


Article

High-Throughput Identification of Putative Antimicrobial Peptides from Multi-Omics Data of the Lined Seahorse (*Hippocampus erectus*)

Xiyang Chen ^{1,2,†}, Yunhai Yi ^{1,2,†} , Xinxin You ^{1,2} , Jie Liu ³ and Qiong Shi ^{1,2,4,*} 

¹ BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, China; chenxiyang@genomics.cn (X.C.); yiyunhai@genomics.cn (Y.Y.); youxinxin@genomics.cn (X.Y.)

² 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

³ BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China; liujie8@genomics.cn

⁴ Laboratory of Aquatic Genomics, College of Life Sciences and Oceanography, Shenzhen University, Shenzhen 518060, China

* Correspondence: shiqiong@genomics.cn; Tel.: +86-185-6627-9826; Fax: +86-755-3630-7807

† These authors contributed equally to this work.

Received: 25 November 2019; Accepted: 25 December 2019; Published: 29 December 2019



Abstract: Lined seahorse (*Hippocampus erectus*), the most widely cultivated seahorse in China, has been in short supply because of its important medicinal value; meanwhile, unnatural deaths caused by various diseases (especially enteritis) have limited their practical large-scale aquaculture. Antimicrobial peptides (AMPs), as the best alternative to antibiotics, have been extensively applied in agricultural practices. In this study, we identified 290 putative AMP sequences from our previously published genome and transcriptome data of the lined seahorse. Among them, 267 are novel, and 118 were validated by our proteome data generated in the present study. It seems that there is a tissue preference in the distribution of AMP/AMP precursor transcripts, such as lectins in the male pouch. In addition, their transcription levels usually varied during development. Interestingly, the representative lectins kept extremely high levels at the pre-pregnancy stage while at relatively lower levels at other stages. Especially *Lectin25*, with the highest transcription levels and significant developmental changes, has been reported to be involved in seahorse and human pregnancy. The comparison of transcriptome data between one-day and three-month juveniles indicated that *Hemoglobin2* (*Hemo2*) was significantly upregulated in the body, haslet, and brain. Our proteome data of female and male individuals revealed three putative AMP precursors with sexual specificity, including two male-biased cyclin-dependent kinases (CDK-like16 and CDK-like23) and one female-biased bovine pancreatic trypsin inhibitor 2 (BPTI2). In conclusion, our present high-throughput identification of putative AMP sequences from multi-omics (including genomics, transcriptomics, and proteomics) data provides an overview of AMPs in the popular lined seahorse, which lays a solid foundation for further development of AMP-based fish food additives and human drugs.

Keywords: lined seahorse; antimicrobial peptide (AMP); genome; transcriptome; proteome; sexual specificity

1. Introduction

Seahorses, belonging to Syngnathidae, are well known for their special shape, male pregnancy, and pharmaceutical values [1]. Their shapes are usually characterized by an elongated snout, a body covered with hard bones but without scales, and a shortage of pelvic and caudal fins. Their reproductive behaviors are also unique to male pregnancy. In addition, seahorses have been a precious resource for traditional Chinese medicine [2,3]. According to many old records, at least five seahorse species had been applied in the production of these medicines; especially, the lined seahorse (*Hippocampus erectus*), introduced mainly from America, has been the most widely cultivated species in China. Recent studies have shown that extracts and isolates from seahorses have a variety of pharmaceutical activities, including antioxidant [4], immune-enhancing [5], sex hormone-like [6], anti-tumor [7], and anti-inflammatory [8] functions.

Seahorses have been in great demand due to their medicinal values, although they are mainly supplied by field fishing. Nevertheless, large-scale aquaculture of seahorses has been seriously limited by fatal diseases, which are usually caused by exogenous pathogens, such as vibrio [9–11] and ciliate [12]. Enteritis is a serious disease restricting the practical aquaculture of the lined seahorse with strong infectiousness and high fatality rates [11,13], especially occurring most frequently among juveniles during a short period from the newborn to approximately one month after birth. The popular antibiotics for fish aquaculture practices, however, often cause drug resistance [14], and residual drug issues. Thus, it is urgent to find a safer antibacterial alternative to improve the quality and quantity of cultivated seahorse products and sustainable development of these fish species.

Antimicrobial peptides (AMPs), a group of amphiphilic and cationic short peptides, are in the first line of defense for the host against infectious pathogens with extensive existence in organisms. Currently, the public database of antimicrobial peptides (APD3, <http://aps.unmc.edu/AP/main.php>) contains 3,138 AMPs from six kingdoms [15]. Most AMPs either have their own genes or come from the hydrolysis of other immune-related proteins. There is growing evidence for derivation of AMPs from the hydrolysis of some large proteins, like histone, lectin, thrombin, hemoglobin, and so on. For example, Buforin I, a histone H2A-derived AMP from the Asian toad (*Bufo gargarizans*), was proved to have antimicrobial activity [16]. The most known antibacterial mechanisms for these AMPs include cell membrane osmosis and internalization [17]; in the former, AMPs form transmembrane pore channels and damages to the cell membrane, while AMPs in the latter enter the bacterial cells directly and act on specific intracellular targets to inhibit the synthesis of nucleic acids and proteins. Antimicrobial activities of some AMPs are related to their structures. A review [18] already investigated and discussed major classes of AMPs in fishes; among them, β -defensins and hepcidins could give rise to a conserved β -sheet topology with several disulfide bridges [19,20]. Recently, some of these AMPs have been extensively studied and widely used in the fields of pharmacy, health care products, and agriculture practices [21–23]. As one of the best alternatives for antibiotics, AMPs have broad application prospects, especially in world-wide freshwater and marine aquaculture [23,24].

Our previous study [25] identified 507 AMPs/AMP precursors from the genome and transcriptome data of two amphibious mudskippers, *Boleophthalmus pectinirostris* (BP), and *Periophthalmus magnuspinnatus* (PM); among them, 23 AMPs/AMP precursors were screened out and synthesized for antibacterial experiments; hemoglobin β 1- and amylin-derived AMPs were further proved to effectively inhibit the growth of *Micrococcus luteus*, indicating that relevant research is offering a theoretical basis for disease prevention in fishes. In the present study, we aim to identify putative AMP sequences from our previously published genome and transcriptome data of the lined seahorse in a high-throughput way, and then to check their transcriptional changes in various tissues during development. Production in the lined seahorse of some potential AMPs/AMP precursors is validated at the protein level based on our newly generated proteome data in the present study. Furthermore, differential gene expression between male and female individuals based on the proteome data was also analyzed for potential sexual variances and specificity.

2. Results

2.1. Summary of Achieved Data from Our Previously Published Genome and Transcriptomes

We collected 2,927 AMP amino acid sequences from the APD3 database (File S1), which was used as a local database of query sequences. Genome sequences and 19 transcriptome datasets of the lined seahorse from our previous reports [1,26] were employed as the seed database. The 19 transcriptome datasets cover several tissues at different developmental stages, with each sample from a single lined seahorse (Tables 1 and 2). After running BLAST and in-house script to search, filter and integrate our results, we obtained 256, 186 and 267 putative AMP sequences from the annotated gene set (i.e., the coding sequence of each predicted gene [26]), assembled genome scaffolds and transcriptome sequences, respectively (Files S2 and S3). After removal of redundant sequences, we finally determined a total of 290 putative AMPs and AMP precursors in the lined seahorse (File S4). After removal of 18 sequences identified from genome scaffolds (for their long lengths), we listed full-length coding sequences of other 272 AMPs/AMP precursors with yellow marks of the aligned regions in File S5.

Statistics of these identified putative AMP/AMP precursor numbers from the 19 transcriptome datasets are summarized in Tables 1 and 2. According to the annotations in the public APD3 database, we classified these putative AMPs/AMP precursors into 24 groups, including thrombin, histone, lectin, defensin, cyclin-dependent protein kinases (*CDK-like*) [27], antiproteinase (*antipro*) [28], bovine pancreatic trypsin inhibitor (*BPTI*) [29], phospholipase A2 (*PLA2*) [30], liver-expressed antimicrobial peptide 2 (*LEAP-2*) [31], Beta2-microglobulin ($\beta 2m$) [32], enhancers of rudimentary (*ERH*) [33], glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) [34], and so on (see more details in Figure 1).

Table 1. Statistics of the identified putative AMP/AMP precursor numbers from the transcriptome datasets [1] of various tissues.

Tissues	Juvenile ¹	Rudimentary ²	Pre-Pregnancy ³	Pregnancy ⁴	Post-Pregnancy ⁵
Brain	233	234	237	240	241
testis	241 (haslet)	252	246	256	241
pouch	240 (body)	225	225	242	227

¹ Three months post-birth; ² Five months post-birth; ³ Seven months post-birth; ⁴ Approximately eight months post-birth, carrying 10-day embryos in the brood pouch; ⁵ Males collected in 10 days after being released from their offspring.

Table 2. Statistics of the identified putative AMP/AMP precursor numbers from the transcriptome datasets [1] of embryos and juveniles.

Sample	1-D Embryo *	3-D Embryo *	10-D Embryo *	1-D Juvenile #
Number	219	245	242	246

* 1/3/10-day post-fertilization; # Young seahorses collected entirely in the first day after hatching.

Among the 290 putative AMP sequences, 267 are novel (with at least one residue variance from the corresponding query; File S2), and only 23 are identical to the reported query sequences, such as those from six putative ubiquitins (Ubiq, 74-aa peptides originally isolated from the gill of Pacific oyster, *Crassostrea gigas* [35]) and 17 putative histone-derived AMPs (e.g., histone H4-derived AMPs isolated from American cupped oysters *Crassostrea virginica* [36], H3-derived AMPs isolated from calf thymus *Bos Taurus* [37], and H2A-derived AMPs isolated from hemocytes of Pacific white shrimp (*Litopenaeus vannamei*) [38]). As shown in Figure 1, thrombin (*Throm*) is the largest group of AMP precursors, with the mapping of 62 thrombin-derived C-terminal peptides (TCPs). In fact, TCPs usually aggregate with lipopolysaccharide (LPS) and Gram-negative bacteria to form an amorphous amyloid-like substance, contributing to the clearance of those aggregates by phagocytosis [39]. Recently, a high structure resolution of TCPs revealed the vital peptide regions for interaction with LPS and CD14, providing a detailed molecular explanation for the practical antimicrobial activities of TCPs [40].

