ultraviolet-B radiation. Environmental Science and Pollution Research. 1999 Jun; 6:69–71. https://doi.org/10.1007/BF02987550 PMID: 19009405

57. Parodi TV, Cunha MA, Becker AG, Zeppenfeld CC, Martins DI, Koakoski G, et al. Anesthetic activity of the essential oil of Aloysia triphylla and effectiveness in reducing stress during transport of albino and gray strains of silver catfish, Rhamdia quelen. Fish Physiology and Biochemistry. 2014 Apr; 40:323–34. https://doi.org/10.1007/s10695-013-9845-z PMID: 23974669

58. Reum JC, Paulsen CE, Pietsch TW, Parker-Stetter SL. First record of an albino chimaeriform fish, *Hydrolagus colliei*. Northwestern Naturalist. 2008 Apr 1:60–2. https://doi.org/10.1898/1051-1733(2008)89[60:froaac]2.0.co;2

59. Shiraki T, Kawakami K. A tRNA-based multiplex sgRNA expression system in zebrafish and its application to generation of transgenic albino fish. Scientific Reports. 2018 Sep 6; 8(1):13366. https://doi.org/10.1038/s41598-018-31476-5 PMID: 30190522

60. Slavík O, Horký P, Valchářová T, Pfauserová N, Velíšek J. Comparative study of stress responses, laterality and familiarity recognition between albino and pigmented fish. Zoology. 2022 Feb 1; 150:125982. https://doi.org/10.1016/j.zool.2021.125982 PMID: 34896758

61. Lechner W, Ladich F. How do albino fish hear? Journal of Zoology. 2011 Mar; 283(3):186–92. https://doi.org/10.1111/j.1469-7998.2010.00762.x PMID: 21552308

62. Xing L, Sun L, Liu S, Li X, Zhang L, Yang H. Transcriptome analysis provides insights into the mechanism of albinism during different pigmentation stages of the albino sea cucumber *Apostichopus japonicus*. Aquaculture. 2018 Feb 3; 486:148–60. https://doi.org/10.1016/j.aquaculture.2017.12.016

63. Wishkerman A, Boglino A, Darias MJ, Andree KB, Estévez A, Gisbert E. Image analysis-based classification of pigmentation patterns in fish: A case study of pseudo-albinism in Senegalese sole. Aquaculture. 2016 Nov 1; 464:303–8. https://doi.org/10.1016/j.aquaculture.2016.06.040

64. Resh CA, Galaska MP, Mahon AR. Genomic analyses of Northern snakehead (*Channa argus*) populations in North America. PeerJ. 2018 Apr 4; 6:e4581. https://doi.org/10.7717/peerj.4581 PMID: 29637024

65. Lee SY, Lee HJ. Comprehensive RNA-seq analysis to evaluate the pigmentation-related genes involved in albinism of cichlid fish, *Aulonocara baenschi*. Frontiers in Marine Science. 2020 Aug 25; 7:723. https://doi.org/10.3389/fmars.2020.00723

66. Zou M, Zhang X, Shi Z, Lin L, Ouyang G, Zhang G, et al. A comparative transcriptome analysis between wild and albino yellow catfish (*Pelteobagrus fulvidraco*). PLoS One. 2015 Jun 26; 10(6): e0131504. https://doi.org/10.1371/journal.pone.0131504 PMID: 26114548

67. Ngamniyom A. Molecular assessment and transcriptome profiling of wild fish populations of *Oryzias mekongensis* and *O. songkhramensis* (Adrianichthyidae: Beloniformes) from Thailand. PLoS One. 2020 Nov 19; 15(11):e0242382. https://doi.org/10.1371/journal.pone.0242382 PMID: 33211755

68. He H, Li J, Weng S, Li M, Yu Y. S100A11: Diverse Function and Pathology Corresponding to Different Target Proteins. Cell Biochemistry and Biophysics. 2009, 55:117–126 (2009). https://doi.org/10.1007/s12013-009-9061-8 PMID: 19649745

