

Article

High Throughput Identification of Antihypertensive Peptides from Fish Proteome Datasets

Yunhai Yi ^{1,2} , Yunyun Lv ^{2,3}, Lijun Zhang ¹, Jian Yang ^{1,*} and Qiong Shi ^{2,3,*} 

¹ School of Applied Chemistry and Biotechnology, Shenzhen Polytechnic, Shenzhen 518055, China; yiyunhai@genomics.cn (Y.Y.); c7zlj@szpt.edu.cn (L.Z.)

² BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, China; lvyunyun@genomics.cn

³ Shenzhen Key Lab of Marine Genomics, Guangdong Provincial Key Lab of Molecular Breeding in Marine Economic Animals, BGI Academy of Marine Sciences, BGI Marine, BGI, Shenzhen 518083, China

* Correspondence: jiany@szpt.edu.cn (J.Y.); shiqiong@genomics.cn (Q.S.); Tel.: +86-136-265-512-99 (J.Y.); +86-185-662-798-26 (Q.S.)

Received: 30 August 2018; Accepted: 24 September 2018; Published: 2 October 2018



Abstract: Antihypertensive peptides (AHTPs) are a group of small peptides with the main role to block key enzymes or receptors in the angiotensin genesis pathway. A great number of AHTPs have been isolated or digested from natural food resources; however, comprehensive studies on comparisons of AHTPs in various species from the perspective of big data are rare. Here, we established a simplified local AHTP database, and performed in situ mapping for high throughput identification of AHTPs with high antihypertensive activity from high-quality whole proteome datasets of 18 fish species. In the 35 identified AHTPs with reported high activity, we observed that Gly-Leu-Pro, Leu-Pro-Gly, and Val-Ser-Val are the major components of fish proteins, and AHTP hit numbers in various species demonstrated a similar distributing pattern. Interestingly, Atlantic salmon (*Salmo salar*) is in possession of far more abundant AHTPs compared with other fish species. In addition, collagen subunit protein is the largest group with more matching AHTPs. Further exploration of two collagen subunits (col4a5 and col8a1) in more fish species suggested that the hit pattern of these conserved proteins among teleost is almost the same, and their phylogeny is consistent with the evolution of these fish species. In summary, our present study provides basic information for the relationship of AHTPs with fish proteins, which sheds light on rapid discovery of marine drugs or food additives from fish protein hydrolysates to alleviate hypertension.

Keywords: antihypertensive peptide; fish protein hydrolysate; high throughput identification; hypertension; collagen

1. Introduction

The prevalence of cardiovascular diseases is mainly caused by high blood pressure [1,2]. Clinical trials have observed that lowering systolic blood pressure would reduce cardiovascular morbidity and mortality significantly [3]. However, serious side effects, such as kidney function deterioration and electrolyte abnormalities, occurred with utilization of various antihypertensive medication [3,4]. Lisinopril, captopril, losartan, azilsartan, atenolol and hydralazine are the common drugs for hypertension treatment [3]. Among them, angiotensin converting enzyme (ACE) inhibitors are preferred antihypertensive drugs [5], as ACE is responsible for the conversion of angiotensin I (Ang I) into vasoconstrictor angiotensin II (Ang II) and the inactivation of vasodilator bradykinin [6,7]. Given the adverse effects of synthetic drugs, there is a growing demand for antihypertensive agents from natural food resources over the past decades since the discovery of natural antihypertensive peptides (AHTPs) from a snake venom [8,9].

AHTPs, a group of small bioactive peptides affecting the workflow of renin-angiotensin system (RAS), commonly are enzyme inhibitors (such as renin and ACE inhibitors) and angiotensin receptor antagonists [6,7]. The most effective and popularly studied molecule is ACE inhibitory peptide [10]. Food-derived AHTPs have been isolated and further validated in vitro and in vivo from various repositories of plants and animals, including milk, soybean, chicken and fish protein hydrolysates [11–13]. Commercial products from milk and fish hydrolysates have been popular in markets [14,15]. VPP and IPP are well-known ACE inhibitory lactotriptides for functional foods, although their effects cannot be comparable to drugs while no adverse effects happened [12].

These natural food proteins provide thousands of templates to develop antihypertensive drugs and functional food supplements [16,17]. The public Database of Antihypertensive Peptides (AHTPDB, <http://crdd.osdd.net/raghava/ahtpdb/>) had collected large amounts of AHTPs by 2015 for better exploitation of antihypertensive agents [18]. Among these, fish hydrolysates have aroused enormous interests for their inexhaustible feature and high nutritional value [19,20]. Bioactive peptides with antimicrobial, antioxidant, anti-obesity, antihypertensive and cryoprotective effects from fish proteins in muscle, skin and even waste products have been reported [19,21–23]. Characterization of AHTPs in fish protein hydrolysates have been conducted in various species, such as salmon, tuna, grass carp and dace [24–27].

Protease hydrolysate components of fish, especially from fish waste products, have been studied for ages. Related engineering reports dedicating to purification techniques, enzymes, crude materials and efficiency assessment of mature products are increasing rapidly [28]. However, the overall knowledge about fish protein content from perspective of big data, and what sort of proteins are potential candidates with the most abundant AHTPs, have not been well examined yet. To figure these out, we performed comparative analyses in this present study for high throughput identification of AHTPs by in situ mapping of whole proteomes from 18 fish species, ranging from ancient coelacanth, spotted gar to teleost fish in Perciformes. We found that Atlantic salmon has more AHTP hits than other fish species. We also identified that collagen subunit protein is in possession of the most plentiful AHTPs. Anyway, our current work initiated a comprehensive investigation into potent AHTPs from protein content in various fishes using big data, which will present instructions for alternatives of raw materials in production of AHTPs.

2. Results

2.1. Construction of a Local AHTP Database

We established a local AHTP database (Table S1) with 963 records, which was created manually from the public AHTPDB and previously published reports. We employed the same unique AHTPDB ID, except for 42 entries (larger than 6979 in ID number) that we appended. Detailed information of sequence, length, molecular weight, half maximal inhibitory concentration (IC_{50}), isolating source, animal model for in vivo test, purification method, IC_{50} determination assay, isoelectric point (PI) and decrease in systolic blood pressure is summarized in Table S1. Interestingly, nearly 20% AHTPs are tripeptides, and 85% AHTPs are between two and eight amino acids (aa) in length (Figure 1). IC_{50} ranges from 0.01 to 12,000 μ M, while half of them are less than 50 μ M. Among them, 148 peptides are with in vivo tests on animal model of spontaneously hypertensive rats, in which 118 entries generated decrease in systolic blood pressure.

2.2. Identification of AHTPs with High Activity

To identify AHTPs more explicitly, we picked 50 AHTPs with the highest activity (see more details in Table S1) to search against the protein datasets of 18 fish species (find corresponding names in Table 1 and Table S2), including coelacanth (*Latimeria chalumnae*) and 17 teleost species. The species selection obviously tends to elucidate AHTPs down the evolutionary timeline. The specific alignment results were provided in Table S3, with a clear exhibition of the locations for each matched AHTP on its corresponding protein.

