ISJ 20: 55-64, 2023 ISSN 1824-307X

RESEARCH REPORT

Comparative genomic analysis revealed the ancient duplication of Factor D genes in horseshoe crabs

Y Jia[1#], Q Shen[1#], Z Zhu[1], J Wang[2], X Yu[2], X Du[2], X Liu[1*]

[#]These authors contributed equally to this work

[1]School of Fishery, Zhejiang Ocean University, Zhoushan, Zhejiang, China
[2]CGN (Shengsi) New Energy Holdings Co., Ltd., Zhoushan, Zhejiang, China;

 Accepted November 7, 2023

Abstract

Horseshoe crabs are famous for their blood lysate that is extensively used for the detection of bacterial endotoxin contamination in injectable drugs and medical equipment. However, their existence is now threatened because of overexploitation for bleeding. Synthetic antimicrobial peptides or proteins created by bioengineering may be effective in solving the present predicament. Factor D (FD) is a bioactive substance that shows significant sensitivity to gram-negative bacteria. However, little is known about the expansion information of FD orthologs. In the present study, a total of nine FD orthologs were detected in horseshoe crabs according to the reference genomes. The composition of the FD orthologs in different species was highly conserved, although the common ancestor of the extant horseshoe crabs dates back to the Silurian period. These data suggested that the composition of FD orthologs in horseshoe crabs was formed before the speciation of these species. Considering that three rounds of whole-genome duplication (WGD) events may expanded the homologs in horseshoe crabs, while only one pair of FD orthologs showed significant collinearity according to the intraspecies comparative analysis. Unexpectedly, tandem duplication events that occurred before speciation also expanded the FD orthologs. In addition, several separately distributed FD orthologs also showed high conservation with each other. These data revealed that transposon-mediated duplication may have expanded the FD orthologs in horseshoe crabs. Intriguingly, the C-terminus of the FD orthologs is extremely conserved, although the expansion occurred more than 400 million years ago. In summary, our research provides evidence for the duplication of FD orthologs in horseshoe crabs. Additionally, by demonstrating the conservation of the segments in the C-terminus of these duplicated orthologs, our results also contributed to understanding the functional differentiation mechanism of the diversified FD orthologs.

Key Words: horseshoe crab; antimicrobial peptides; comparative genomic; duplication; functional differentiation

Introduction

Horseshoe crabs are ancient marine invertebrates that belong to the order Xiphosura. Fossil records revealed that the morphology of horseshoe crabs was nearly unchanged for 500 million years, and they are often described as “living fossils” (Rudkin *et al*., 2008). At present, there are only four extant species of horseshoe crabs. Three of them inhabit tropical and subtropical Asia (*Carcinoscopius rotundicauda*, Cro; *Tachypleus tridentatus*, Ttr; *Tachypleus gigas*, Tgi), and the fourth (*Limulus polyphemus*, Lpo) is found along the Atlantic coast of North America (Tarazona *et al*., 2016; Shingate *et al*., 2020a).

Horseshoe crabs live in different coastal habitats during their complete life cycle (Bicknell *et al*., 2022). Generally, it breeds on sandy beaches and then moves to adjacent mudflats to finish hatching and early development (Kendrick *et al*., 2021). Considering the dynamic and highly pathogenic nature of coastal environments, horseshoe crabs have developed multiple efficient innate immune systems in response to a myriad of microbial antigens (Iwanaga, 1993). Previous studies have identified that horseshoe crabs have special hemocytes that are significantly sensitive to the lipopolysaccharides in bacterial endotoxins (Iwanaga, 2007). In this context, the blood lysate of

Corresponding author:
Xiao Liu
School of Fishery
Zhejiang Ocean University
Zhoushan, Zhejiang, China
E-mail: liuxiao@zjou.edu.cn