69. Eames BF, Amores A, Yan YL, Postlethwait JH. Evolution of the osteoblast: skeletogenesis in gar and zebrafish. BMC Evolutionary Biology. 2012 Dec; 12(1):1–3. https://doi.org/10.1186/1471-2148-12-27 PMID: 22390748

70. Benjakul S, Nalinanon S, Shahidi F. "Fish Collagen" in Food Biochemistry and Food Processing, ed. Simpson B.K. (Ames, Iowa: Wiley-Blackwell), 2012;365–387. https://doi.org/10.1002/9781118308035.ch20

71. Meena C, Mengi SA, Deshpande SG. Biomedical and industrial applications of collagen. In Proceedings of the Indian Academy of Sciences-Chemical Sciences; 1999 Apr: 111:319–329. Springer India. https://doi.org/10.1007/bf02871912

72. Timpl R. "Immunology of the collagens" in Extracellular matrix biochemistry, ed. Piez K.A., Reddi A.H. New York: Elsevier. 1984. pp. 159–190.

73. Krieg T, Aumailley M, Koch M, Chu M, Uitto J. Collagens, elastic fibers, and other extracellular matrix proteins of the dermis. Fitzpatrick's dermatology in general medicine. ed. Goldsmith L.A, Katz S.I., Gilchrest B.A, Paller A.S, Leffell D.J, Wolff K. New York: McGraw Hill. 2012. (Accessed May 04, 2023).

74. Bella J, Hulmes D. "Fibrillar Collagens", in Fibrous Proteins: Structures and Mechanisms, Subcellular Biochemistry, vol 82., ed. Parry D.A.D., Squire J.M. Cham: Springer. 2017. pp. 457–490. https://doi.org/10.1007/978-3-319-49674-0_14 PMID: 28101870

75. Sorushanova A, Delgado LM, Wu Z, Shologu N, Kshirsagar A, Raghunath R, et al. The collagen suprafamily: from biosynthesis to advanced biomaterial development. Advanced Materials. 2019 Jan; 31 (1):1801651. https://doi.org/10.1002/adma.201801651 PMID: 30126066

76. Gistelinck C, Gioia R, Gagliardi A, Tonelli F, Marchese L, Bianchi L, et al. Zebrafish collagen type I: molecular and biochemical characterization of the major structural protein in bone and skin. Scientific Reports. 2016 Feb 15; 6(1):21540. https://doi.org/10.1038/srep21540 PMID: 26876635

77. Karsdal M. Biochemistry of collagens, laminins and elastin: structure, function and biomarkers. Amsterdam: Elsevier. 2019 May 28. https://doi.org/10.1016/c2015-0-05547-2

78. Maia FR, Correlo VM, Oliveira JM, Reis RL. Natural origin materials for bone tissue engineering: Properties, processing, and performance. In Principles of Regenerative Medicine. Boston: Academic Press. 2019 Jan 1. pp. 535–558. https://doi.org/10.1016/b978-0-12-381422-7.10032-x

79. Fedarko NS. Osteoblast/Osteoclast Development and Function in Osteogenesis Imperfecta. Osteogenesis Imperfecta. 2013; 45–56. https://doi.org/10.1016/b978-0-12-397165-4.00005–8

80. Nguyen TT, Gobinet C, Feru J, Pasco SB, Manfait M, Piot O. Characterization of type I and IV collagens by Raman microspectroscopy: Identification of spectral markers of the dermo-epidermal junction. Journal of Spectroscopy. 2012 Jan 1; 27(5–6):421–7. https://doi.org/10.1155/2012/686183

81. Katsnelson A. Cosmetics: molecular beauty. Nature. 2015 Oct 8; 526(7572):S4–5. https://doi.org/10.1038/526S4a PMID: 26444373

82. Subhan F, Kang HY, Lim Y, Ikram M, Baek SY, Jin S, et al. Fish scale collagen peptides protect against CoCl 2/TNF-α-induced cytotoxicity and inflammation via inhibition of ROS, MAPK, and NF-κB pathways in HaCaT cells. Oxidative Medicine and Cellular Longevity. 2017 Oct; 2017: 1–17. https://doi.org/10.1155/2017/9703609 PMID: 28717410