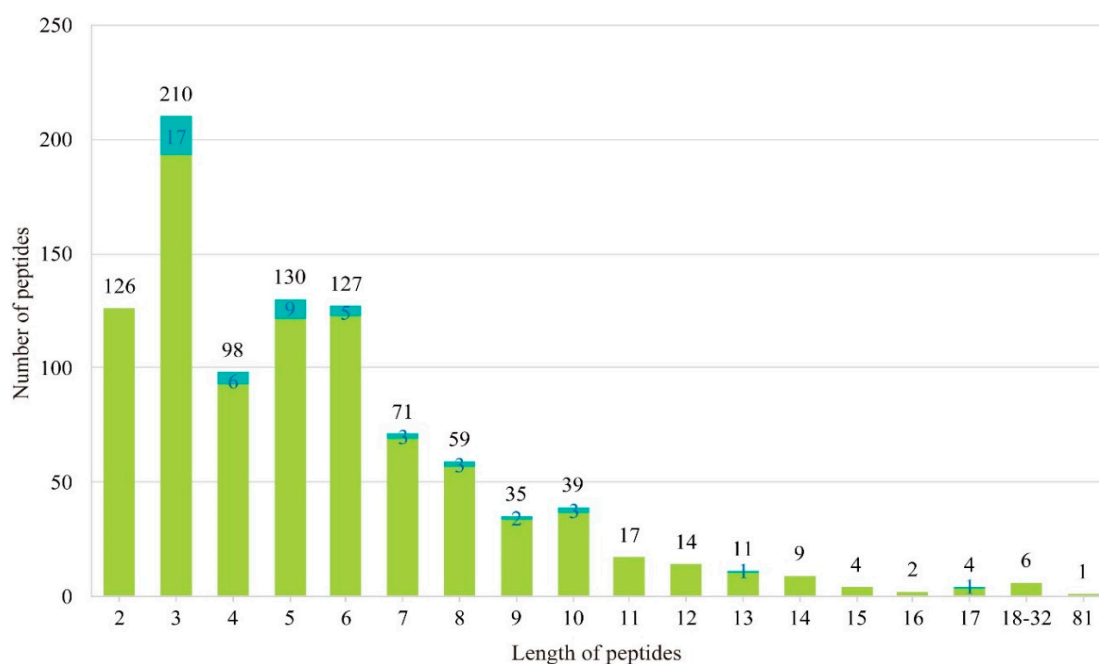


Figure 1. Overview of peptide length vs. numbers in our local AHTP database. Blue boxes within some groups denote the numbers of peptides that we generated in this study, which potentially have high activities for antihypertension.

However, 15 AHTPs did not match any protein, and among them 13 peptides are longer than 6 aa and 2 peptides are hexapeptides. In the 35 mapped AHTPs (Figure 2a), most tripeptides (such as GLP, LGP, LPG and VSV) have numerous hits in the 18 examined fish species (Figure 2b). Tetrapeptide, pentapeptide and hexapeptide were much less likely to match a protein. Interestingly, it seems that the hit number was also influenced by the aa composition of AHTPs. For example, the difference among IRP (Ile-Arg-Pro), IRY (Ile-Arg-Tyr) and IRW (Ile-Arg-Trp) is the last C-terminal aa; however, their hit numbers varied greatly (Figure 2a and Table S4).

Values of Pearson correlation coefficient (Pearson's r ; relationship between total protein number and mapped protein number) of AHTP hit numbers in each type (Table S4) between spotted gar (*Lepisosteus oculatus*) and other 17 fish species, as well as between caved Mexican tetra (*Astyanax mexicanus*) and other 17 fish species, were perfectly correlated ($r > 0.98$). That is to say, when the hit number of an AHTP in one species is greater than another species, the hit number of other AHTPs would also be greater in general. The variance of each mapped AHTP and correlations of AHTP hit number between fishes suggest that different types of AHTPs have variable tendency to occur in organisms, while a similar AHTP hit pattern was conserved among the 18 examined fish species.

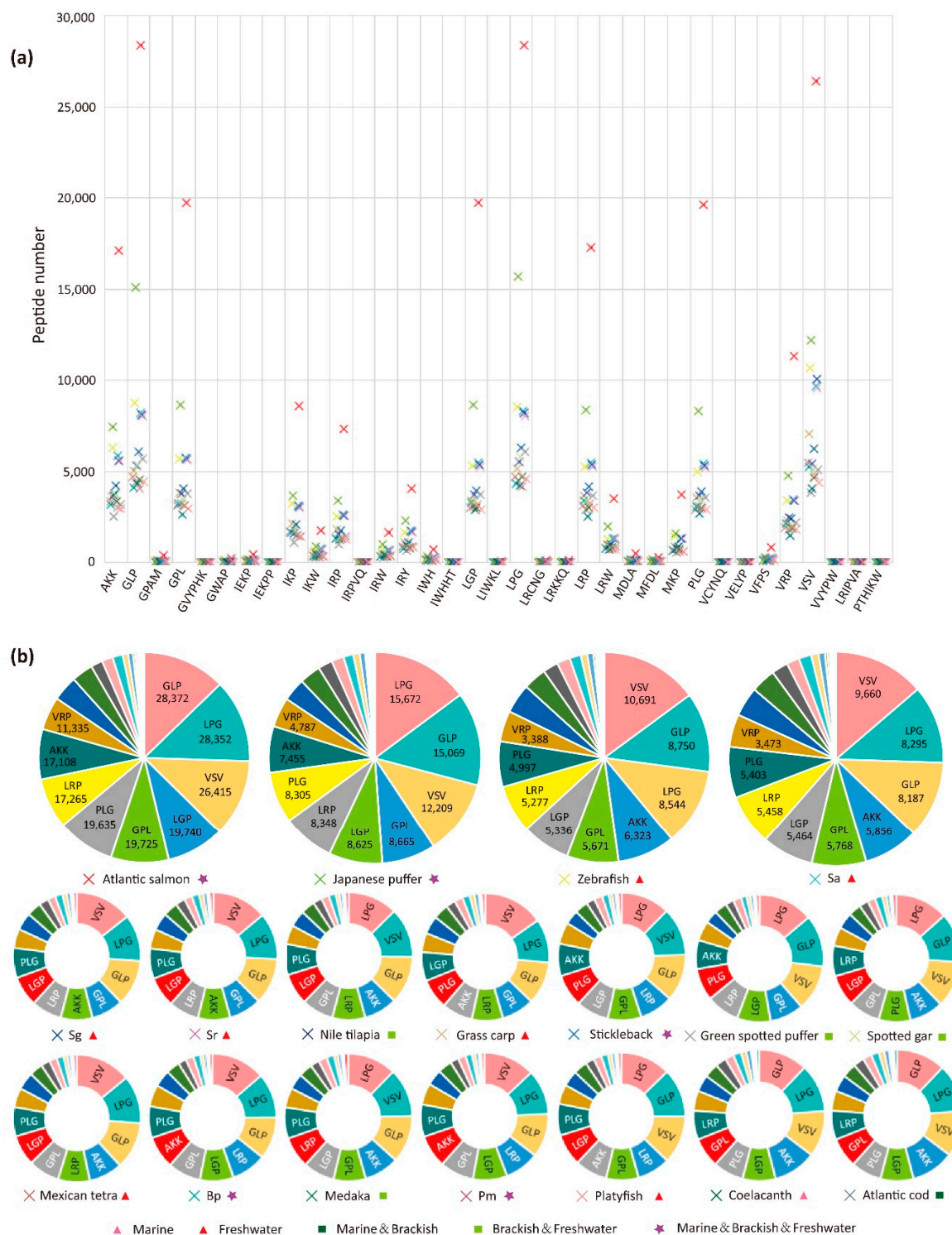


Figure 2. Statistics of the mapped AHTPs. **(a)** AHTP hit number in each species. Crosses with the same color represent the same species. **(b)** Composition of the mapped AHTPs in each protein dataset from the 18 examined fish species. Arrangement of the pie charts is based on the total hit number (from the greatest to the least) in one species. Abbreviations of fish species: Sa, *Sinocyclocheilus anshuiensis*; Sg, *S. grahami*; Sr, *S. rhinocerosus*; Bp, *Boleophthalmus pectinirostris* (Blue-spotted mudskipper); Pm, *Periophthalmus magnuspinnatus* (Giant-fin mudskipper).

In all datasets, Atlantic cod (*Gadus morhua*) has fewest AHTP hits while Atlantic salmon (*Salmo salar*) is the most abundant in every type of AHTPs (Figures 2 and 3a). Interestingly, the Top 11 AHTP

categories (marked in Figure 2b) are the same in all 18 fish species. All these massive peptides scattered around a considerable amount of proteins in each species, with a mapping rate ranging from 0.59 to 0.74 (Figure 3b). Pearson’s r values of the 18 fish species indicated that the more total protein a species possesses, the more protein would be mapped ($r = 0.995$). It was the same between total protein number and total hit number ($r = 0.981$) in all species. Average AHTP hit number in each protein of the 18 fish species ranges from 2.3 to 3.3 (Figure 3c; with more details in Table S4). However, there are hundreds of outliers beyond the mean level (Table S5), accounting for almost a third of the mapped proteins.

Overall, the potential AHTPs produced by proteins of Atlantic salmon would be more abundant than any other fish, suggesting its advantage as a healthy food. In addition, based on the classification of water ecotypes for residence of the 18 fish species, we found that it was not related to the total amount of matched AHTPs (Figure 3).