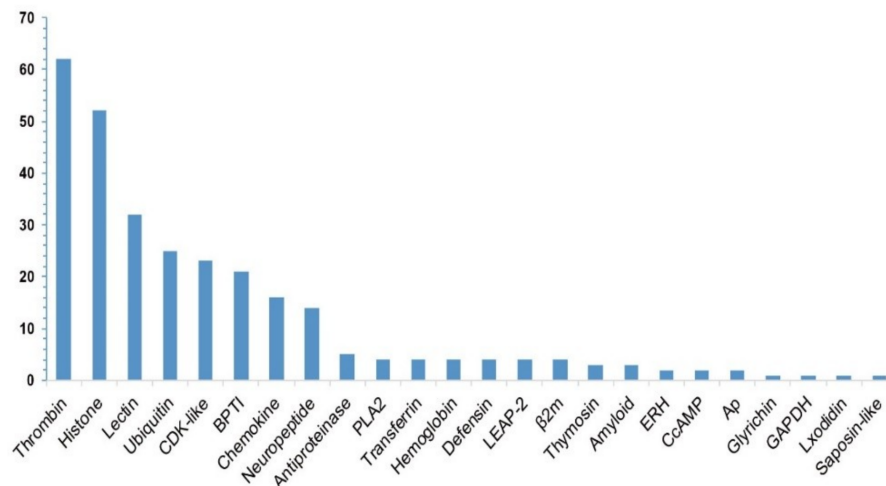


Figure 1. Summary of the identified putative AMP/AMP precursor numbers (the y-axis) from the lined seahorse. A total of 290 putative AMP sequences were determined with a subsequent categorization of 24 groups (the x-axis represents each classified group).

The histone family, the second major group with the mapping of known AMPs in the present study, consists of histones *H2A*, *H4*, *H3*, and *H2B*. Lectins, also a major group among the putative AMP precursors with a reported role in male pregnancy [41], consist of 23 RegIIIgamma (a secreted C-type lectin from the intestine of the mouse) and nine RegIIIalpha (regenerating islet-derived protein 3-alpha from human) [42]. Others, including ubiquitin, *CDK-like*, *BPTI*, and chemokine (*Chem*), also accounted for a large proportion separately.

We performed the Kyoto Encyclopedia of Genes and Genomes (KEGG) [43] clustering analysis on the 272 newly predicted AMP sequences, including 256 from the genome gene set [26] and 16 from our reported transcriptomes [1]. These AMP sequences were searched against the KEGG database with BLAST. They were finally clustered into 193 KEGG pathways. Representative classes include “cancers: an overview”, “immune diseases”, “substance dependence”, and “immune system” (see Figure 2), indicating that most of the putative AMPs/AMP precursors are potentially involved in immune and disease resistance. In addition, we noticed that 27 and 24 putative AMP/AMP precursor genes were clustered into “infectious diseases: bacterial” and “infectious diseases: viral” terms, respectively, which may be relevant to antimicrobial defense.

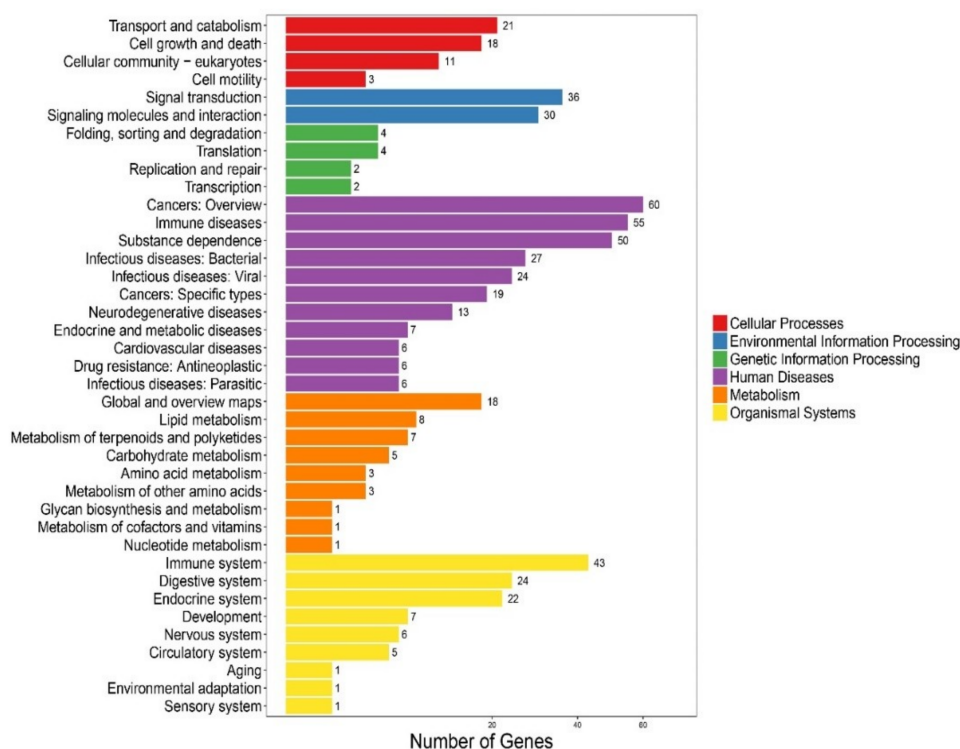


Figure 2. KEGG annotation of the putative AMP/AMP precursor genes in the lined seahorse. A total of 272 predicted genes were clustered into 193 KEGG pathways.

2.2. Representative AMPs/AMP Precursors with High Transcription Levels in the Published Transcriptome Datasets

A total of 19 transcriptomes from several tissues of the lined seahorse were assembled from our previously reported raw data [1]. Finally, we obtained 267 putative AMP sequences by searching against these datasets (File S3). Corresponding transcription values (fragments per kilobase of transcript per million mapped reads, FPKM) were calculated for the quantification of mRNA levels. Top 20 putative AMPs/AMP precursors in each transcriptome were picked out (File S6) for comparison.

Interestingly, we observed that the transcription levels of some lectins were usually extremely high among these identified putative AMP/AMP precursor transcripts. Three with FPKM values over 18,000 are all lectins in the brood pouch. Especially, the FPKM values of *Lectin25* were the highest in the pouch at both pre-pregnancy (58,060.70) and pregnancy (24,092.77) stages, followed by *Lectin11* with an FPKM value of 21,205.30 and *Lectin20* with an FPKM value of 20,844.54 in the pouch at the pre-pregnancy stage. *Lectin20* also highly transcribed in the pouch at the rudimentary stage, with an FPKM value of 18,732.16.

Statistics of these Top 20 putative AMPs/AMP precursors revealed that histones and ubiquitins were highly transcribed in embryos; ubiquitins and neuropeptides (*Neuropep*) were the major classes in the brain; in the pouch, however, lectins and ubiquitins were with high transcription levels; in the testis, ubiquitins and histones accounted for the major part. It seems that ubiquitins were widely distributed in various tissues.

2.3. Transcriptional Changes of the Predicted AMPs/AMP Precursors: Based on Our Previously Published Transcriptome Datasets

Those Top 20 putative AMPs/AMP precursors with low transcription levels and insignificant changes (less than two folds) in the testis were removed, and the remaining data were displayed in Figure 3. In embryos, *Lectin20*, *Lectin23*, *Lectin21*, *Antipro2*, and *Antipro3* were upregulated, while *CDK-like12* and *CDK-like21* were downregulated during the period from 1 to 10 days post fertilization

(ptf1-10); *histone42*, *histone4*, *histone3*, and *histone10* were highly transcribed in ptf1 and low transcribed in ptf3 and ptf10 (Figure 3a). In the brain, *Hemo4* (hemoglobin 4) is the most downregulated putative AMP precursor throughout the period from rudimentary to pregnant stages with very high transcription levels in the juvenile, rudimentary, and post-pregnancy stages. In addition, the synthesized peptide of hemoglobin $\beta 1$, identified by the present high-throughput method, has been proven to have high inhibitive activities on *Micrococcus luteus* [25]. The peptide of hemoglobin $\beta 1$ was downloaded to search against the gene set of the lined seahorse [26] with BLAST (an e-value of 1×10^{-5}), and surprisingly we found that it matched to *Hemo4* with an identity of 68.70% in the protein sequence. The developmental change trend of *Hemo2* was similar to *Hemo4* with a low transcription and change rangeability. *Transfe1* and *Ubiq22* were upregulated and then downregulated, and their transcription levels peaked at the pre-pregnancy and pregnancy stages, respectively, while *Throm63* maintained high transcription levels with a modest change rangeability (Figure 3b).

In the testis, *Throm64*, varying significantly, was downregulated until the pregnancy stage and then upregulated. *Ubiq16* was downregulated and then maintained a relatively stable level after a high transcription level at the pre-pregnancy stage. The transcription levels of *Hemo4*, *Hemo2*, *Ubiq22*, *Transfe1*, *Antipro3*, *Antipro2* and $\beta 2m1$ reached the highest at the pregnancy stage with a similar change tendency. Among them, *Antipro3* and *Antipro2* were lowly transcribed at other stages (Figure 3c). Furthermore, in the pouch, *Lectin25*, *Lectin11*, and *Lectin20* were upregulated and then downregulated, while their transcription levels at the pre-pregnancy stage reached the highest. In addition, the transcription levels of *Lectin25* were always extremely high; *Transfe1* was highly transcribed at the pregnancy stage but with low levels at other stages, and the transcription levels of *Hemo4* kept going up during development (Figure 3d).