83. Huang G, Ge G, Wang D, Gopalakrishnan B, Butz DH, Colman RJ, et al. Greenspan DS. α3 (V) collagen is critical for glucose homeostasis in mice due to effects in pancreatic islets and peripheral tissues. The Journal of Clinical Investigation. 2011 Feb 1; 121(2):769–83. https://doi.org/10.1172/jci45096 PMID: 21293061

84. Imamura Y, Scott IC, Greenspan DS. The Pro-α3 (V) Collagen Chain: complete primary structure, expression domains in adult and developing tissues, and comparison to the structures and expression domains of the other types V and XI procollagen chains. Journal of Biological Chemistry. 2000 Mar 24; 275(12):8749–59. https://doi.org/10.1074/jbc.275.12.8749 PMID: 10722718

85. Jarrold BB, Osborne R, Tamura M, Chan W, Bascom CC, Isfort RJ, et al. Collagen type V expression and localization in aging skin. Journal Of The American Academy Of Dermatology. 2017 Jun 1; 76(6): AB75–AB75. https://doi.org/10.1016/j.jaad.2017.04.308

86. Duran I, Csukasi F, Taylor SP, Krakow D, Becerra J, Bombarely A, et al. Collagen duplicate genes of bone and cartilage participate during regeneration of zebrafish fin skeleton. Gene Expression Patterns. 2015 Sep 1; 19(1–2):60–9. https://doi.org/10.1016/j.gep.2015.07.004 PMID: 26256560

87. Potekaev NN, Borzykh OB, Medvedev GV, Petrova MM, Gavrilyuk OA, Karpova EI, et al. Genetic and epigenetic aspects of skin collagen fiber turnover and functioning. Cosmetics. 2021 Sep 29; 8(4):92. https://doi.org/10.3390/cosmetics8040092

88. Babraj JA, Cuthbertson DJ, Smith K, Langberg H, Miller B, Krogsgaard MR, et al. Collagen synthesis in human musculoskeletal tissues and skin. American Journal of Physiology-Endocrinology and Metabolism. 2005 Nov; 289(5):E864–9. https://doi.org/10.1152/ajpendo.00243.2005 PMID: 15972270

89. Birk DE. Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. Micron. 2001 Apr 1; 32(3):223–37. https://doi.org/10.1016/s0968-4328(00)00043-3 PMID: 11006503

90. Chanut-Delalande H, Bonod-Bidaud C, Cogne S, Malbouyres M, Ramirez F, Fichard A, et al. Development of a functional skin matrix requires deposition of collagen V heterotrimers. Molecular and Cellular Biology. 2004 Jul 1; 24(13):6049–57. https://doi.org/10.1128/MCB.24.13.6049-6057.2004 PMID: 15199158

91. Birk DE, Brückner P. Collagens, suprastructures, and collagen fibril assembly. InThe extracellular matrix: an overview. Heidelberg: Springer. 2010 Dec 10. pp. 77–115. https://doi.org/10.1007/978-3-642-16555-9_3

92. Jafari H, Lista A, Siekapen MM, Ghaffari-Bohlouli P, Nie L, Alimoradi H, et al. Fish collagen: Extraction, characterization, and applications for biomaterials engineering. Polymers. 2020 Sep 28; 12(10):2230. https://doi.org/10.3390/polym12102230 PMID: 32998331

93. Eyre D. Articular cartilage and changes in arthritis: collagen of articular cartilage. Arthritis Research & Therapy. 2001 Oct; 4:1–6. https://doi.org/10.1186/1471-2148-12-27 PMID: 22390748

94. Noma Y, Asakawa Y. "Biotransformation of Monoterpenoids", in Comprehensive Natural Products II, ed. Liu H-W. Oxford: Elsevier. 2010. pp. 669–801. https://doi.org/10.1016/b978-008045382-8.00742–5