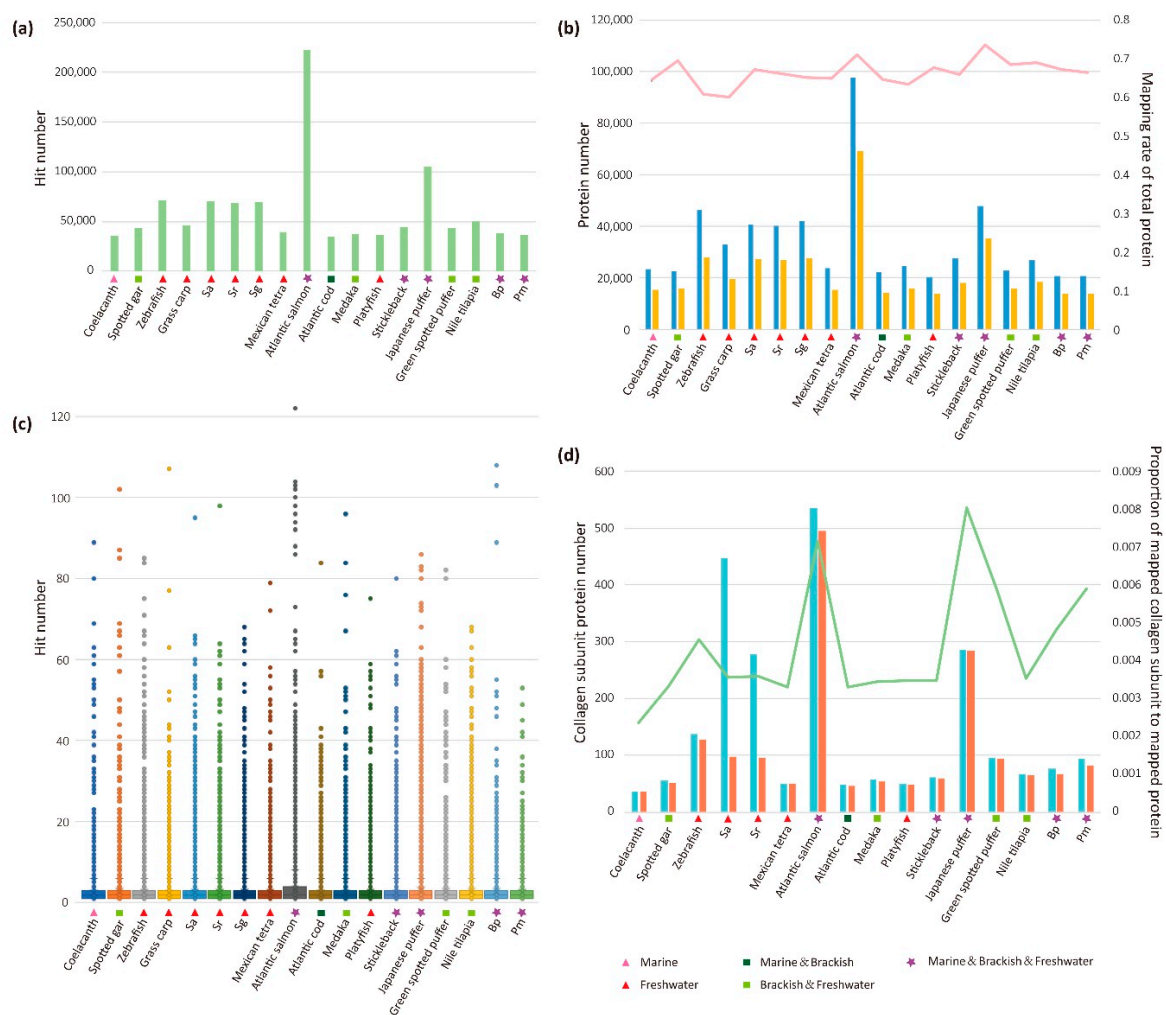


Figure 3. Statistics of AHTP hit numbers and mapped proteins. (a) Total AHTP hit numbers from the 18 examined fish species. (b) Summary of the total and mapped protein numbers, and corresponding mapping rates (the pink line). Blue boxes denote the total proteins while yellow boxes indicate the mapped proteins. (c) Statistics of AHTP hit number of each protein in the 18 fish species. Box plots indicate the median (bold line), the 25th and 75th percentages (box edges), the range (whiskers), and the outliers (dots). (d) Summary of the numbers for total and mapped collagen subunit proteins, and their ratio in mapped proteins (the green line). Light blue boxes denote the total collagen subunit proteins while orange boxes indicate the mapped collagen subunit proteins. The order was arranged in accordance with the phylogenetic relationship of the 18 examined fish species.

Table 1. Categories of proteins with the first top hit and score in 18 examined fish species.

Species	Common Name	The Top Hit	Hit Number	The Top Score	Hit Number
<i>Latimeria chalumnae</i>	Coelacanth	Col4a5	89	Col8a1	39
<i>Lepisosteus oculatus</i>	Spotted gar	Col4a5	102	Col8a1b	57
<i>Danio rerio</i>	Zebrafish	Elastin b	85	Birc2	1
<i>Ctenopharyngodon idella</i>	Grass carp	Titin	107	Col8a1	52
<i>Sinocyclocheilus anshuiensis</i>	Sa	Titin	95	Unknown ¹	9
<i>Sinocyclocheilus rhinoceros</i>	Sr	Titin	98	Col8a1	61
<i>Sinocyclocheilus grahami</i>	Sg	Col4a6	68	Col8a1	59
<i>Astyanax mexicanus</i>	Mexican tetra	Col4a5	79	Col8a2	46
<i>Salmo salar</i>	Atlantic salmon	Xylanase	122	Xylanase	122
<i>Gadus morhua</i>	Atlantic cod	Titin	84	Col8a1	29
<i>Oryzias latipes</i>	Medaka	Col4a6	68	Col25a1	24
<i>Xiphophorus maculatus</i>	Platyfish	Titin	75	Col8a1a	53
<i>Gasterosteus aculeatus</i>	Stickleback	Titin	80	Col8a1	47
<i>Takifugu rubripes</i>	Japanese puffer	Col4a5	86	Col8a1a	42
<i>Tetraodon nigroviridis</i>	Green spotted puffer	Col4a5	82	Col14a1	5
<i>Oreochromis niloticus</i>	Nile tilapia	Col4a6	68	Col25a1	33
<i>Boleophthalmus pectinirostris</i>	Bp	Titin	108	Epsin	89
<i>Periophthalmus magnuspinnatus</i>	Pm	Col4a5	53	Unknown ²	5

¹ Protein with 74 aa. ² Protein with 71 aa.

2.3. Analysis of Mapped Proteins

Since the AHTP hit number of every protein in each fish species (Figure 3c and Figure S5) supported a variety of matched AHTPs for the numerous mapped proteins, we summarized the categories of 50 proteins with top hits and top scores for reference (Figure 4 and Figure S6). As a large amount of sequences are unannotated in the protein datasets of golden-line nabeel fish Sg (*Sinocyclocheilus grahami*) and grass carp (*Ctenopharyngodon idella*), statistics was subsequently applied to the 16 other fish species. Surprisingly, among the 50 proteins with the most hits, collagen subunit proteins accounted for the largest part in general (Figure 4). Another protein, titin, prevailing in most species especially in Atlantic salmon, is the largest giant protein known so far. Taking the length discrepancy into account, categories of the Top 50 proteins with the highest scoring values were more diverse. The proportion of collagen subunit proteins descended in most species, and titin disappeared while proteins of histone family were identified in some fishes. It seems that titin has a great deal with matched AHTPs mainly due to its long length.