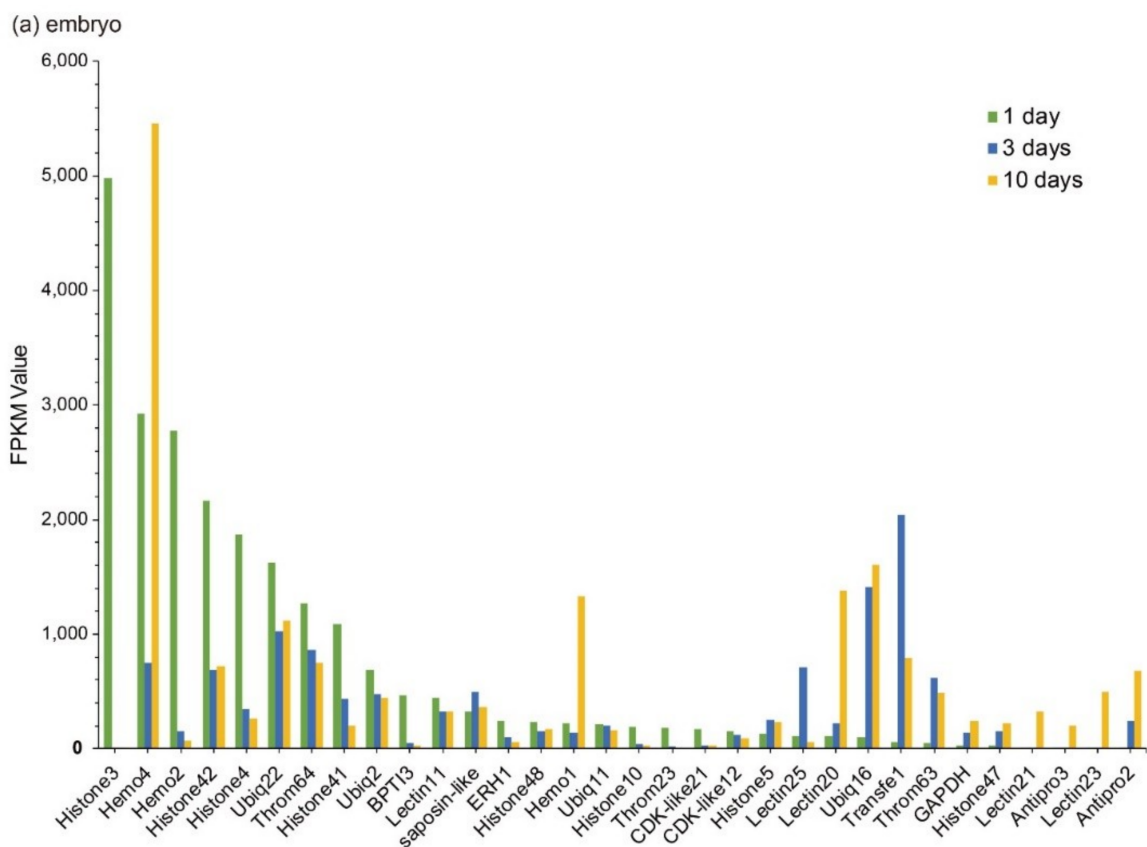
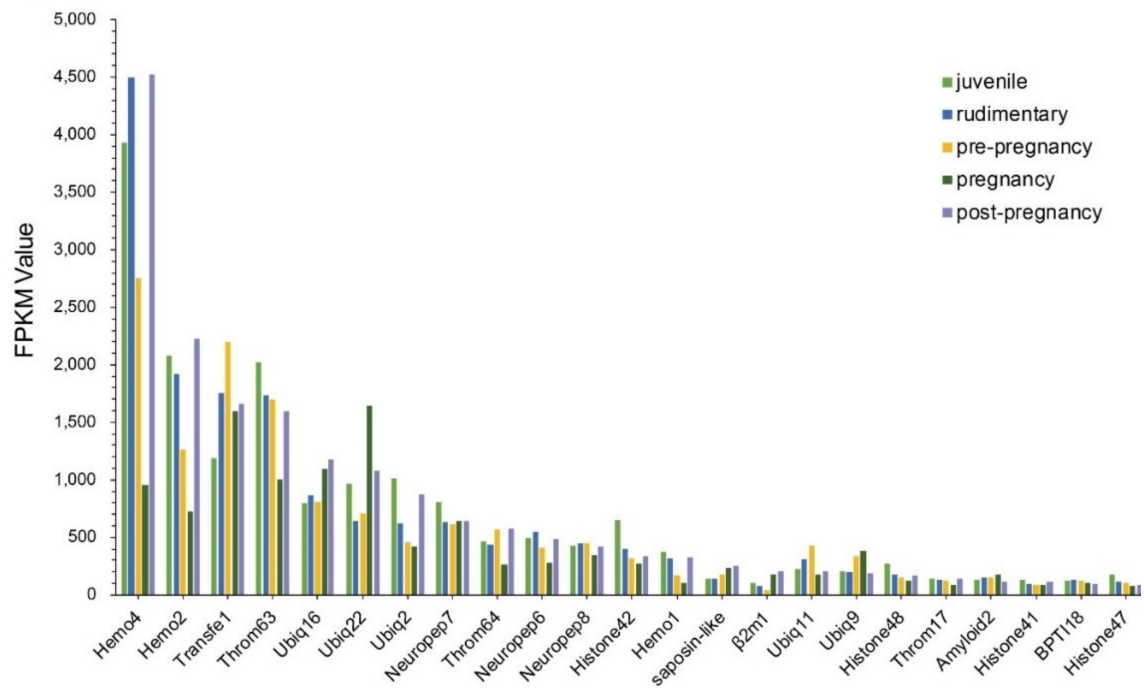


Figure 3. Cont.

(b) brain



(c) testis

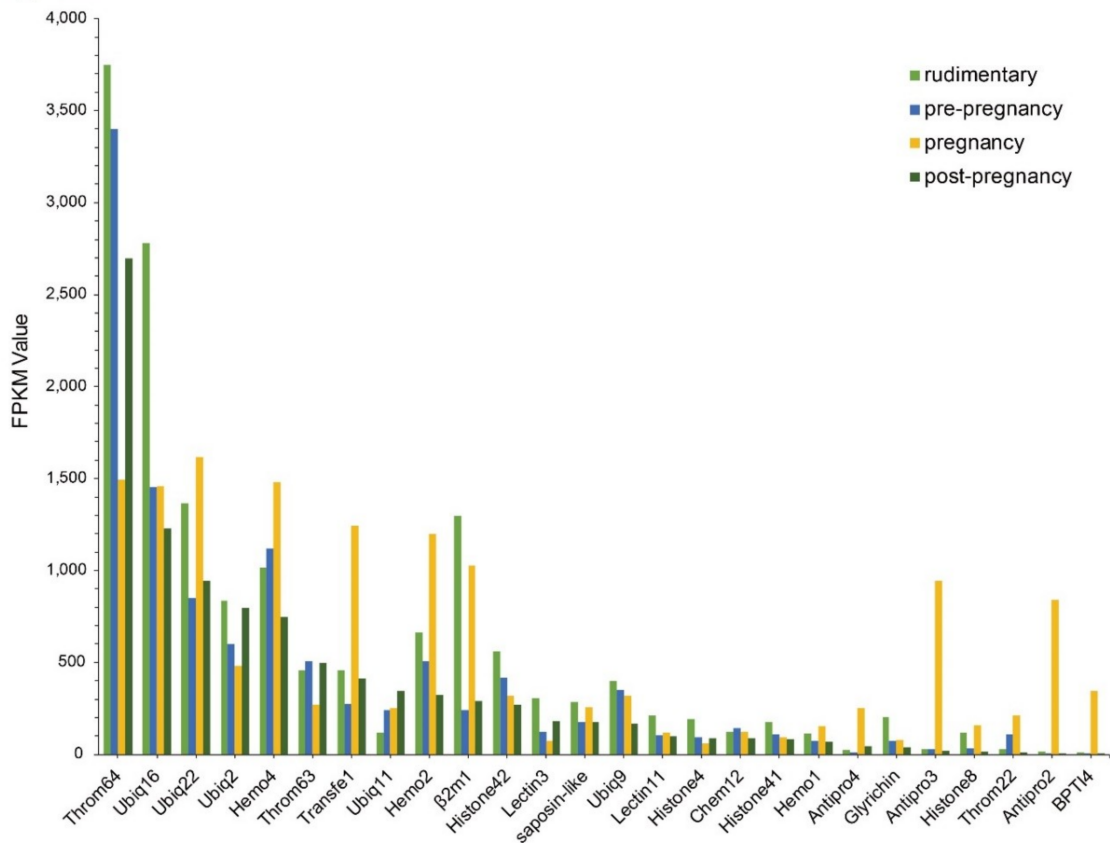


Figure 3. Cont.

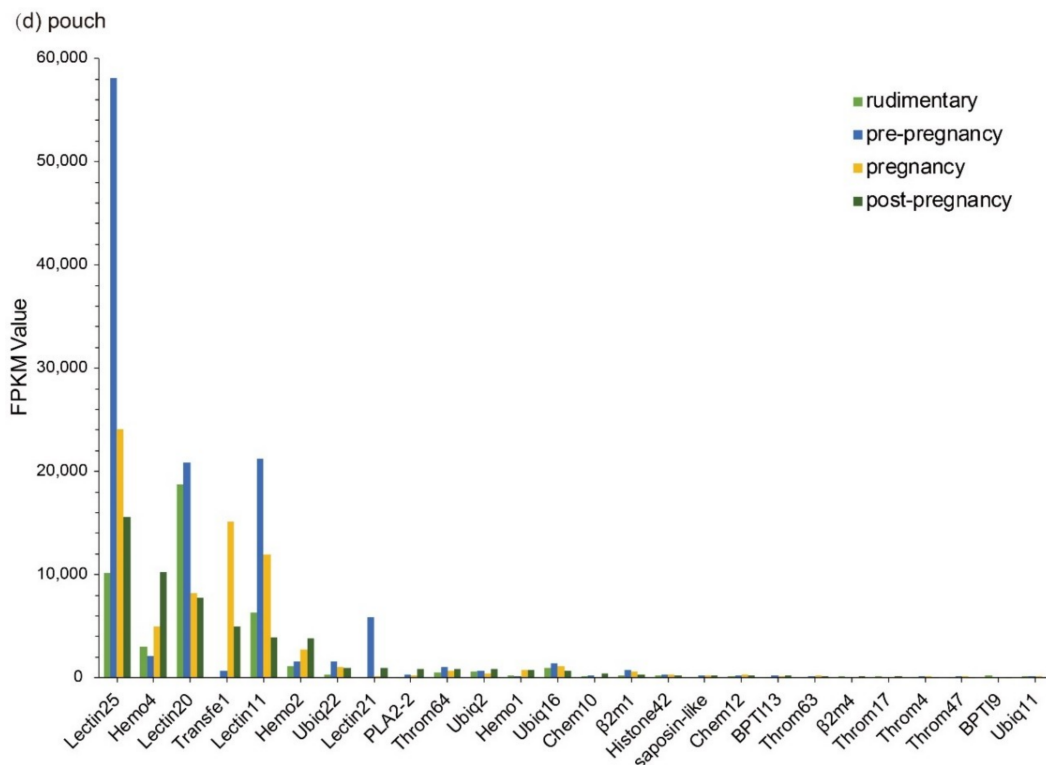


Figure 3. Top 20 putative AMPs/AMP precursors based on their transcription levels (FPKM values) at different stages. Various tissues include (a) embryo, (b) brain, (c) testis, and (d) pouch. The x-axis is the identified putative AMPs/AMP precursors, and the y-axis represents the FPKM values.