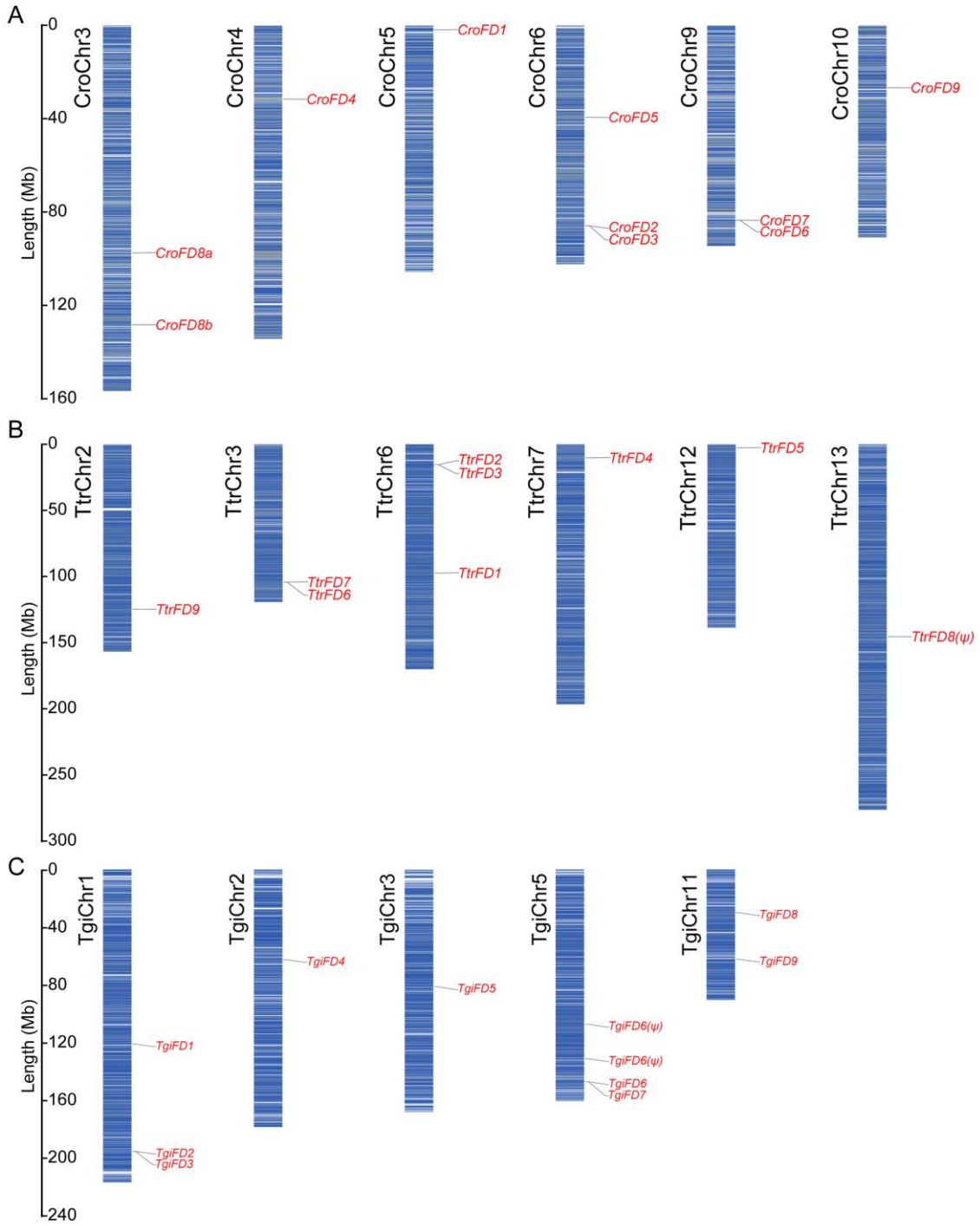

Fig. 1 Chromosome distribution of the FD orthologs in different horseshoe crabs including *C. rotundicauda* (A), *T. tridentatus* (B) and *T. gigas* (C)

horseshoe crabs has been extensively used to detect bacterial endotoxin contamination in injectable drugs and medical equipment. Since the 1970s, the US Food and Drug Administration (FDA) has enforced the horseshoe crab test on experimental drugs and surgical implants to eliminate the harm caused to the human body by endotoxin contamination. As a result, a large number of horseshoe crabs have been overharvested for bleeding by humans. In addition, habitat destruction and marine pollution are two potential reasons for the decline in horseshoe crab populations (Faurby *et al*., 2010; Yang *et al*., 2019).