95. Zanotti S, Magri F, Salani S, Napoli L, Ripolone M, Ronchi D, et al. Extracellular Matrix Disorganization and Sarcolemmal Alterations in COL6-Related Myopathy Patients with New Variants of COL6 Genes. International Journal of Molecular Sciences. 2023 Mar 14; 24(6):5551. https://doi.org/10.3390/ijms24065551 PMID: 36982625

96. Luo YY, Karsdal MA. "Type XI Collagen", in Biochemistry of Collagens, Laminins and Elastin: structure, function and biomarkers, ed. Karsdal M.A. Cambridge: Academic Press. 2019 may 28. pp. 77–80. https://doi.org/10.1016/b978-0-12-809847-9.00011–8

97. Salvatore L, Gallo N, Natali ML, Terzi A, Sannino A, Madaghiele M. Mimicking the hierarchical organization of natural collagen: Toward the development of ideal scaffolding material for tissue regeneration. Frontiers in Bioengineering and Biotechnology. 2021 Apr 27; 9:644595. https://doi.org/10.3389/fbioe.2021.644595 PMID: 33987173

98. Shen G. The role of type X collagen in facilitating and regulating endochondral ossification of articular cartilage. Orthodontics & Craniofacial Research. 2005 Feb; 8(1):11–7. https://doi.org/10.1111/j.1601-6343.2004.00308.x PMID: 15667640

99. Debiais-Thibaud M, Simion P, Ventéo S, Muñoz D, Marcellini S, Mazan S, et al. Skeletal mineralization in association with type X collagen expression is an ancestral feature for jawed vertebrates. Molecular Biology and Evolution. 2019 Oct 1; 36(10):2265–76. https://doi.org/10.1093/molbev/msz145 PMID: 31270539

100. Manon-Jensen T, Karsdal MA. "Type XII Collagen", in Biochemistry of Collagens, Laminins and Elastin: structure, function and biomarkers, ed. Karsdal M.A. Cambridge: Academic Press. 2016. pp. 81–85. https://doi.org/10.1016/b978-0-12-809847-9.00012-x

101. Purcel G, Meliță D, Andronescu E, Grumezescu AM. Collagen-based nanobiomaterials: Challenges in soft tissue engineering. In Nanobiomaterials in Soft Tissue Engineering. Oxford: William Andrew Publishing. 2016 Jan 1. pp. 173–200. https://doi.org/10.1016/b978-0-323-42865-1.00007–6

102. Bächinger HP, Mizuno K, Vranka JA, Boudko SP. "Collagen Formation and Structure", in Comprehensive Natural Products II, ed. Liu H.B. and Mander L. Oxford: Elsevier. 2010. pp. 469–530. https://doi.org/10.1016/b978-008045382-8.00698–5

103. Järvinen TA, Prince S. Decorin: a growth factor antagonist for tumor growth inhibition. BioMed Research International. 2015 Oct; 2015. https://doi.org/10.1155/2015/654765 PMID: 26697491

104. Senadheera TRL, Dave D, Shahidi F. Sea Cucumber Derived Type I Collagen: A Comprehensive Review. *Marine Drugs*. 2020, 18: 471. https://doi.org/10.3390/md18090471 PMID: 32961970

105. Pozzolini M, Scarfì S., Giovine M., 2020. Marine Collagen and its Biotechnological Applications. Encyclopedia of Marine Biotechnology, 2020 2:1007–1030. https://doi.org/10.1002/9781119143802.ch39

106. Rajabimashhadi Z, Gallo N, Salvatore L, Lionetto F. Collagen derived from fish industry waste: progresses and challenges. Polymers. 2023 Jan 20; 15(3):544. https://doi.org/10.3390/polym15030544 PMID: 36771844

107. León-López A, Morales-Peñaloza A, Martínez-Juárez VM, Vargas-Torres A, Zeugolis DI, Aguirre-Álvarez G. Hydrolyzed collagen—Sources and applications. Molecules. 2019 Nov 7; 24(22):4031. https://doi.org/10.3390/molecules24224031 PMID: 31703345