Representative putative AMPs or AMP precursors with high transcription levels or significant changes (at least two folds) in each tissue were picked out. Their change patterns at different developmental stages were marked in Figure 4. We speculate that those putative AMPs and AMP precursors with high transcription levels or significant fold changes at a certain stage may play certain vital roles, which are worthy of in-depth investigations into their disease resistance for aquaculture practices.

According to previous studies, juvenile seahorses within one month after birth are more likely to be susceptible to exogenous bacterial infection, and enteritis frequently occurs in this period. Comparisons of transcriptional changes between 1-day and 3-month juveniles were therefore performed. We noted that the sequenced juvenile_body [1] was the mixture of pouch and trunk. Interestingly, we found that, compared with the one-day juvenile stage, *Lectin25*, *Hemo2*, *BPTI10*, and *Chem10* were upregulated significantly in the body (Figure 5a). In the haslet, *Antipro2*, *Transfe1*, *BPTI4*, *Throm15*, *Hemo2*, and *Antipro3* were upregulated significantly (Figure 5b). Furthermore, *Hemo2*, *Neuropep8*, *Throm63*, *Neuropep7*, *Ubiq2*, and *Transfe1* were upregulated significantly in the brain (Figure 5c). Notably, *Hemo2* was upregulated significantly in all examined tissues. *Transfe1* was upregulated in both haslet and brain, and it was predicted as eukaryotic translation initiation factor 4 gamma 1.

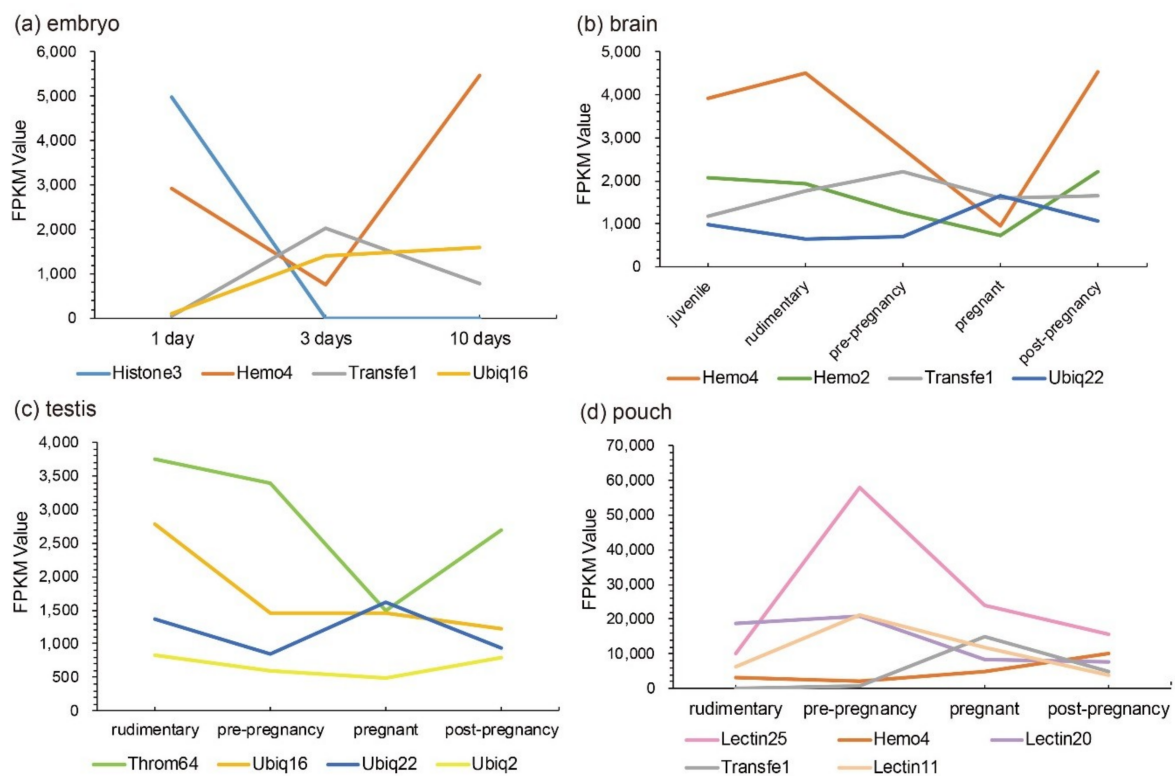


Figure 4. Representative AMPs or AMP precursors with significant changes (over two folds) or high transcription levels (FPKM values). Various tissues include (a) embryo, (b) brain, (c) testis, and (d) pouch. The x-axis is different stages (1 to 10-day embryos or from juvenile to post-pregnancy), and the y-axis represents FPKM values.

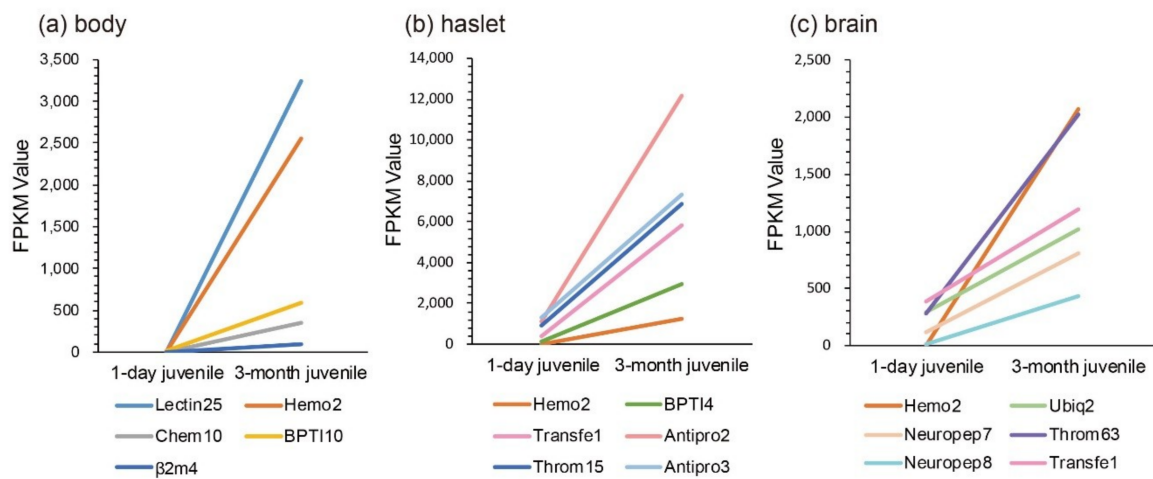


Figure 5. Transcriptional changes of representative AMPs or AMP precursors with significant upregulation and high FPKM values during the development from one-day to three-month juveniles. (a) body; (b) haslet; (c) brain. The x-axis is different stages, and the y-axis represents the FPKM values.

Since these significantly upregulated AMPs or AMP precursors may be involved in the process of antimicrobial defense during the first trimester, they are expected to benefit for improvement of survival rates of juvenile seahorses in practical aquaculture.

2.4. Validation of the Newly Generated Proteome Data

In this study, proteome sequencing of six male and six female lined seahorses was achieved (see more details in Section 4.3). Statistics of total spectra, sequenced peptides, and predicted proteins are summarized in Table S1. MS.F.20 and MS.M.20 were detected by the routine mass spectrometry (MS) analysis for female (F) and male (M) individuals, respectively; their peptides were separated into 20 fractions before MS acquisition. MS.F.1 and MS.M.1 were similar for females and males, respectively, but without fractionation. Moreover, DIA_F and DIA_M covered identification and quantification of proteins by Data Independent Acquisition (DIA) for female and male seahorses, respectively (see more details in Section 4.3.2). After the integration of all the achieved results and removal of redundant data, we finally obtained a total of 6,340 proteins from the examined samples of the lined seahorse. Among them, 6055 proteins were determined for the female and 4,849 proteins for the male seahorses. We also mapped these proteome datasets to the predicted putative AMPs/AMP precursors, with confirmations of 118 (in total), 112 (female), and 94 (male) AMP sequences, respectively. Details of the total 118 AMPs/AMP precursors were listed in File S7.

To explore the sexual differences of putative AMP or AMP precursors, we compared them between male and female seahorses based on the proteome data. Interestingly, we found that 24 putative AMPs or AMP precursors were only detected in the female, and six AMPs or AMP precursors were only available in the male seahorses. Their transcription levels as well, as aligned regional sequences, are shown in File S7. However, we confirmed that several putative AMP/AMP precursor genes were also detectable in transcriptome data of both genders. Consequently, after validation with the transcriptome data, we determined three sexually biased putative AMP precursors, including two male-biased (CDK-like16 and CDK-like23) and one female-biased (BPTI2; File S8).

3. Discussion

3.1. Variances and Consistences of Putative AMPs/AMP Precursors from the Multi-Omics Datasets

We performed high-throughput identification of AMP sequences in the lined seahorse based on multi-omics data, including genome, transcriptome, and proteome datasets. The genome provides an overview of the potential AMPs/AMP precursors, and the transcriptomes provide transcriptional details of these genes; however, the proteome reveals the bioactive peptides (or derived) from these genes. As we discussed before [25], multi-omics data supply more chances to identify more AMPs/AMP precursors and to verify them from different views in gene expression. Especially, the proteome data provide strong evidence for verification of these AMP sequences. In the lined seahorse, we identified 118 putative AMP or AMP precursors based on the proteome datasets, which is different from the total number of genes or transcripts, although they are overlapped among the three resources of data (see more details in Figure 6).

The multi-omics data provide mutual verification and support; however, the quality of sequencing and assembly, as well as different analysis technologies, may lead to indeterminacy in the identified gene number. To confirm that the 16 novel predicted transcripts based on transcriptome data (Figure 6) are inconsistent with the genome-annotated gene set, we employed BLAST (an e -value of 1×10^{-5}) to search against the gene set sequences and observed that these transcripts were not identical to any genes in the annotated gene set [26].

Transcriptome data provide clues for transcriptional changes of AMPs or AMP precursors at several developmental stages or in different tissues, and the proteome data validate proteins for potential production of AMPs. However, analyses based on multi-omics datasets also have limitations in the present study because many AMP sequences are from AMP precursors, although hydrolysis of these large proteins (such as histone, lectin, and thrombin) may not have always occurred. Hence, transcription analysis based on the transcriptome data can't accurately reflect changes of AMPs, and validation based on only proteome data can hardly discriminate the presence of AMPs.