The immune system of horseshoe crabs is carried by hemolymph and is mainly composed of granular and nongranular cells. The granular cells

that comprise 99% of the total hemocytes are further divided into two populations named large and small granules (Toh *et al*., 1991). The amoebocytes of horseshoe crabs contain various bioactive substances and can be roughly divided into several categories, such as anti-lipopolysaccharide factors, tachyplesins, polyphemusins, big defensins, tachycitins, tachystatins and factor D (FD) (Iwanaga, 2002; Wang *et al*., 2020). In vitro research revealed that these bioactive substances have strong inhibitory effects against bacteria, fungi and some viruses while having low toxicity to normal cells (Kawabata and Iwanaga, 1999). Considering the continuous reduction in horseshoe crabs and the low extraction efficiency of natural bioactive substances, the synthesis of antimicrobial peptides or proteins by bioengineering technology is an efficient way to solve the shortage of natural resources faced by the pharmaceutical field (Owings *et al*., 2019). Moreover, it is also a vital problem that needs to be solved for the protection of the wild resources of horseshoe crabs.

Previous research revealed that the common ancestor of extant horseshoe crabs experienced three rounds of whole-genome duplication (WGD) events during its long evolutionary trajectory (Shingate *et al*., 2020b). Generally, diversification and functional differentiation of most duplicated genes often subsequently occurred after WGD events. For instance, there are types of tachyplesins and polyphemusins that have been identified in horseshoe crabs (Miyata *et al*., 1989; Muta *et al*., 1990). To date, recombinant research on bioactive substances in horseshoe crabs has been performed in both prokaryotic and eukaryotic expression systems (Xu *et al*., 2008; Bolden *et al*., 2020). Functional analysis revealed that these small-molecule peptides were broad-spectrum antimicrobial peptides and had strong inhibitory effects on gram-negative bacteria, gram-positive bacteria, and fungi. In contrast, FD is a high-molecular-weight protein in horseshoe crabs that exhibits specific antibacterial activity. Studies have shown that FD only has inhibitory effects on gram-negative bacteria (Kawabata *et al*., 1996). As a serine protease homolog in which the N-terminal signal sequence is similar to part of the mammalian interleukin-6 receptor α-chain, FD was speculated to play important roles in host defense mechanisms (Kawabata *et al*., 1996). However, little research has been conducted on the functional mechanism of FD in horseshoe crabs. Additionally, similar orthologs have never been described in other arthropods. Considering the WGD events that commonly occurred in the ancient ancestor of horseshoe crabs, duplication and deletion of the FD genes in these species have not been clarified. Recently, the genome assemblies of the four extant horseshoe crab species have been published (Tarazona *et al*., 2016; Gong *et al*., 2019). Remarkably, three of them were assembled at the chromosome level. These reference genomes provided the feasibility to analyze the evolution of FD genes in horseshoe crabs.

In the present study, the complete FD genes in horseshoe crabs were de novo identified by utilizing the available genome assemblies. Then, the phylogenetic relationship of each ortholog was analyzed according to the sequence information of the FD genes. Moreover, the expansion mechanism of each ortholog in different species was determined. At the same time, the conserved domain of FD orthologs was analyzed according to the gene sequence information. The characteristics of the FD genes identified in the present study lay a foundation for subsequent biosynthesis and efficient expression and even provide a comprehensive understanding of the mechanism of FD orthologs in host defense.

Materials and methods

Materials

Available genome assemblies for the four extant horseshoe crab species were selected to perform the related analysis in the present study (Tarazona *et al*., 2016; Gong *et al*., 2019; Shingate *et al*., 2020a; Shingate *et al*., 2020b).