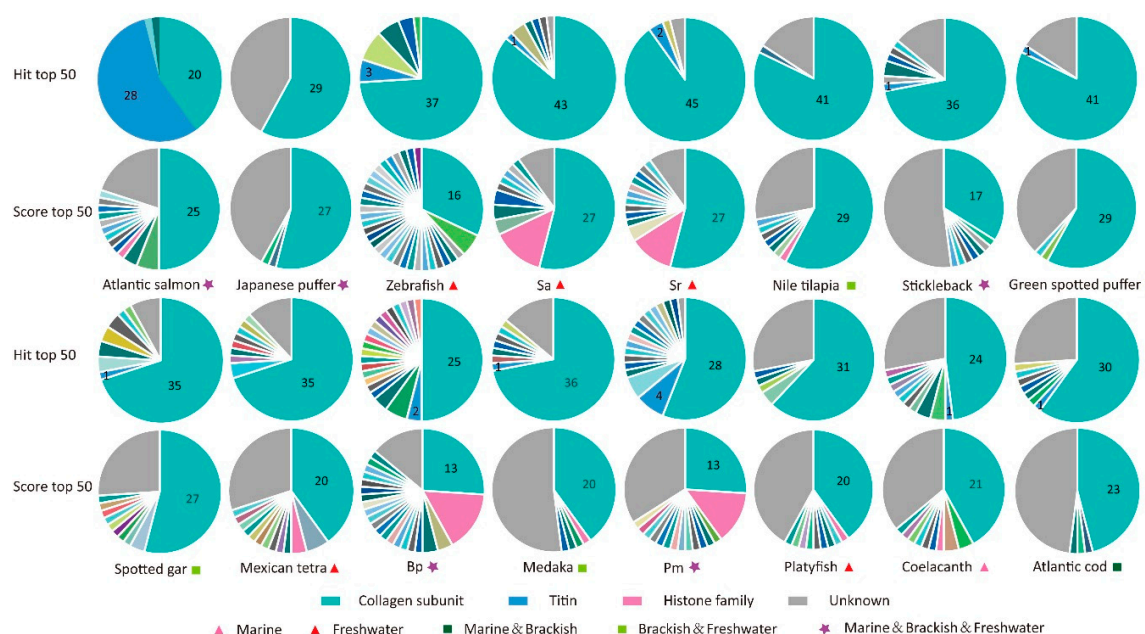


Figure 4. Statistics of protein category for the Top 50 hit and the Top 50 score in 16 fish species.

Categories of proteins with the first top hit and top score from the 18 examined fish species are listed in Table 1. Collagen type IV alpha 5 (col4a5) subunit and collagen type VIII alpha 1 (col8a1) subunit were commonly identified in most species. Meanwhile, in proteins ranked in the Top 50 of score values (Table S6), type IV, VIII and IX collagen subunits were the most extensive types. It is notable that none are fibril forming collagen. However, ancient coelacanth has a limited number of collagen subunit proteins (Figure 3d). Teleost fishes seem to evolve more members of collagen gene family (such as *col25a1*, *col8a1* and *col9a1*) and generate various transcripts, especially in the tetraploid Atlantic salmon and Atlantic cod. Similar observation exists in the tetraploid *Sinocyclocheilus* fishes (Sg, Sr and Sa) and Tetraodontiformes. These splice transcripts and isoforms (Figures 3d and 4) would contribute to the generation of tremendous AHTPs under the hydrolysis with various hydrolases.

2.4. Deep Investigations into Two Collagen Subunits

To learn more about the conserved collagen proteins, we collected homologous sequences of col4a5 and col8a1 (approximately 1685 and 740 aa, respectively) from more fish species (see more details in Tables S7–S9) for deep analysis. They are both structural component of basement membranes in extracellular matrix [29,30]. Phylogenetic analysis and localization of AHTPs in these two macromolecules were realized (Figures 5 and 6). Interestingly, the phylogenetic analysis revealed that the evolutionary relationship of col4a5 and col8a1 coincided with the phylogeny topology of the selected fish species, when spotted gar was used as the outgroup (left in Figures 5 and 6). The tree topology showed that if there is a closer relationship between certain two fish species, their sequences would have a more similar AHTP mapping pattern (right in Figures 5 and 6).

There are multiple GLPG repeating motifs in the collagen sequences, resulting in constant occurrence of triple AHTPs of GLP and LPG (Tables S8 and S9 and Figures 5 and 6). The hits in C-terminal of col4a5 sequences, showing collinear alignment among 24 species in the positions around 1550–1560 (right of Figure 5), are IKP except one hit for IRP in Nile tilapia (*Oreochromis niloticus*). It was also revealed that col4a5 is more conserved than col8a1, as the latter has collinear alignment only among closely related species (right of Figure 6). In addition, all matched AHTPs in the col8a1 sequences are composed of three amino acids of G, L and P.

In the phylogeny of col4a5 (left of Figure 5), four major groups were obviously distinguished. Mainly clustered by genetic distance, the examined species included six fish species in Ostariophysi, salmonids in Protacanthopterygii, Beloniformes and Cyprinodontiformes in Acanthopterygii, as well as other fishes in Acanthopterygii (Table S8). However, in the phylogeny of col8a1, members in Acanthopterygii were converged into a larger group (Table S9). Atlantic cod, the species in Paracanthopterygii for separating Acanthopterygii from Protacanthopterygii perfectly, was a sister group to them. We then observed three major groups (left of Figure 6).

Interestingly, the branch lengths of platyfish (*Xiphophorus maculatus*) and medaka (*Oryzias latipes*, *Oryzias melastigma*) in col4a5 are longer than those of other teleost, while salmonids (Atlantic salmon, Rainbow trout: *Oncorhynchus mykiss*) have shorter branch lengths (Figure 5), suggesting their protein evolutionary rates are distinguished for some reasons. Meanwhile, it seems that there is no correlation between the aa proportion of matched AHTPs and total hit numbers in col4a5 and col8a1 among different water ecotypes (Figures 5 and 6) across the total examined 33 fish species (Table S7), suggesting a more complex collagen evolutionary pattern in the numerous fish species.

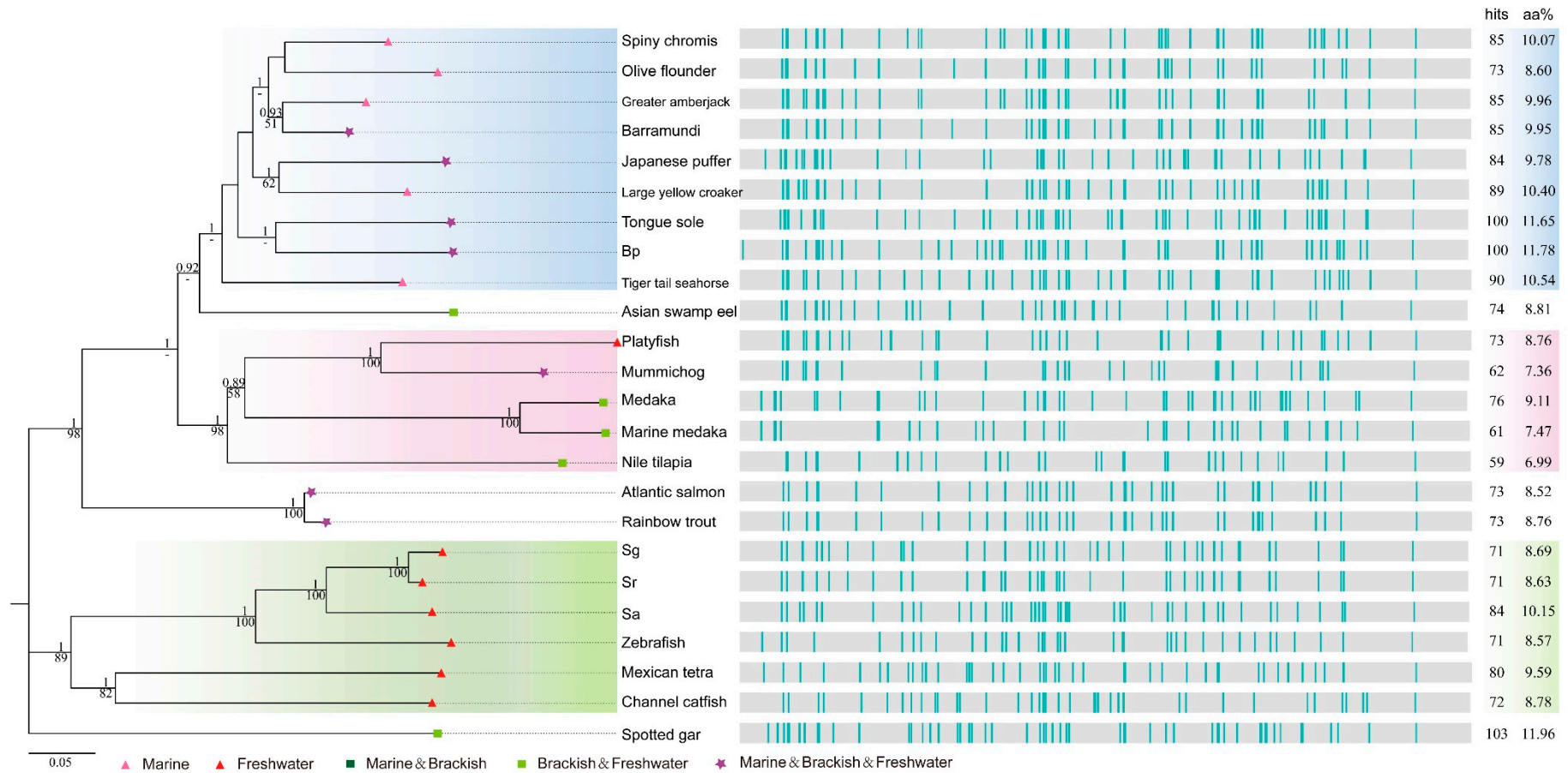


Figure 5. Phylogenetic analysis and AHTP localization of col4a5 in 24 fish species. Corresponding accession number of each protein sequence can be found in Table S7 and detailed AHTP localizations in Table S8.