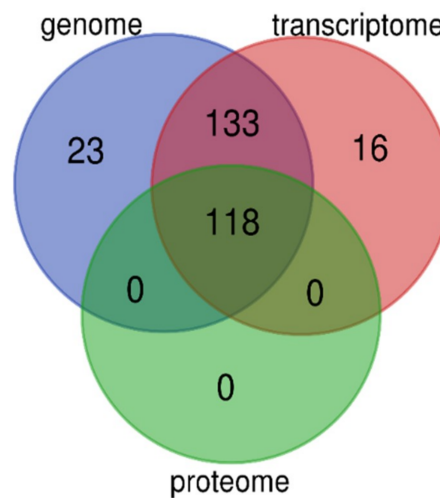


Figure 6. Overlapping of putative AMPs and AMP precursors identified from the genome, transcriptome, and proteome datasets. A total of 118 AMP sequences were validated by these multi-omics resources.

3.2. Species Difference in AMPs/AMP Precursors: Using the Lined Seahorse and Amphibious Mudskippers as the Representatives

AMP studies based on fish genomes and transcriptomes are rare, whereas identification of AMPs or AMP precursors from mudskippers, a group of representative amphibious fishes, were reported by us previously [25]. The comparison between putative AMP sequences from mudskippers and the lined seahorse demonstrated that these teleost fishes have similar putative AMP/AMP precursor numbers (BP, 257; PM, 250; the lined seahorse, 290), although the putative AMPs/AMP precursors identified from the lined seahorse transcriptome and proteome data are approximately twice as many as those from mudskipper transcriptomes.

Based on the detailed results, however, we found that putative AMP groups in the lined seahorse lack piscidin, hepcidin, lysozyme, RNase, misgurin [44], SK84, and lipid transfer protein (LTP). Interestingly, thrombin was the biggest group in both mudskippers and the lined seahorse, indicating their vital roles in fishes. The second biggest group in the lined seahorse, histone, was approximately twice as many as those in BP and three times more than those in PM (BP, 21; PM, 10; the lined seahorse, 52). Previous works have proven that histones have broad-spectrum activities against fish pathogens [45]. Our finding, therefore, suggests that histones may play important roles in the lined seahorse for defending against microorganisms. In general, the major groups of putative AMPs in mudskippers and lined seahorse are similar, but their numbers are a little bit different.

3.3. Representative AMPs/AMP Precursors: Tissue Distribution and Developmental Changes

3.3.1. Lectin25

There exist remarkable differences for many representative AMPs/AMP precursors in tissue distribution and their transcription levels during development. In the pouch where male seahorses nourish embryos, lectins were usually the most abundant AMP precursors with high transcription levels. Among them, *Lectin25* showed obvious developmental changes with extremely high transcription levels, especially at the pre-pregnancy stage (Figure 7a). Based on the proteome datasets, we observed that the *Lectin25* was detectable in the MS data but not in the DIA data. Gene structure prediction indicated that the *Lectin25* has six exons, and the AMP matching region started from the second exon and ended in the fourth exon (172–390th bp; see detailed protein sequence in Figure 7b). *Lectin25* was predicted to be “Galactose-specific lectin nattolectin” by the Swiss-Prot annotation. In fact, nattolectin belongs to the family of C-type lectins and was firstly isolated from the venom of a Brazilian venomous fish

(*Thalassophryne nattereri*) [46]. Nattectin-like protein, a C-type lectin from Qihe crucian carp (*Carassius auratus*), was proven to have an agglutinating activity against bacteria [47,48] and play an important role in antimicrobial defense, also suggesting that this lectin may have antimicrobial activities.

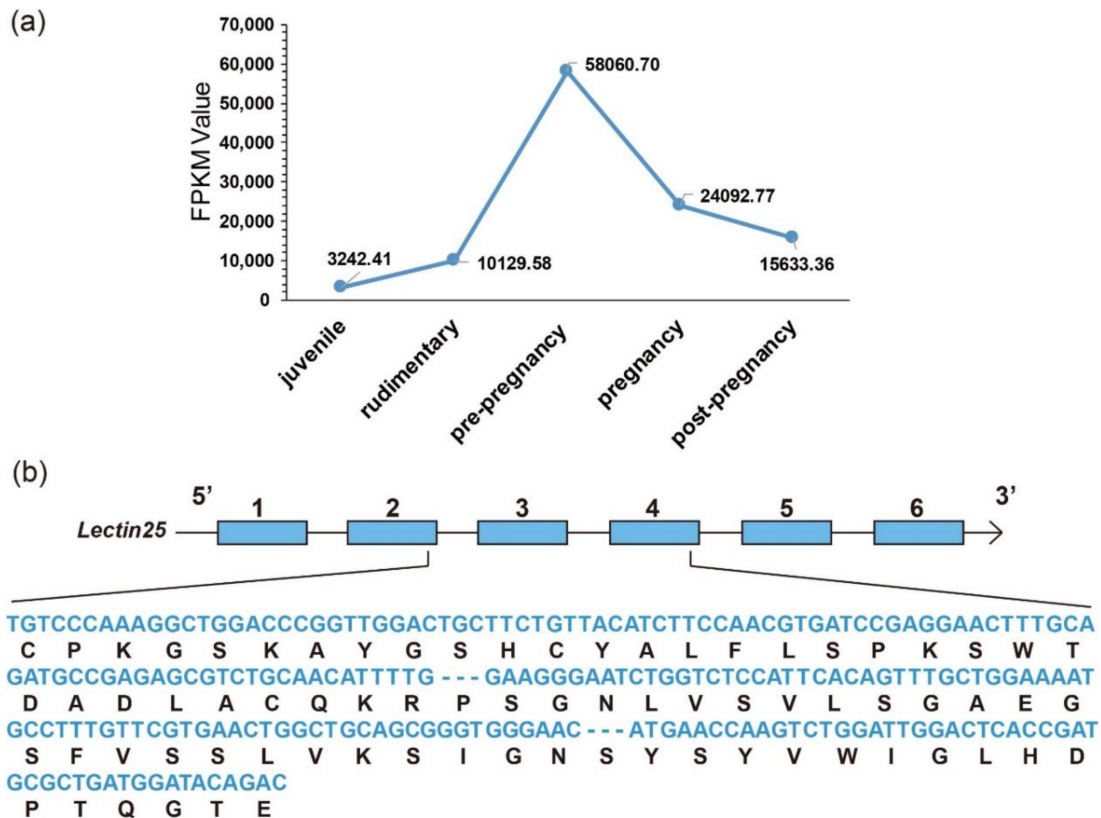


Figure 7. Representative *Lectin25* (Galactose-specific lectin nattectin). (a) Transcriptional changes (FPKM values, the y-axis) during development (the x-axis for various development stages) in the pouch. (b) Predicted gene structure with the aligned AMP sequences (nucleotides in blue, and translated amino-acids in black).

According to previous work [41], *hcCTL III*, one of C-type lectins (CTLs) in the tiger tail seahorse (*Hippocampus comes*), was secreted into the pouch during early pregnancy; bacteriostatic tests showed that *hcCTL III* inhibited the growth of *Escherichia coli*, demonstrating its antibacterial activity. Gene set of the lined seahorse were searched against by BLAST (e-value = 1×10^{-5}) with protein sequences of *hcCTL I*, *II*, and *III* as the queries, and those hits with the highest identity for each *hcCTL* sequence were picked out. Interestingly, we found that *hcCTL III* mapped to *Lectin25* with an identity of 80.2% (Figure 8). It seems that *Lectin25* may play a significant role in embryo antimicrobial protection during male pregnancy [41,49]. Interestingly, lectins have been reported to be associated with human pregnancy. For example, a previous work [50] has proven that deficiency of mannan-binding lectin (MBL) was related to recurrent miscarriage, and MBL concentration significantly increased during pregnancy. Another work [51] found that codon 54 polymorphism of the *MBL* gene may lead to pregnancy-related problems.

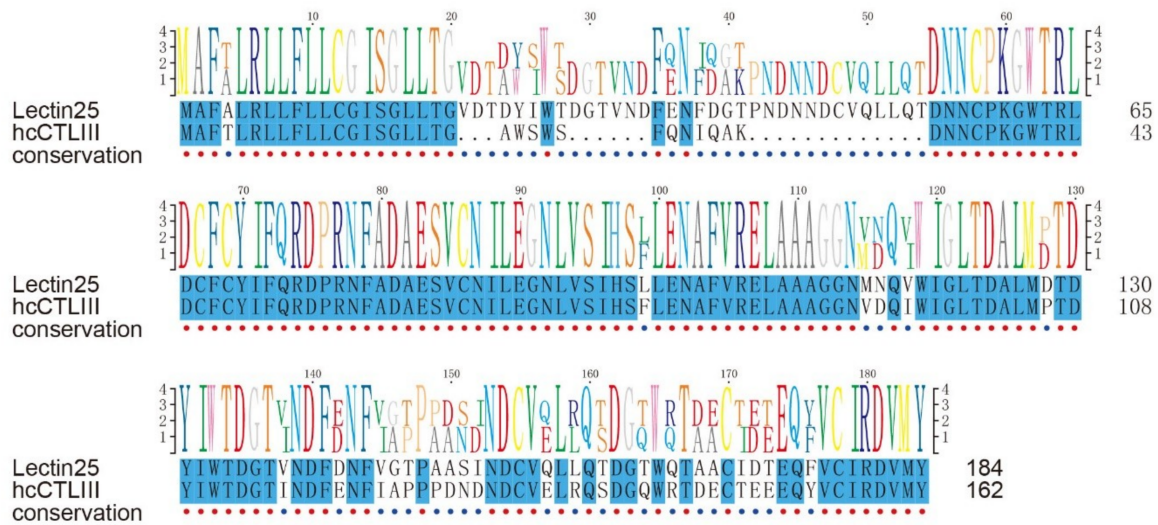


Figure 8. A Pairwise sequence alignment of Lectin25 and hcCTL III. Blue blocks mark the conservation (identity of 100%).

Lectin11 was also highly transcribed at the pre-pregnancy stage in the pouch with an FPKM value of 21,205.30. Interestingly, we found that both *Lectin25* and *Lectin11* located in the same scaffold45 with adjacent localizations. In summary, it seems that lectins are a vital resource of AMPs during seahorse and human pregnancy. We hence propose further studies on these genes to make a great contribution to at least the seahorse breeding and aquaculture.

3.3.2. *Hemo2*

A comparison between 1-day and 3-month juveniles showed that *Hemo2* was significantly upregulated in the seahorse body, haslet, and the brain (Figure 5). *Hemo2* was also detectable in the proteome data, with the relative quantitative values of 2,814,388 and 3,167,032 in female and male lined seahorse, respectively.