Methods

Identification of the FD genes in four species

The complete amino acid sequence of factor D (Accession No.: BAA13312.1) and other coagulation factors that belong to the serine protease superfamily (clotting factor B: XP_013784210.1; clotting factor C: AAB34362.1; clotting factor G β-subunit: Q27083.1) were used as query sequences to conduct de novo identification in horseshoe crabs using local Tblastn (2.12.0) and Hmmer (v3.3.2) with the default parameters (Camacho *et al*., 2009; Wheeler and Eddy, 2013). The custom HMM profile was autogenerated by utilizing the hmmbuild program based on the query sequences. Meanwhile, the transcriptomes of Ttr (SRR3401116 and SRR12023447) were collected and de novo assembled by utilizing Trinity (v2.12) software (Grabherr *et al*., 2011). The identified sequences in different horseshoe crabs were calibrated by screening the de novo assembled transcripts. For convenience in the subsequent descriptions, we named the FD genes in Tgi according to the order of chromosome location. Then, the corresponding FD orthologs in Cro, Ttr and Lop were named following the order in Tgi.

Phylogenetic analysis of the FD genes in horseshoe crabs

The whole set of identified protein sequences in four horseshoe crab species was initially aligned by utilizing the ClustalX (v2.1) with the default parameters (Larkin *et al*., 2007). The sequence alignments were visualized in DNAMAN9 software. The heatmap that used for the visualization of sequence identity was generated using the plugin TBtools (v1.098769) (Chen *et al*., 2020). Next, to define the identified FD genes in horseshoe crabs, a phylogenetic tree was constructed with the maximum-likelihood (ML) method using IQ-tree software (Nguyen *et al*., 2014). CpREV+I+G4 was the best-fit model chosen according to the Bayesian information criterion (BIC). The tree topology was tested by 5000 ultrafast bootstrap re-samplings. The