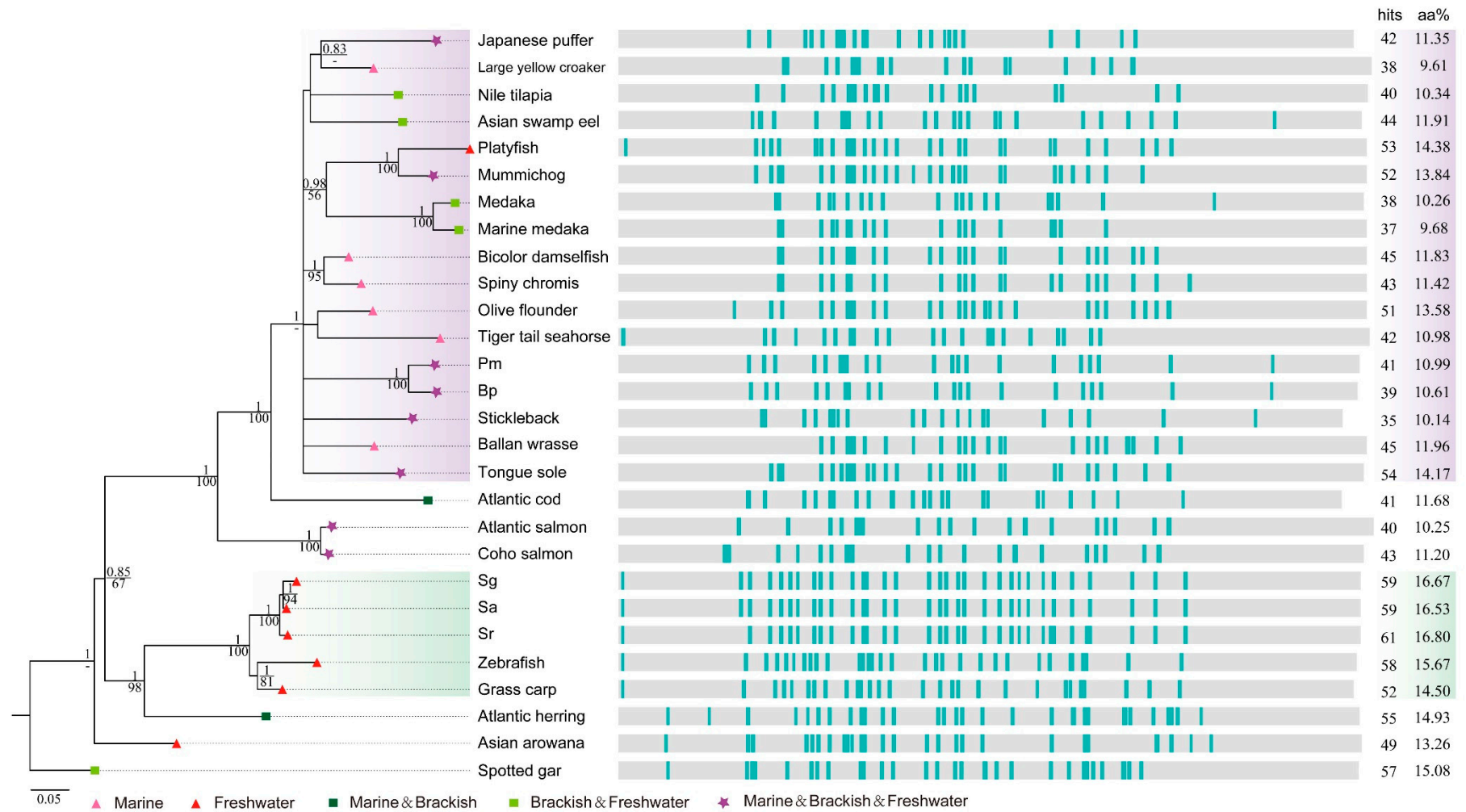


Figure 6. Phylogenetic analysis and AHTPs localization of col8a1 in 28 fish species. Corresponding accession number of each protein sequence can be found in Table S7 and detailed AHTP localizations in Table S9.

3. Discussion

3.1. High-Throughput Identification of AHTPs

In this study, we identified myriad proteins (accounting for nearly 60% of each whole proteome from the examined 18 fish species) matching 50 AHTPs with the highest antihypertensive activity. In this case, with more AHTPs utilized, we would be able to map more proteins. In reverse, the more proteins a species possesses, the more AHTPs could be matched as indicated by our results.

If all AHTPs were used to identify potential peptides for antihypertension in fish proteins, it would be tremendous amount of mapping results; hence, we had to choose Top 50 AHTPs to conduct a representative analysis. This approach relies heavily on previously reported active AHTPs, thus the quality of established AHTP database is crucial for identification of novel bioactive AHTPs [31]. To facilitate data reutilization [13], we recommend unification of the concentration unit for enzyme inhibition assays, and establishment of a standard procedure to construct an updated real-time database.

In a number of studies, BIOPEP database [32] has been regarded as a comprehensive tool to identify peptides with various activities. Salmon col8a1 protein displayed different activities (as shown in Table S10), which were calculated by this *in silico* analysis. Interestingly, 118 peptides with ACE inhibitor activity were scored by PeptideRanker [33], but the majority are dipeptides and some peptides employed by our present study were in the front rank (Table S10). However, it is unclear whether they could be produced under physiological conditions. ExPASy PeptideCutter was employed to simulate digestive reactions of col8a1 by pepsin and trypsin, with 214 cleavages sites generated from all possible fragments (Table S10). These data indicate that the combination of various enzymes increases the number of peptides, although not all predicted AHTPs could be produced or some peptides occurred with unknown activities [34,35]. For example, AHTPs of GLP, GPL and LPG are always split by pepsin, while PGF, GDK, PGK and PQY, whose activities have not been proved yet, appeared as potential candidates according to the prediction of AHTpin [36] (<http://crdd.osdd.net/raghava/ahtpin/>). In fact, many platforms could identify AHTPs in protein hydrolysates *in silico*, demonstrating the advantages of bioinformatics for instructing us to perform subsequent physiological tests [34,37]. The actual situation, nonetheless, has numerous complex tasks to complete. The exploitation of novel enzymes is one of the keys to successful releasing of AHTPs from proteins. Except for pepsin, trypsin, alcalase, flavourzyme, neutrase and protamex, other novel proteinases are also waiting for exploration [38] to generate bioactive peptides more efficiently and productively.

Meanwhile, antihypertensive activity of many peptides *in vivo* is largely unknown. Among our 35 mapped AHTPs, only seven peptides were tested *in vivo* (Table S1): IWH, IKW, MKP, IWHHT, LIWKL, VELYP and IKP. Generally speaking, they are not rich in most species except the IKP. The top hit peptides usually have no data from *in vivo* test yet. It is worth investigating the relationship between their antihypertensive activity and their amount from a whole proteome view. Therefore, further bioactive validation is indispensable to the understanding of AHTPs. What is more, the composition of peptides is a crucial factor to affect the activity of AHTPs. Composition analysis of AHTPs have revealed that Gly (G) and Pro (P) are prevalent in dipeptides and tripeptides [36]. Quantitative structural activity relationship (QSAR) is a powerful tool to predict their activity [39]. Molecular docking and other prediction model are also dedicated to find novel AHTPs and further provide guidelines to subsequent preliminary tests [37,40]. It was also reported that peptides rich in Arg (R), Val (V) and Leu (L) with lower molecular weight exhibited a higher antihypertensive activity [41]. Based on these characteristics, we should make a restricted choice to select crude materials for AHTPs. Thereby, related exploration into functional supplements and drugs originated from fish still have a long way to go.