Hemoglobin-derived AMPs were proven to have broad-spectrum activities against pathogens in many previous works. For example, from specimens of human endometrium mucus, they were reported to present strong inhibitive activities against *E. coli* ML-35p, ATCC 25 922, and 54 080 [52]. Hb 98-114, a hemoglobin-derived AMP from the midgut of cattle tick (*Rhipicephalus microplus*), is the first hemoglobin reported with specific antifungal activity [52,53]. In addition, HbbetaP-1, derived from hemoglobin beta-subunit in Channel catfish (*Ictalurus punctatus*), was proven to have an anti-parasite activity against *Ichthyophthirius multifiliis* and *Tetrahymena pyriformis* [54]. In summary, hemoglobin's activities against bacteria, fungal and parasites, suggesting their potential roles in enhancing immune defense. Therefore, the representative *Hemo2* may be useful for practical aquaculture of seahorses.

3.4. Sex Differences of Putative AMP/AMP Precursor Genes

In the present study, we identified three sex-differential putative AMP precursors based on both transcriptome and proteome data. Among them, two male-biased putative AMP precursors (*CDK-like16* and *CDK-like23*) were aligned to Scolopendin1 (query ID 2453 in File S1), which was originally identified from Chinese red-headed centipede (*Scolopendra subspinipe*) with antimicrobial and antifungal activities [23]. Other Scolopendins, like Scolopendin and Scolopendin2, was also proven to have broad-spectrum antimicrobial and antifungal activities [55–57]. Although the relationships between male pregnancy and *CDK-like* genes are still unknown, their antimicrobial activities support their potential applications in the improvement of aquaculture qualities of seahorses.

4. Materials and Methods

4.1. Data Collection

A total of 2,624 AMP amino-acid sequences (File S1) were downloaded from the public APD3 database. Genome sequence of the lined seahorse [26], previously reported by us, included assembled genome scaffolds and an annotated gene set. The 19 transcriptome raw datasets, covering multiple developmental stages and various tissues, were publically available in our another paper [1]. We applied a reference genome-based strategy to analyze the transcriptome data. Firstly, the transcriptome raw reads [1] were filtered by SOAPnuke (version 1.5.6) [58] to remove low-quality reads, and those reads with adapter sequences or PCR duplicates. Subsequently, clean reads were aligned to the assembled genome [26] with HISAT2 (version 2.0.4) [59]. Finally, we employed RSEM (version 1.2.12) [60] to calculate the FPKM value of each transcript. The proteome data were generated by us in this study (see more details in Section 4.3).

4.2. Prediction and Identification of Putative AMPs/AMP Precursors in the Lined Seahorse

We used BLAST to predict AMPs, with the 2624 AMP sequences as the queries and above-mentioned 21 multi-omics datasets as the subject sequences. Before running the BLAST ($e\text{-value} = 1 \times 10^{-5}$), we applied a `makeblastdb` command to build an index separately. We subsequently used in-house scripts to deal with alignment information, filtered those hits with query align ratios less than 0.5 and removed redundant data. The aligned region sequences and full-length gene sequences were captured from the archived subject sequences by an in-house script. Classification of putative AMPs was referred to as the annotation of AMPs in the AMD3 database. Finally, we obtained the predicted AMP/AMP precursor genes or sequences with detailed information for each hit.

4.3. Proteomics of the Lined Seahorse

4.3.1. Identification by Mass Spectrometry (MS)

4.3.1.1. Sample Preparation and Protein Extraction

We collected ten 5-month-old lined seahorses, with half male and half female, from BGI-Marine Aquaculture Base in Baguang, Shenzhen, Guangdong Province, China. Abdominal tissues for identification proteomics included skin and muscle, which were cut from the abdomen of each anesthetized seahorse. Tissues were vortexed with 5-mm magnetic beads, and lysis buffer (8 M urea, 40 mM Tris-HCl, 2 mM ethylenediaminetetraacetic acid (EDTA), 1 mM Phenylmethanesulfonyl fluoride (PMSF), pH 8.5). Subsequently, tissues were homogenized in a tissue lyser for two min. The samples were centrifuged at $25,000 \times g$ at 4°C for 20 min. The supernatant liquid was reduced in 10 mM dithiothreitol (DTT; Amresco, Solon, OH, USA) at 56°C for one hour prior to alkylation in 55 mM iodoacetamide (Sigma, St. Louis, MO, USA) at room temperature in the dark for 30 min. The protein concentration was determined by a Bradford kit (Bio-Rad, Hercules, CA, USA).

4.3.1.2. Protein Digestion and Peptide Fractionation

Fractions of 100- μg alkylated proteins were diluted (to less than 2 M urea) with 50 mM tetraethylammonium borohydride (TEAB; Applied Biosystems, Milan, Italy), and then digested by trypsin at a ratio 1:40 (enzyme to protein) at 37°C with overnight incubation. The digested peptides were desalted using a Strata X C18 column (Phenomenex, Torrance, CA, USA) according to the manufacturer's instructions. The peptides then were dried by a speed-vac (LaboGene, Lyngby, Denmark).

Peptides from three male (MS.M.20) and three female (MS.F.20) seahorses were performed with an extra procedure of fractionation. They were re-dissolved by 2-mL buffer A (5% acetonitrile (ACN), pH 9.8), and then separated by a Gemini C18 column (5 μm , 250×4.6 mm; Phenomenex) equipped on a Shimadzu LC-20AB system (Shimadzu Corporation, Kyoto, Japan). Peptides were fractionated using

a gradient of buffer B (95% ACN, pH 9.8) for 10 min, 5–35% buffer B for 40 min, 35–95% buffer B for 1 min, buffer B for 3 min and 5% buffer B for 10 min at a constant flow rate of 1 mL/min. Fractions were collected each min and pooled into 20 fractions [61]. Subsequently, each fraction was dried by speed-vac (LaboGene).

4.3.1.3. Identification of MS Acquisition

The dried fractionated peptides were re-dissolved in buffer C (2% ACN, 0.1% formic acid (FA)) and centrifuged at 20,000× g for 10 min to collect supernatants. Peptides were separated by Shimadzu LC-20AD, on which a 300 μm × 4 mm (μ-Precolumn, Thermo Scientific, Waltham, MA, USA) for peptide enrichment and a 75 μm × 15 cm in-house column for peptide separation were equipped. After enrichment and desalting in a trap column, each fraction was separated in in-house column by a 65-min gradient at 300 nL/min: 5% buffer D (98% ACN, 0.1% FA) for 8 min, 8–35% buffer D for 35 min, 35–60% buffer D for 5 min, 60–80% buffer D for 2 min, 80% buffer D for 5 min, and 5% buffer D for 10 min. Peptides were detected by the Q-Exactive tandem mass spectrometer (ThermoFisher Scientific, San Jose, CA, USA) that coupled to the Shimadzu LC-20AD at a data-dependent acquisition (DDA) mode. MS parameters were set as follows: spray voltage 1.6kV; scan range 350–1600 m/z; MS resolution 70,000; MS/MS resolution 17,500; dynamic exclusion duration 15s.

4.3.1.4. MS Raw Data Analysis

MS raw data were converted to MGF files and searched against the annotated gene set of the lined seahorse with Mascot 2.3.02 [62]. Mascot parameters were listed as follows: enzyme: trypsin; peptide Mass tolerance: 20 ppm; fragment Mass tolerance: 0.05 Da; fixed modifications: Carbamidomethyl (C); variable modifications: Oxidation (M), Deamidation (N, Q); max missed cleavages: 1. To perform quality control for identification, peptide spectrum matches (PSMs) were re-scored by Percolator [63]. Subsequently, a protein inference was performed using Occam's razor approach. The protein false discovery rate (FDR) was calculated by Pickedprotein FDR [64]. Finally, the credible protein identifications were obtained with FDR ≤ 0.01 at both PSM and protein levels.

4.3.2. Quantification by Data Independent Acquisition (DIA)

To acquire the expression level of each identified protein, we performed DIA quantification of individual male and female seahorse. We fully grounded the five-month-old male and female seahorses and collected the homogenates. Protein extraction, digestion, and peptide fraction were performed as mentioned in Sections 4.3.1.1 and 4.3.1.2 except that peptides were separated into 10 fractions in the DDA mode and without fractionation in the DIA mode.

The procedure of MS identification is similar to Section 4.3.1.3 but using different instruments and parameters. In brief, re-dissolved peptides were separated in a 150 μm × 25 cm in-house column equipped on UltiMate 3000 UHPLC (ThermoFisher Scientific) by a 180 min gradient at 500 nL/min: 5% buffer D (98% ACN, 0.1% FA) for 5 min, 5–35% buffer D for 155 min, 35–80% buffer D for 10 min, 80% buffer D for 5 min, and 5% buffer D for 5 min. Fractionated peptides were detected by the Q-Exactive HF (ThermoFisher Scientific) at the DDA and DIA modes, respectively. MS parameters were set as follows: spray voltage, 1.6kV; scan range, 350–1500 m/z; MS resolution at DDA mode, 60,000; MS resolution at DIA mode, 120,000; MS/MS resolution, 15,000; dynamic exclusion duration, 30s.

The DDA MS raw data were identified by MaxQuant (Version 1.5.3.30) [65], and the identification with FDR ≤ 0.01 on both PSM and protein levels was used to build a spectra library [65] for subsequent DIA analysis. MaxQuant parameters were set as follows: enzyme, trypsin; minimal peptide length, 7; PSM-level FDR, 0.01; protein-level FDR, 0.01; fixed modifications: Carbamidomethyl (C); variable modifications: Oxidation (M), Acetyl (Protein N-term); Database, the gene set of the lined seahorse. Based on the spectrum library of DDA data, DIA data was analyzed by Spectronaut 12 [66]. The retention time was corrected by the Spectronaut with indexed retention time (iRT). FDR was calculated by the mProphet approach in Spectronaut.

4.4. Validation of the Proteome Data

We integrated the proteome data and removed those repeat sequences to obtain the final non-redundant protein lists of the total, female, and male lined seahorses. These identified protein lists were mapped to the predicted AMP/AMP precursor genes for overlapping determination.