108. Mearns-Spragg A, Tilman J, Tams D, Barnes A. The biological evaluation of jellyfish collagen as a new research tool for the growth and culture of iPSC derived microglia. Frontiers in Marine Science. 2020 Aug 31; 7:689. https://doi.org/10.3389/fmars.2020.00689

109. Lupi O. Prions in dermatology. Journal of the American Academy of Dermatology. 2002 May 1; 46(5):790–3. https://doi.org/10.1067/mjd.2002.120624 PMID: 12004327

110. Walters BD, Stegemann JP. Strategies for directing the structure and function of three-dimensional collagen biomaterials across length scales. Acta Biomaterialia. 2014 Apr 1; 10(4):1488–501. https://doi.org/10.1016/j.actbio.2013.08.038 PMID: 24012608

111. Terzi A, Gallo N, Bettini S, Sibillano T, Altamura D, Madaghiele M, et al. Sub-and supramolecular X-ray characterization of engineered tissues from equine tendon, bovine dermis, and fish skin type-I collagen. Macromolecular Bioscience. 2020 May; 20(5):2000017. https://doi.org/10.1002/mabi.202000017 PMID: 32163225

112. Reátegui-Pinedo N, Salirrosas D, Sánchez-Tuesta L, Quiñones C, Jáuregui-Rosas SR, Barraza G, et al. Characterization of collagen from three genetic lines (gray, red and F1) of *Oreochromis niloticus* (tilapia) skin in young and old adults. Molecules. 2022 Feb 8; 27(3):1123. https://doi.org/10.3390/molecules27031123 PMID: 35164387

113. Elbialy ZI, Atiba A, Abdelnaby A, Al-Hawary II, Elsheshtawy A, El-Serehy HA, et al. Collagen extract obtained from Nile tilapia (*Oreochromis niloticus* L.) skin accelerates wound healing in rat model via up regulating VEGF, bFGF, and α-SMA genes expression. BMC Veterinary Research. 2020 Dec; 16(1):1–1.

114. Sahid NA, Hayati F, Rao CV, Ramely R, Sani I, Dzulkarnaen A, et al. Snakehead consumption enhances wound healing? From tradition to modern clinical practice: A prospective randomized controlled trial. Evidence-Based Complementary and Alternative Medicine. 2018 Jan 1;2018:3032790. https://doi.org/10.1155/2018/3032790 PMID: 30538757

115. Baie SH, Sheikh KA. The wound healing properties of *Channa striatus*-cetrimide cream—tensile strength measurement. Journal of Ethnopharmacology. 2000 Jul 1; 71(1–2):93–100. https://doi.org/10.1016/s0378-8741(99)00184-1 PMID: 10904151

116. Sveen LR, Robinson N, Krasnov A, Daniels RR, Vaadal M, Karlsen C, et al. Transcriptomic landscape of Atlantic salmon (Salmo salar L.) skin. *G3*: *Genes*, *Genomes*, *Genetics*. 2023 November 13(11), p. jkad215. https://doi.org/10.1093/g3journal/jkad215 PMID: 37724757

117. Felician FF, Yu RH, Li MZ, Li CJ, Chen HQ, Jiang Y, et al. The wound healing potential of collagen peptides derived from the jellyfish Rhopilema esculentum. Chinese Journal of Traumatology. 2019 Feb 1; 22(01):12–20. https://doi.org/10.1016/j.cjtee.2018.10.004 PMID: 30827814

118. Zata HF, Chiquita P, Shafira K. Collagen from marine source for regenerative therapy: A literature review. In AIP Conference Proceedings; 2020 Dec 9 (Vol. 2314, No. 1). AIP Publishing. https://doi.org/10.1063/5.0036110

119. Yamada S, Yamamoto K, Nakazono A, Matsuura T, Yoshimura A. Functional roles of fish collagen peptides on bone regeneration. Dental Materials Journal. 2021 Nov 25; 40(6):1295–302. https://doi.org/10.4012/dmj.2020-446 PMID: 34334505