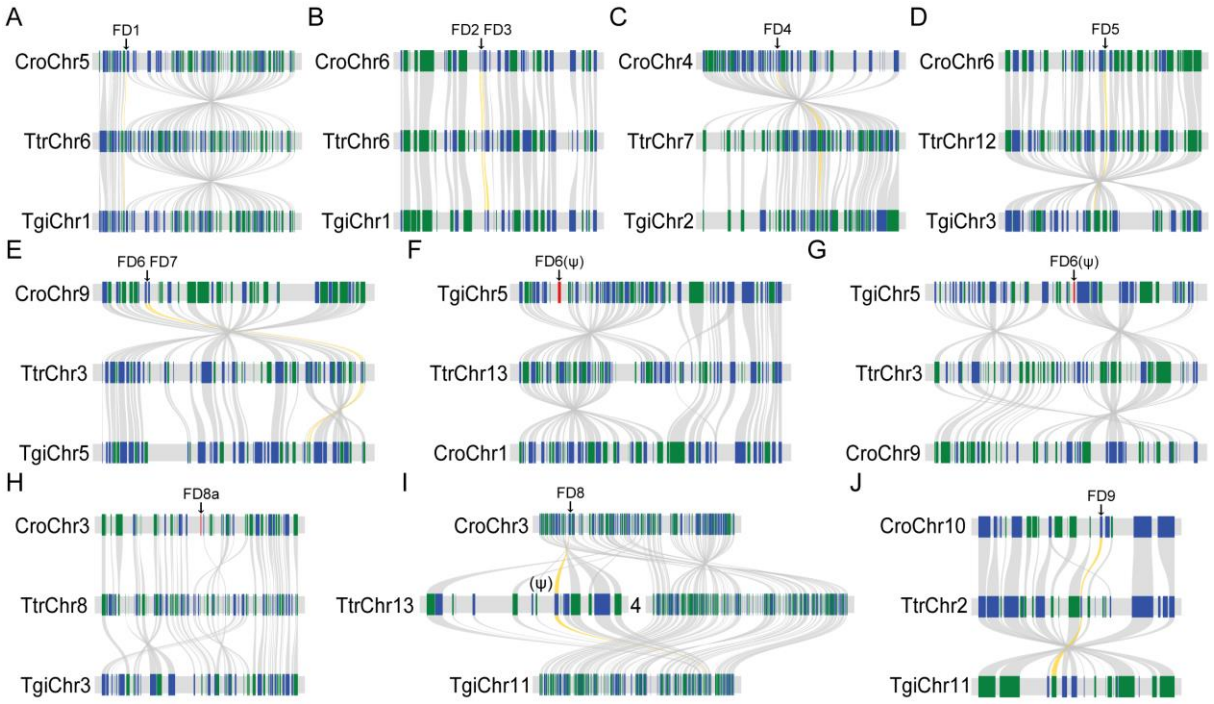

Fig. 2 Synteny analysis of the regions that encoding the nine FD orthologs in different horseshoe crabs. The rectangles colored in green and blue represent the genes that encoded in different reading directions. The pseudogenized FD genes were marked with red rectangles. The orthologs in different species were linked with grey curves. Collinearities between the FD orthologs were linked with yellow curves

phylogenetic tree was visualized using the FigTree (v1.4.4) program. To further analyze the conservation of amino acid sequences among the FD genes, the functional domain of each FD ortholog was searched with the online Pfam program (http://pfam-legacy.xfam.org) (Mistry *et al*., 2021).

Synteny analysis of the FD genes in horseshoe crabs

First, the genomic distribution of the FD orthologs was visualized based on annotation information. To detect the specific expansion or deletion of the diversified FD genes in different species, synteny and collinearity analyses were conducted based on the genome assemblies and corresponding annotation files using the JCVI (v1.1.15) program (Wang *et al*., 2012). In addition, to analyze the micro-synteny relationship of each ortholog, multiple collinearities were visualized using the plugin “Find Gene Block Evolutionary Path by Gene Pairs” in TBtools. To explore the collinearity between the WGD events and the expansion of the FD orthologs, intraspecies genomic comparisons were also performed according to the Tgi genome information by utilizing the one-step MCScanX program in TBtools. Similarly, micro-synteny collinearity was also visualized using the “Find Gene Block Evolutionary Path by Gene Pairs” plugin in TBtools.

Results

Massive amplification of FD genes in horseshoe crabs

To further characterize the consumption of the FD genes in horseshoe crabs, we initially performed the genome-wide identification according to the genome assemblies of four horseshoe crab species. In summary, a total of nine FD orthologs were identified in each of the four horseshoe crab species. Intriguingly, some FD orthologs showed independent expansion or pseudogenization patterns in different species after the speciation of horseshoe crabs. For instance, two FD8 orthologs that were separately distributed on the chromosome were independently identified in the Cro genome (Fig. 1A). In addition, two FD6 orthologs were also partially detected in the genome of Tgi (Fig. 1C). Unexpectedly, segments that encoding the FD6 and FD8 were not detected in the genome of Lpo (Fig. 5C). As a consequence, the number of FD genes in horseshoe crabs varied from 7 to 11 between different species (Fig. 1).

To detect the relative relationship of each FD ortholog on the chromosomes in horseshoe crabs, the distribution of the complete set of FD genes from the three chromosome-level assembled species was visualized according to the genome annotation information. It is obvious that the FD genes were distributed across 6 chromosomes in