5. Conclusions

Using a high-throughput strategy, we identified 290 putative AMPs/AMP precursors from our previously reported genome and transcriptome data of the lined seahorse. Among them, 267 are novel, and 118 were further validated by the proteome data that were generated in this study. Transcription levels of top 20 putative AMPs/AMP precursors (with the highest FPKM values) revealed tissue preference of different AMP groups, especially for lectins. For example, *Lectin25* was highly transcribed at the pre-pregnancy and pregnancy stages in the pouch. These data are consistent with previous reports that *Lectin25* is involved in seahorse male pregnancy. A comparison between one-day and three-month juveniles indicated that *Hemo2* was significantly upregulated in the seahorse body, haslet, and the brain, suggesting that *Hemo2* may be useful for the treatment of the popular lethal enteritis in juveniles. Proteome validation of the female and male seahorses identified three sex-differential AMP precursors. Among them, *CDK-like16* and *CDK-like23* were male-biased. In summary, we identified putative AMPs and AMP precursors from multi-omics data of the lined seahorse. These data will contribute to the development of AMP-based fish food additives and human drugs.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1660-3397/18/1/30/s1>. File S1. Collection of reported AMPs. File S2. Alignment results from the lined seahorse genome. File S3. Alignment results from the lined seahorse transcriptomes. File S4. Summary of the 290 identified putative AMPs/AMP precursors. File S5. Full-length coding sequences of the identified putative 272 AMPs/AMP precursors with yellow marks of the aligned regions. File S6. Top 20 AMPs/AMP precursors in each transcriptome. File S7. Identified AMPs/AMP precursors from the lined seahorse proteome data. File S8. FPKM values and sequences of the 24 AMPs/AMP precursors only detected in the female and the 6 AMPs only available in the male lined seahorses based on our proteome data. Table S1. Statistics of spectra, peptides, and proteins in each proteome dataset.

Author Contributions: Q.S. conceived and designed the project; X.C. and Y.Y. performed data analysis; X.Y. and J.L. participated in data analysis and figure preparation; X.C. and Y.Y. wrote the manuscript; Q.S. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Shenzhen Special Project for High-Level Talents (No. SZYSGZZ-2018001), and Shenzhen Dapeng Special Program for Industrial Development (Nos. KY20180205 and PT201901-08).

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

References

- Lin, Q.; Fan, S.; Zhang, Y.; Xu, M.; Zhang, H.; Yang, Y.; Lee, A.P.; Woltering, J.M.; Ravi, V.; Gunter, H.M.; et al. The seahorse genome and the evolution of its specialized morphology. *Nature* **2016**, *540*, 395–399. [[CrossRef](#)] [[PubMed](#)]
- Zhang, N.; Xu, B.; Mou, C.; Yang, W.; Wei, J.; Lu, L.; Zhu, J.; Du, J.; Wu, X.; Ye, L.; et al. Molecular profile of the unique species of traditional Chinese medicine, Chinese seahorse (*Hippocampus kuda* Bleeker). *FEBS Lett.* **2003**, *550*, 124–134. [[CrossRef](#)]
- Kumaravel, K.; Ravichandran, S.; Balasubramanian, T.; Sonneschein, L. Seahorses—A source of traditional medicine. *Nat. Prod. Res.* **2012**, *26*, 2330–2334. [[CrossRef](#)] [[PubMed](#)]
- Sanaye, S.V.; Pise, N.M.; Pawar, A.P.; Parab, P.P.; Sreepada, R.A.; Pawar, H.B.; Revankar, A.D. Evaluation of antioxidant activities in captive-bred cultured yellow seahorse, *Hippocampus kuda* (Bleeker, 1852). *Aquaculture* **2014**, *434*, 100–107. [[CrossRef](#)]
- Lin, T.; Zhang, D.; Liu, X.; Xiao, D. Parental care improves immunity in the seahorse (*Hippocampus erectus*). *Fish Shellfish Immunol.* **2016**, *58*, 554–562. [[CrossRef](#)] [[PubMed](#)]
- Xu, D.H.; Wang, L.H.; Mei, X.T.; Li, B.J.; Lv, J.L.; Xu, S.B. Protective effects of seahorse extracts in a rat castration and testosterone-induced benign prostatic hyperplasia model and mouse oligospermia model. *Environ. Toxicol. Phar.* **2014**, *37*, 679–688. [[CrossRef](#)]

7. Kim, Y.T.; Kim, S.K.; Jeon, Y.J.; Park, S.J. Seahorse-derived peptide suppresses invasive migration of HT1080 fibrosarcoma cells by competing with intracellular alpha-enolase for plasminogen binding and inhibiting uPA-mediated activation of plasminogen. *BMB Rep.* **2014**, *47*, 691–696. [[CrossRef](#)]
8. Chen, L.; Shen, X.; Chen, G.; Cao, X.; Yang, J. Effect of three-spot seahorse petroleum ether extract on lipopolysaccharide induced macrophage RAW264.7 inflammatory cytokine nitric oxide and composition analysis. *J. Oleo Sci.* **2015**, *64*, 933–942. [[CrossRef](#)]
9. Alcaide, E.; Gil-Sanz, C.; Sanjuán, E.; Esteve, D.; Amaro, C.; Silveira, L. *Vibrio harveyi* causes disease in seahorse, *Hippocampus* sp. *J. Fish Dis.* **2001**, *24*, 311–313. [[CrossRef](#)]
10. Qin, G.; Wang, X.; Tan, S.; Lin, Q. A bacterial infection by *Vibrio harveyi* causing heavy reduction of cultured lined seahorse *Hippocampus erectus*. *J. Fish Dis.* **2017**, *40*, 601–605. [[CrossRef](#)]
11. Lin, T.; Zhang, D.; Liu, X.; Xiao, D. Variations of immune parameters in the lined seahorse *Hippocampus erectus* after infection with enteritis pathogen of *Vibrio parahaemolyticus*. *Fish Shellfish Immunol.* **2016**, *50*, 247–254. [[CrossRef](#)] [[PubMed](#)]
12. Di Cicco, E.; Paradis, E.; Stephen, C.; Turba, M.E.; Rossi, G. Scuticociliatid ciliate outbreak in Australian potbellied seahorse, *Hippocampus abdominalis* (Lesson, 1827): Clinical signs, histopathologic findings, and treatment with metronidazole. *J. Zoo Wildl. Med.* **2013**, *44*, 435–440. [[CrossRef](#)] [[PubMed](#)]
13. Wang, X.; Zhang, Y.; Qin, G.; Luo, W.; Lin, Q. A novel pathogenic bacteria (*Vibrio fortis*) causing enteritis in cultured seahorses, *Hippocampus erectus* Perry, 1810. *J. Fish Dis.* **2016**, *39*, 765–769. [[CrossRef](#)] [[PubMed](#)]
14. Okocha, R.C.; Olatoye, I.O.; Adedeji, O.B. Food safety impacts of antimicrobial use and their residues in aquaculture. *Public Health Rev.* **2018**, *39*, 21. [[CrossRef](#)]
15. Wang, G.; Li, X.; Wang, Z. APD3: The antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.* **2016**, *44*, D1087–D1093. [[CrossRef](#)]
16. Kim, H.S.; Yoon, H.; Minn, I.; Park, C.B.; Lee, W.T.; Zasloff, M.; Kim, S.C. Pepsin-mediated processing of the cytoplasmic histone H2A to strong antimicrobial peptide buforin I. *J. Immunol.* **2000**, *165*, 3268–3274. [[CrossRef](#)]
17. Falanga, A.; Lombardi, L.; Franci, G.; Vitiello, M.; Iovene, M.R.; Morelli, G.; Galdiero, M.; Galdiero, S. Marine antimicrobial peptides: Nature provides templates for the design of novel compounds against pathogenic bacteria. *Int. J. Mol. Sci.* **2016**, *17*, 785. [[CrossRef](#)]
18. Masso-Silva, J.A.; Diamond, G. Antimicrobial peptides from fish. *Pharmaceuticals* **2014**, *7*, 265–310. [[CrossRef](#)]
19. Zhu, S.; Gao, B. Evolutionary origin of beta-defensins. *Dev. Comp. Immunol.* **2013**, *39*, 79–84. [[CrossRef](#)]
20. Chaturvedi, P.; Dhanik, M.; Pande, A. Characterization and structural analysis of hepcidin like antimicrobial peptide from *Schizothorax richardsonii* (Gray). *Protein J.* **2014**, *33*, 1–10. [[CrossRef](#)]
21. Mishra, B.; Reiling, S.; Zarena, D.; Wang, G. Host defense antimicrobial peptides as antibiotics: Design and application strategies. *Curr. Opin. Chem. Biol.* **2017**, *38*, 87–96. [[CrossRef](#)]
22. Wang, S.; Zeng, X.; Yang, Q.; Qiao, S. Antimicrobial peptides as potential alternatives to antibiotics in food animal industry. *Int. J. Mol. Sci.* **2016**, *17*, 603. [[CrossRef](#)]
23. Nakatsuji, T.; Gallo, R.L. Antimicrobial peptides: Old molecules with new ideas. *J. Investig. Dermatol.* **2012**, *132*, 887–895. [[CrossRef](#)]
24. Choudhary, A.; Naughton, L.M.; Montanez, I.; Dobson, A.D.W.; Rai, D.K. Current status and future prospects of marine natural products (MNPs) as antimicrobials. *Mar. Drugs* **2017**, *15*, 272. [[CrossRef](#)]
25. Yi, Y.; You, X.; Bian, C.; Chen, S.; Lv, Z.; Qiu, L.; Shi, Q. High-Throughput identification of antimicrobial peptides from amphibious mudskippers. *Mar. Drugs* **2017**, *15*, 364. [[CrossRef](#)]
26. Lin, Q.; Qiu, Y.; Gu, R.; Xu, M.; Li, J.; Bian, C.; Zhang, H.; Qin, G.; Zhang, Y.; Luo, W.; et al. Draft genome of the lined seahorse, *Hippocampus erectus*. *Gigascience* **2017**, *6*. [[CrossRef](#)]
27. Choi, H.; Hwang, J.S.; Lee, D.G. Identification of a novel antimicrobial peptide, scolopendin 1, derived from centipede *Scolopendra subspinipes mutilans* and its antifungal mechanism. *Insect Mol. Biol.* **2014**, *23*, 788–799. [[CrossRef](#)]
28. Bishop, B.M.; Juba, M.L.; Devine, M.C.; Barksdale, S.M.; Rodriguez, C.A.; Chung, M.C.; Russo, P.S.; Vliet, K.A.; Schnur, J.M.; van Hoek, M.L. Bioprospecting the American alligator (*Alligator mississippiensis*) host defense peptidome. *PLoS ONE* **2015**, *10*, e0117394. [[CrossRef](#)]
29. Bleackley, M.R.; Hayes, B.M.; Parisi, K.; Saiyed, T.; Traven, A.; Potter, I.D.; van der Weerden, N.L.; Anderson, M.A. Bovine pancreatic trypsin inhibitor is a new antifungal peptide that inhibits cellular magnesium uptake. *Mol. Microbiol.* **2014**, *92*, 1188–1197. [[CrossRef](#)]