3.2. Potential Utility of Collagens

Compared with other proteins, the collagen family has a distinct characteristic of triple-helical Gly-X-Y structure [30]. This gives rise to the repeat occurrence of AHTPs of GLP and LPG. Furthermore, their bioactivities demonstrate their remarkable advantages in various fields [42,43]. For example, Vastatin, an endogenous antiangiogenic polypeptide from non-collagenous C-terminal globular domain of col8a1, has been reported to be promising in fighting against glioblastoma [44]. Elevated expression of type VIII collagen in the serum of patients with various cancers suggested its relationship with angiogenesis [45]. In addition, type IX collagen was required for the formation of vascular plexus [46], and the enzymatic fragments of salmon col7a1 isoform X5 also showed ACE inhibitory activity [37]. Previous studies also reported that collagen subunits contain multiple host defense peptides [47]. What is more, some collagen-derived peptides have antioxidant and anti-inflammatory capacities [48], and even can suppress allergic responses [49]. In this present study, multiple collagen subunits aroused our attention since their mapping ratios are much higher compared with other proteins. Titin, troponin, actin and myosin are frequently employed for generation of AHTPs [12]. Nevertheless, in the mapping results of myosin, there are only 1–12 AHTP hits. In summary, the bioactivity of fish collagens is highly interesting, especially for their wide applications in the cosmetic industry [50].

Scale and skin of fish contain more abundant collagens than muscle, thus the choice of raw materials for collagen preparation is also a challenge [51]. In addition, it is noticeable that the evolutionary rate of platyfish is much higher than any other examined fish species in our phylogenetic analysis (Figures 5 and 6). In our attempt to explore the relationships between water ecology and potential AHTP hit number or mapping ratio of collagens, we could not reach a consistent conclusion. The species with similar habitat niches did not exhibit similar mapping pattern in col4a5 and col8a1, although the closely-related fish species did. Obviously, genetic basis, functional analysis and molecular mechanisms of collagens are worth deep investigation.

3.3. Prospective Development of Fish-Derived Antihypertensive Agents

Fishes have always been regarded as omnipotent food resources [15], but it is hard to evaluate the overall nutritional value discrepancy among various fish species. Previous studies reported that marine fish supply proteins equivalent to freshwater fish but with lower fat [52]. In this study, salmon proteins were revealed to have the most abundant AHTPs. From this view, salmons may be a good supplement for treatment of patients with cardiovascular diseases.

There are many reports with identification of bioactive peptides from salmon proteins, especially from those trimmings [24,28]. Tuna dark muscle have also been suggested as a potential functional food or pharmaceutical for AHTPs with antihypertensive effects [26]. Nevertheless, big-data-based analysis of the whole proteomes from various fish species are rare. This present study supports that fish proteins are good raw materials for further development of antihypertensive agents. Meanwhile, processing methods of fish protein hydrolysates are also ongoing engineering problems for commercial applications. The bioactivity of fish proteins after treatment and manufacture should be considered seriously [25,53].

4. Materials and Methods

4.1. Data Collection

AHTPs were originally downloaded from the public AHTPDB, which contained 6979 entries (on 28 June 2018). We corrected some confusing records based on published references and deleted those items that had not been well defined. Most IC₅₀ values of the AHTPs were confirmed further manually, with retaining of those in the unit of μM . Those duplicated peptides were deleted and the entry for each AHTP was retained with the lowest IC₅₀ value. Another 42 sequences were obtained from public literatures of reviews and articles.

Finally, we compiled 963 peptides in our local database (Table S1). To identify potential AHTPs in fish protein datasets more explicitly and efficiently, we selected the Top 50 AHTPs with the highest inhibitory activity (marked in Table S1), without dipeptides as their high likelihood of random matches, to search against protein datasets. High-quality whole proteome sequences of 17 fish species were downloaded from National Center for Biotechnology Information (NCBI) and Ensemble database, while the grass carp protein dataset was acquired from an exclusively database for this fish in National Center for Gene Research under Chinese Academy of Sciences (Table S2).

4.2. Screening for AHTPs

To obtain the target proteins and extract their corresponding matched AHTPs with localization information, we developed a local custom Perl script pipeline to analyze the data. Considering the length discrepancy among sequences, a scoring rule represented by an absolute value was applied to sum the proportion of various AHTPs mapped to a protein. Statistics of total hit number in each mapped protein of the 18 examined fish species and hit number of matched AHTPs in each species were implemented. According to the protein coding gene symbol of each protein sequence, categories of the Top 50 proteins with the most hit number (Hit Top 50) and with the largest scoring value (Score Top 50) were summarized and classified. Collagen subunit proteins were characterized by index terms. Gene symbols of collagens include collagen, persistent plexus, si:dkey-21p1.3, si:dkey-61l1.4, zmp:0000000760, si:dkeyp-44a8.4, si:ch211-196i2.1, and si:dkey-225n22.4.

4.3. Construction of Phylogenetic Trees

The collagen subunit protein sequences were collected from homology searches against collected protein datasets and hits of NCBI Blast results. All sequences were further checked in the NCBI and Ensemble database. Accession numbers were provided in Table S7. Those proteins with low-quality were discarded. Multiple sequences of col4a5 and col8a1 were aligned by MAFFT (version 7.037) [54]. Conserved domains were extracted by Gblocks (version 0.91b) [55]. Likelihood-based methods were applied to establish the phylogenetic trees by PhyML (version 3.1) [56] and MrBayes (version 3.2.6) [57]. Models of amino acid replacement were determined by Prottest (version 3.4.2) [58]. PhyML analyses were executed under the MtREV+I+G+F (col4a5) and JTT+I+G+F (col8a1) model with bootstrap value of 1000. Markov chains were run for 1,000,000 generations in Bayesian Inference, sampling every 100 generations. The first 2500 convergence runs were discarded as burn-in. Chain stationary and run parameter convergence were further checked using Tracer (version 1.7.1) [59]. Phylogenetic trees were visualized by FigTree (version 1.4.3) [60]. Local visualization of peptides to protein sequences was realized by Perl package "SVG" [61].

5. Conclusions

This is the first attempt to develop AHTPs from fish proteins in a high-throughput way. In this study, identification of AHTPs with the highest inhibitory activity from whole proteome datasets of 18 fish species was realized. In addition, the distribution of various AHTPs in protein sequences and protein categories with the most hit number were analyzed. Salmons seem to be a good food resource for people to treat hypertension or improve health, and collagen subunit proteins col4a5 and col8a1 are potent antihypertensive agent substitutes for existence of numerous AHTPs in their sequences. In summary, this is a comprehensive work for discovery of potent AHTPs and their fish proteins/resources at a whole proteome level.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1660-3397/16/10/365/s1>. Table S1: Established local AHTP database. Table S2: Websites for downloading the protein datasets of 18 fish species. Table S3: Specific alignment results of all mapped proteins and their corresponding matched AHTPs in the 18 examined fish species. Table S4: Statistics of AHTP hit number in each species and comprehensive alignment results in the 18 examined fish species. Table S5: Statistics of hit numbers in each mapped protein from the 18 examined fish species. Table S6: Mapped proteins with the Top 50 scores in the 18 examined fish species. Table S7: Accession numbers of collagen subunit protein sequences and water ecotypes of 33 fish species. Table S8: In situ mapping result of col4a5 from 24 fish species with matched AHTPs. Table S9: In situ mapping result of col8a1 from 28 fish species with matched AHTPs. Table S10: In silico analysis of salmon col8a1.

Author Contributions: J.Y. and Q.S. conceived and designed the experiments. Y.Y., Y.L. and L.Z. analyzed the data. Y.Y. wrote the manuscript. Q.S. and J.Y. revised the manuscript.