30. Weinrauch, Y.; Abad, C.; Liang, N.S.; Lowry, S.F.; Weiss, J. Mobilization of potent plasma bactericidal activity during systemic bacterial challenge. Role of group IIA phospholipase A2. *J. Clin. Investig.* **1998**, *102*, 633–638. [[CrossRef](#)]
31. Liu, T.; Gao, Y.; Wang, R.; Xu, T. Characterization, evolution and functional analysis of the liver-expressed antimicrobial peptide 2 (LEAP-2) gene in miuuy croaker. *Fish Shellfish Immunol.* **2014**, *41*, 191–199. [[CrossRef](#)]
32. Kim, J.Y.; Park, S.C.; Lee, J.K.; Choi, S.J.; Hahm, K.S.; Park, Y. Novel antibacterial activity of beta (2)-microglobulin in human amniotic fluid. *PLoS ONE* **2012**, *7*, e47642.
33. Agerberth, B.; Boman, A.; Andersson, M.; Jornvall, H.; Mutt, V.; Boman, H.G. Isolation of three antibacterial peptides from pig intestine: Gastric inhibitory polypeptide (7–42), diazepam-binding inhibitor (32–86) and a novel factor, peptide 3910. *Eur. J. Biochem.* **1993**, *216*, 623–629. [[CrossRef](#)]
34. Seo, J.K.; Lee, M.J.; Go, H.J.; Park, T.H.; Park, N.G. Purification and characterization of YFGAP, a GAPDH-related novel antimicrobial peptide, from the skin of yellowfin tuna, *Thunnus albacares*. *Fish Shellfish Immunol.* **2012**, *33*, 743–752. [[CrossRef](#)]
35. Seo, J.K.; Lee, M.J.; Go, H.J.; Kim, G.D.; Jeong, H.D.; Nam, B.H.; Park, N.G. Purification and antimicrobial function of ubiquitin isolated from the gill of Pacific oyster, *Crassostrea gigas*. *Mol. Immunol.* **2013**, *53*, 88–98. [[CrossRef](#)]
36. Dorrington, T.; Villamil, L.; Gomez-chiarri, M. Upregulation in response to infection and antibacterial activity of oyster histone H4. *Fish Shellfish Immunol.* **2011**, *30*, 94–101. [[CrossRef](#)]
37. Tagai, C.; Morita, S.; Shiraiishi, T.; Miyaji, K.; Iwamuro, S. Antimicrobial properties of arginine- and lysine-rich histones and involvement of bacterial outer membrane protease T in their differential mode of actions. *Peptides* **2011**, *32*, 2003–2009. [[CrossRef](#)]
38. Patat, S.A.; Carnegie, R.B.; Kingsbury, C.; Gross, P.S.; Chapman, R.; Schey, K.L. Antimicrobial activity of histones from hemocytes of the Pacific white shrimp. *Eur. J. Biochem.* **2004**, *271*, 4825–4833. [[CrossRef](#)]
39. Petrlova, J.; Hansen, F.C.; van der Plas, M.J.A.; Huber, R.G.; Morgelin, M.; Malmsten, M.; Bond, P.J.; Schmidtchen, A. Aggregation of thrombin-derived C-terminal fragments as a previously undisclosed host defense mechanism. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E4213–E4222. [[CrossRef](#)]
40. Saravanan, R.; Holdbrook, D.A.; Petrlova, J.; Singh, S.; Berglund, N.A.; Choong, Y.K.; Kjellstrom, S.; Bond, P.J.; Malmsten, M.; Schmidtchen, A. Structural basis for endotoxin neutralisation and anti-inflammatory activity of thrombin-derived C-terminal peptides. *Nat. Commun.* **2018**, *9*, 2762. [[CrossRef](#)]
41. Melamed, P.; Xue, Y.; Poon, J.F.; Wu, Q.; Xie, H.; Yeo, J.; Foo, T.W.; Chua, H.K. The male seahorse synthesizes and secretes a novel C-type lectin into the brood pouch during early pregnancy. *FEBS J.* **2005**, *272*, 1221–1235. [[CrossRef](#)]
42. Cash, H.L.; Whitham, C.V.; Behrendt, C.L.; Hooper, L.V. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **2006**, *313*, 1126–1130. [[CrossRef](#)]
43. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)] [[PubMed](#)]
44. Park, C.B.; Lee, J.H.; Park, I.Y.; Kim, M.S.; Kim, S.C. A novel antimicrobial peptide from the loach, *Misgurnus anguillicaudatus*. *FEBS Lett.* **1997**, *411*, 173–178. [[CrossRef](#)]
45. Noga, E.J.; Ullal, A.J.; Corrales, J.; Fernandes, J.M. Application of antimicrobial polypeptide host defenses to aquaculture: Exploitation of downregulation and upregulation responses. *Comp. Biochem. Physiol. D* **2011**, *6*, 44–54. [[CrossRef](#)] [[PubMed](#)]
46. Lopes-Ferreira, M.; Magalhaes, G.S.; Fernandez, J.H.; Junqueira-de-Azevedo Ide, L.; Le Ho, P.; Lima, C.; Valente, R.H.; Moura-da-Silva, A.M. Structural and biological characterization of Nattectin, a new C-type lectin from the venomous fish *Thalassophryne nattereri*. *Biochimie* **2011**, *93*, 971–980. [[CrossRef](#)] [[PubMed](#)]
47. Lv, C.; Zhang, D.; Wang, Z. A novel C-type lectin, Nattectin-like protein, with a wide range of bacterial agglutination activity in large yellow croaker *Larimichthys crocea*. *Fish Shellfish Immunol.* **2016**, *50*, 231–241. [[CrossRef](#)] [[PubMed](#)]
48. Wang, L.; Zhang, J.; Kong, X.; Zhao, X.; Pei, C.; Li, L. A C-type lectin, Nattectin-like protein (CaNTC) in Qihe crucian carp *Carassius auratus*: Binding ability with LPS, PGN and various bacteria, and agglutinating activity against bacteria. *Fish Shellfish Immunol.* **2017**, *67*, 382–392. [[CrossRef](#)]
49. Whittington, C.M.; Griffith, O.W.; Qi, W.; Thompson, M.B.; Wilson, A.B. Seahorse brood pouch transcriptome reveals common genes associated with vertebrate pregnancy. *Mol. Biol. Evol.* **2015**, *32*, 3114–3131. [[CrossRef](#)]

50. Kilpatrick, D.C. Mannan-binding lectin concentration during normal human pregnancy. *Hum. Reprod.* **2000**, *15*, 941–943. [[CrossRef](#)]
51. Calkavur, S.; Erdemir, G.; Onay, H.; Altun Koroglu, O.; Yalaz, M.; Zekioglu, O.; Aksu, G.; Ozkinay, F.; Akercan, F.; Kultursay, N. Mannose-binding lectin may affect pregnancy outcome. *Turk. J. Pediatr.* **2015**, *57*, 26–33. [[PubMed](#)]
52. Wang, L.L.; Pan, X.L.; Ma, Y.; Li, M.; Feng, Y.; Wu, Q.; Wang, B.Y.; Huang, N. Antimicrobial peptide may originate from hemoglobin and play a role in the antibacterial mechanisms of endometrium. *Natl. Med. J. China* **2006**, *86*, 2044–2048.
53. Belmonte, R.; Cruz, C.E.; Pires, J.R.; Daffre, S. Purification and characterization of Hb 98-114: A novel hemoglobin-derived antimicrobial peptide from the midgut of *Rhipicephalus (Boophilus) microplus*. *Peptides* **2012**, *37*, 120–127. [[CrossRef](#)] [[PubMed](#)]
54. Ullal, A.J.; Noga, E.J. Antiparasitic activity of the antimicrobial peptide HbbetaP-1, a member of the beta-haemoglobin peptide family. *J. Fish Dis.* **2010**, *33*, 657–664. [[CrossRef](#)] [[PubMed](#)]
55. Lee, W.; Hwang, J.S.; Lee, D.G. A novel antimicrobial peptide, scolopendin, from *Scolopendra subspinipes mutilans* and its microbicidal mechanism. *Biochimie* **2015**, *118*, 176–184. [[CrossRef](#)] [[PubMed](#)]
56. Lee, H.; Hwang, J.S.; Lee, J.; Kim, J.I.; Lee, D.G. Scolopendin 2, a cationic antimicrobial peptide from centipede, and its membrane-active mechanism. *Biochim. Biophys. Acta* **2015**, *1848*, 634–642. [[CrossRef](#)]
57. Lee, H.; Hwang, J.S.; Lee, D.G. Scolopendin, an antimicrobial peptide from centipede, attenuates mitochondrial functions and triggers apoptosis in *Candida albicans*. *Biochem. J.* **2017**, *474*, 635–645. [[CrossRef](#)]
58. Chen, Y.; Chen, Y.; Shi, C.; Huang, Z.; Zhang, Y.; Li, S.; Li, Y.; Ye, J.; Yu, C.; Li, Z.; et al. SOAPnuke: A MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* **2018**, *7*, gix120. [[CrossRef](#)]
59. Kim, D.; Paggi, J.M.; Park, C.; Bennett, C.; Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **2019**, *37*, 907–915. [[CrossRef](#)]
60. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [[CrossRef](#)]
61. Hao, P.; Ren, Y.; Dutta, B.; Sze, S.K. Comparative evaluation of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) and high-pH reversed phase (Hp-RP) chromatography in profiling of rat kidney proteome. *J. Proteom.* **2013**, *82*, 254–262. [[CrossRef](#)] [[PubMed](#)]
62. Perkins, D.N.; Pappin, D.J.; Creasy, D.M.; Cottrell, J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20*, 3551–3567. [[CrossRef](#)]
63. Kall, L.; Canterbury, J.D.; Weston, J.; Noble, W.S.; MacCoss, M.J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **2007**, *4*, 923–925. [[CrossRef](#)] [[PubMed](#)]
64. Savitski, M.M.; Wilhelm, M.; Hahne, H.; Kuster, B.; Bantscheff, M. A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Mol. Cell. Proteom.* **2015**, *14*, 2394–2404. [[CrossRef](#)] [[PubMed](#)]
65. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372. [[CrossRef](#)] [[PubMed](#)]
66. Bruderer, R.; Bernhardt, O.M.; Gandhi, T.; Miladinovic, S.M.; Cheng, L.Y.; Messner, S.; Ehrenberger, T.; Zanotelli, V.; Butscheid, Y.; Escher, C.; et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell. Proteom.* **2015**, *14*, 1400–1410. [[CrossRef](#)] [[PubMed](#)]