Funding: This research was supported by the Science and Technology Project of Shenzhen (No. JCYJ20170413155047512), Shenzhen High Technology Research and Development Program (No. JSGG20150916140818695), and Yunnan Innovation and Enhancement Program of Provincial Science and Technology Department (No. 2016AB024).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lurie, Y. Hypertension is the biggest cost center being faced by healthcare today. *Hypertension* **2018**.
2. Chockalingam, A.; Campbell, N.R.; Fodor, J.G. Worldwide epidemic of hypertension. *Can. J. Cardiol.* **2006**, *22*, 553–555. [[CrossRef](#)]
3. SPRINT Research Group. A randomized trial of intensive versus standard blood-pressure control. *N. Engl. J. Med.* **2015**, *373*, 2103–2116. [[CrossRef](#)] [[PubMed](#)]
4. The SPRINT Study Research Group. The design and rationale of a multi-center clinical trial comparing two strategies for control of systolic blood pressure: The Systolic Blood Pressure Intervention Trial (SPRINT). *Clin. Trials Lond. Engl.* **2014**, *11*, 532–546. [[CrossRef](#)] [[PubMed](#)]
5. Kalaitzidis, R.G.; Elisaf, M.S. Treatment of hypertension in chronic kidney disease. *Curr. Hypertens. Rep.* **2018**, *20*, 64. [[CrossRef](#)] [[PubMed](#)]
6. George, A.J.; Thomas, W.G.; Hannan, R.D. The renin–angiotensin system and cancer: Old dog, new tricks. *Nat. Rev. Cancer* **2010**, *10*, 745–759. [[CrossRef](#)] [[PubMed](#)]
7. Wu, J.; Liao, W.; Udenigwe, C.C. Revisiting the mechanisms of ACE inhibitory peptides from food proteins. *Food Bioact. Evid. Health Benefits Underst. Mech.* **2017**, *69*, 214–219. [[CrossRef](#)]
8. Brook, R.D.; Appel, L.J.; Rubenfire, M.; Ogedegbe, G.; Bisognano, J.D.; Elliott, W.J.; Fuchs, F.D.; Hughes, J.W.; Lackland, D.T.; Staffileno, B.A.; et al. Beyond medications and diet: Alternative approaches to lowering blood pressure. *Hypertension* **2013**, *61*, 1360–1383. [[CrossRef](#)] [[PubMed](#)]
9. Ferreira, S.H.; Bartelt, D.C.; Greene, L.J. Isolation of bradykinin-potentiating peptides from Bothrops jararaca venom. *Biochemistry* **1970**, *9*, 2583–2593. [[CrossRef](#)] [[PubMed](#)]
10. Majumder, K.; Wu, J. Molecular targets of antihypertensive peptides: understanding the mechanisms of action based on the pathophysiology of hypertension. *Int. J. Mol. Sci.* **2015**, *16*, 256–283. [[CrossRef](#)] [[PubMed](#)]
11. FitzGerald, R.J.; Murray, B.A.; Walsh, D.J. Hypotensive peptides from milk proteins. *J. Nutr.* **2004**, *134*, 980–988. [[CrossRef](#)] [[PubMed](#)]
12. Bhat, Z.F.; Kumar, S.; Bhat, H.F. Antihypertensive peptides of animal origin: A review. *Crit. Rev. Food Sci. Nutr.* **2017**, *57*, 566–578. [[CrossRef](#)] [[PubMed](#)]
13. Lee, S.Y.; Hur, S.J. Antihypertensive peptides from animal products, marine organisms, and plants. *Food Chem.* **2017**, *228*, 506–517. [[CrossRef](#)] [[PubMed](#)]
14. Hartmann, R.; Meisel, H. Food-derived peptides with biological activity: From research to food applications. *Plant. Biotechnol. Food Biotechnol.* **2007**, *18*, 163–169. [[CrossRef](#)] [[PubMed](#)]
15. Gevaert, B.; Veryser, L.; Verbeke, F.; Wynendaele, E.; De Spiegeleer, B. Fish hydrolysates: A regulatory perspective of bioactive peptides. *Protein Pept. Lett.* **2016**, *23*, 1052–1060. [[CrossRef](#)] [[PubMed](#)]
16. Daliri, E.B.M.; Oh, D.H.; Lee, B.H. Bioactive peptides. *Foods* **2017**, *6*, 32. [[CrossRef](#)] [[PubMed](#)]
17. Nasri, M. Protein hydrolysates and biopeptides: Production, biological activities, and applications in foods and health Benefits. A review. In *Advances in Food and Nutrition Research*; Toldrá, F., Ed.; Academic Press: Cambridge, MA, USA, 2017; pp. 109–159. ISBN 1043-4526.

18. Kumar, R.; Chaudhary, K.; Sharma, M.; Nagpal, G.; Chauhan, J.S.; Singh, S.; Gautam, A.; Raghava, G.P. AHTPDB: A comprehensive platform for analysis and presentation of antihypertensive peptides. *Nucleic Acids Res.* **2015**, *43*, 956–962. [[CrossRef](#)] [[PubMed](#)]
19. Ryan, J.T.; Ross, R.P.; Bolton, D.; Fitzgerald, G.F.; Stanton, C. Bioactive peptides from muscle sources: Meat and fish. *Nutrients* **2011**, *3*, 765–791. [[CrossRef](#)] [[PubMed](#)]
20. Jensen, I.J.; Mæhre, K.H. Preclinical and clinical studies on antioxidative, antihypertensive and cardioprotective effect of marine proteins and peptides—A review. *Mar. Drugs* **2016**, *14*, 211. [[CrossRef](#)] [[PubMed](#)]
21. Wilson, J.; Hayes, M.; Carney, B. Angiotensin-I-converting enzyme and prolyl endopeptidase inhibitory peptides from natural sources with a focus on marine processing by-products. *Food Chem.* **2011**, *129*, 235–244. [[CrossRef](#)]
22. Kim, M.R.; Kim, J.W.; Park, J.B.; Hong, Y.K.; Ku, S.K.; Choi, J.S. Anti-obesity effects of yellow catfish protein hydrolysate on mice fed a 45% kcal high-fat diet. *Int. J. Mol. Med.* **2017**, *40*, 784–800. [[CrossRef](#)] [[PubMed](#)]
23. Nikoo, M.; Benjakul, S.; Xu, X. Antioxidant and cryoprotective effects of Amur sturgeon skin gelatin hydrolysate in unwashed fish mince. *Food Chem.* **2015**, *181*, 295–303. [[CrossRef](#)] [[PubMed](#)]
24. Neves, A.C.; Harnedy, P.A.; O’Keeffe, M.B.; FitzGerald, R.J. Bioactive peptides from Atlantic salmon (*Salmo salar*) with angiotensin converting enzyme and dipeptidyl peptidase IV inhibitory, and antioxidant activities. *Food Chem.* **2017**, *218*, 396–405. [[CrossRef](#)] [[PubMed](#)]
25. Elavarasan, K.; Shamasundar, B.A.; Badii, F.; Howell, N. Angiotensin I-converting enzyme (ACE) inhibitory activity and structural properties of oven- and freeze-dried protein hydrolysate from fresh water fish (*Cirrhinus mrigala*). *Food Chem.* **2016**, *206*, 210–216. [[CrossRef](#)] [[PubMed](#)]
26. Qian, Z.J.; Je, J.Y.; Kim, S.K. Antihypertensive effect of angiotensin I converting enzyme-inhibitory peptide from hydrolysates of bigeye tuna dark muscle, *Thunnus obesus*. *J. Agric. Food Chem.* **2007**, *55*, 8398–8403. [[CrossRef](#)] [[PubMed](#)]
27. Chen, J.; Wang, Y.; Zhong, Q.; Wu, Y.; Xia, W. Purification and characterization of a novel angiotensin-I converting enzyme (ACE) inhibitory peptide derived from enzymatic hydrolysate of grass carp protein. *Peptides* **2012**, *33*, 52–58. [[CrossRef](#)] [[PubMed](#)]
28. Harnedy, P.A.; FitzGerald, R.J. Bioactive peptides from marine processing waste and shellfish: A review. *J. Funct. Foods* **2012**, *4*, 6–24. [[CrossRef](#)]
29. Shuttleworth, C.A. Type VIII collagen. *Int. J. Biochem. Cell. Biol.* **1997**, *29*, 1145–1148. [[CrossRef](#)]
30. Gordon, M.K.; Hahn, R.A. Collagens. *Cell. Tissue Res.* **2009**, *339*, 247. [[CrossRef](#)] [[PubMed](#)]
31. Win, T.S.; Schaduangrat, N.; Prachayasittikul, V.; Nantasenamat, C.; Shoombuatong, W. PAAP: A web server for predicting antihypertensive activity of peptides. *Future Med. Chem.* **2018**, *10*, 1749–1767. [[CrossRef](#)] [[PubMed](#)]
32. Minkiewicz, P.; Dziuba, J.; Iwaniak, A.; Dziuba, M.; Darewicz, M. BIOPEP database and other programs for processing bioactive peptide sequences. *J. AOAC Int.* **2008**, *91*, 965–980. [[PubMed](#)]
33. Mooney, C.; Haslam, N.J.; Pollastri, G.; Shields, D.C. Towards the improved discovery and design of functional peptides: Common features of diverse classes permit generalized prediction of bioactivity. *PLoS ONE* **2012**, *7*, 45012. [[CrossRef](#)] [[PubMed](#)]
34. Bleakley, S.; Hayes, M.; O’Shea, N.; Gallagher, E.; Lafarga, T. Predicted release and analysis of novel ACE-I, renin, and DPP-IV inhibitory peptides from common oat (*Avena sativa*) protein hydrolysates using in silico analysis. *Foods* **2017**, *6*, 108. [[CrossRef](#)] [[PubMed](#)]
35. Lafarga, T.; O’Connor, P.; Hayes, M. Identification of novel dipeptidyl peptidase-IV and angiotensin-I-converting enzyme inhibitory peptides from meat proteins using in silico analysis. *Peptides* **2014**, *59*, 53–62. [[CrossRef](#)] [[PubMed](#)]
36. Kumar, R.; Chaudhary, K.; Singh Chauhan, J.; Nagpal, G.; Kumar, R.; Sharma, M.; Raghava, G.P. An in silico platform for predicting, screening and designing of antihypertensive peptides. *Sci. Rep.* **2015**, *5*, 12512. [[CrossRef](#)] [[PubMed](#)]
37. Zhipeng, Y.; Yang, C.; Wenzhu, Z.; Jianrong, L.; Jingbo, L.; Feng, C. Identification and molecular docking study of novel angiotensin-converting enzyme inhibitory peptides from *Salmo salar* using in silico methods. *J. Sci. Food Agric.* **2018**, *98*, 3907–3914.

38. Ngo, D.H.; Ryu, B.; Kim, S.K. Active peptides from skate (*Okamejei kenojei*) skin gelatin diminish angiotensin-I converting enzyme activity and intracellular free radical-mediated oxidation. *Food Chem.* **2014**, *143*, 246–255. [[CrossRef](#)] [[PubMed](#)]
39. Wu, J.; Aluko, R.E.; Nakai, S. Structural requirements of angiotensin I-converting enzyme inhibitory peptides: quantitative structure–activity relationship study of di- and tripeptides. *J. Agric. Food Chem.* **2006**, *54*, 732–738. [[CrossRef](#)] [[PubMed](#)]
40. Ondetti, M.; Rubin, B.; Cushman, D. Design of specific inhibitors of angiotensin-converting enzyme: New class of orally active antihypertensive agents. *Science* **1977**, *196*, 441–444. [[CrossRef](#)] [[PubMed](#)]
41. Bhat, I.; Karunasagar, I. Antihypertensive activity of fish protein hydrolysates and its peptides. *Crit. Rev. Food Sci. Nutr.* **2018**, 1–12. [[CrossRef](#)] [[PubMed](#)]
42. Felician, F.F.; Xia, C.; Qi, W.; Xu, H. Collagen from marine biological sources and medical applications. *Chem. Biodivers.* **2018**, *15*, e1700557. [[CrossRef](#)] [[PubMed](#)]
43. Boonmaleerat, K.; Wanachewin, O.; Phitak, T.; Pothacharoen, P.; Kongtawelert, P. Fish collagen hydrolysates modulate cartilage metabolism. *Cell. Biochem. Biophys.* **2018**, *76*, 279–292. [[CrossRef](#)] [[PubMed](#)]
44. Li, Y.; Li, J.; Woo, Y.M.; Shen, Z.; Yao, H.; Cai, Y.; Lin, M.C.; Poon, W.S. Enhanced expression of Vastatin inhibits angiogenesis and prolongs survival in murine orthotopic glioblastoma model. *BMC Cancer* **2017**, *17*, 126. [[CrossRef](#)] [[PubMed](#)]
45. Hansen, N.U.B.; Willumsen, N.; Sand, J.M.B.; Larsen, L.; Karsdal, M.A.; Leeming, D.J. Type VIII collagen is elevated in diseases associated with angiogenesis and vascular remodeling. *Clin. Biochem.* **2016**, *49*, 903–908. [[CrossRef](#)] [[PubMed](#)]
46. Huang, C.; Wang, T.C.; Lin, B.H.; Wang, Y.W.; Johnson, S.L.; Yu, J. Collagen IX is required for the integrity of collagen II fibrils and the regulation of vascular plexus formation in Zebrafish caudal fins. *Dev. Biol.* **2009**, *332*, 360–370. [[CrossRef](#)] [[PubMed](#)]
47. Abdillahi, S.M.; Maaß, T.; Kasetty, G.; Strömstedt, A.A.; Baumgarten, M.; Tati, R.; Nordin, S.L.; Walse, B.; Wagener, R.; Schmidtchen, A.; et al. Collagen VI contains multiple host defense peptides with potent in vivo activity. *J. Immunol.* **2018**, *201*, 1007. [[CrossRef](#)] [[PubMed](#)]
48. Chen, Y.P.; Liang, C.H.; Wu, H.T.; Pang, H.Y.; Chen, C.; Wang, G.H.; Chan, L.P. Antioxidant and anti-inflammatory capacities of collagen peptides from milkfish (*Chanos chanos*) scales. *J. Food Sci. Technol.* **2018**, *55*, 2310–2317. [[CrossRef](#)] [[PubMed](#)]
49. Nishikimi, A.; Koyama, Y.; Ishihara, S.; Kobayashi, S.; Tometsuka, C.; Kusubata, M.; Kuwaba, K.; Hayashida, O.; Hattori, S.; Katagiri, K. Collagen-derived peptides modulate CD4(+) T-cell differentiation and suppress allergic responses in mice. *Immun. Inflamm. Dis.* **2018**, *6*, 245–255. [[CrossRef](#)] [[PubMed](#)]
50. Tang, Y.; Jin, S.; Li, X.; Li, X.; Hu, X.; Chen, Y.; Huang, F.; Yang, Z.; Yu, F.; Ding, G. Physicochemical properties and biocompatibility evaluation of collagen from the skin of giant croaker (*Nibea japonica*). *Mar. Drugs* **2018**, *16*, 222. [[CrossRef](#)] [[PubMed](#)]
51. Jie, L.; MingYan, Y.; JunHong, L.; Song, Q. Research progress of fish collagen. *J. Food Saf. Qual.* **2015**, *6*, 3941–3946.
52. Puwastien, P.; Judprasong, K.; Kettwan, E.; Vasanachitt, K.; Nakngamanong, Y.; Bhattacharjee, L. Proximate composition of raw and cooked Thai freshwater and marine fish. *J. Food Compos. Anal.* **1999**, *12*, 9–16. [[CrossRef](#)]
53. Shaviklo, A.R. Development of fish protein powder as an ingredient for food applications: A review. *J. Food Sci. Technol.* **2015**, *52*, 648–661. [[CrossRef](#)] [[PubMed](#)]
54. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
55. Talavera, G.; Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **2007**, *56*, 564–577. [[CrossRef](#)] [[PubMed](#)]
56. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]
57. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [[CrossRef](#)] [[PubMed](#)]

58. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinforma. Oxf. Engl.* **2011**, *27*, 1164–1165. [[CrossRef](#)] [[PubMed](#)]
59. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)] [[PubMed](#)]
60. Molecular Evolution, Phylogenetics and Epidemiology. Available online: <http://tree.bio.ed.ac.uk/software/figtree/> (accessed on 4 October 2016).
61. Wallace, S. *Perl Graphics Programming: Creating SVG, SWF (Flash), JPEG and PNG Files with Perl*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2002; ISBN 1-4493-5830-6.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).