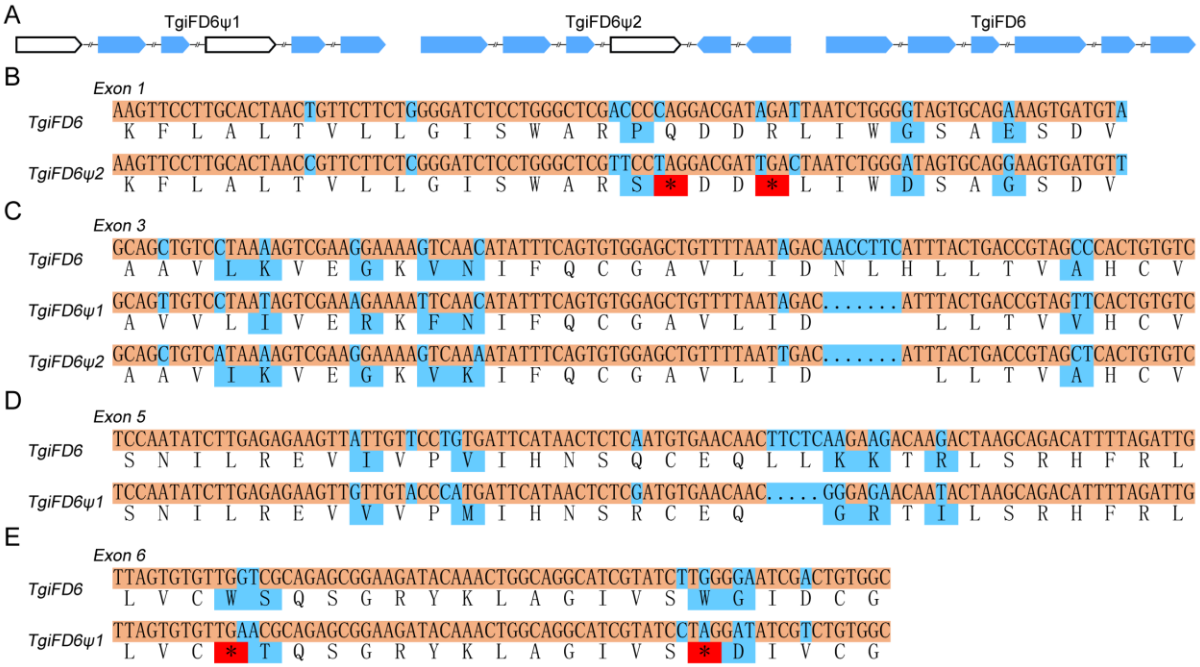

Fig. 3 Pseudogenization of the duplicated FD6 in *T. gigas*. (A) Gene structure of the duplicated FD6 paralogs in *T. gigas*. Blue wedges represent the exon segments. Dashed lines represent the introns. White wedges boxed in black represent the absented exons. (B) Blast analysis of the region that encoding the first exon. (C) Blast analysis of the region that encoding the third exon. (D) Blast analysis of the region that encoding the firth exon. (E) Blast analysis of the region that encoding the sixth exon. The asterisks (*) indicates the in-frame stop codons and were shaded in red. The nucleotide indels or mutations were shaded in light blue

both Cro and Ttr (Fig. 1A and 1B). While in Tgi, the 11 FD genes were distributed across 5 out of 14 chromosomes (Fig. 1C). Unexpectedly, there were only two pairs of FD orthologs that were adjacently localized on the chromosome in these three species (Fig. 1). Except that, the other orthologs were randomly scattered in the genome sequence (Fig. 1). Considering the massive inter-chromosome and inner-chromosome fusions, fissions and even rearrangements after the speciation of different horseshoe crabs (Shingate *et al.*, 2020), the distribution characteristics of the FD orthologs on the chromosomes in different species are inconspicuous.

Synteny analysis of the FD genes in horseshoe crabs

As mentioned above, several FD orthologs were independently duplicated in some horseshoe crabs. In addition, large-scale inner-chromosome rearrangements were also widespread in these species. In this context, we further analyzed the evolutionary conservation of FD orthologs in different horseshoe crab species. Synteny analysis was carried out based on the genome data of the three chromosome-level assembled species. The microsynteny relationship revealed that the majority of FD orthologs were distributed throughout the region that showed high collinearity in these three horseshoe crab species, although some were localized at the edge of inverted regions (Fig. 2). These data demonstrated the high conservation of the FD orthologs in horseshoe crabs, although the speciation of these three species in Asia occurred anciently within the Triassic period (Obst *et al.*, 2012).

It should be noted that in the Tgi genome, except for the conserved FD6 that was encoded in the syntenic segment, two other duplicated FD6 genes were also detected in the other regions (Fig. 2F and 2G). However, similar expansions were not detected in the other two species. These data suggested that duplication events independently occurred in Tgi after the speciation. While blast analysis revealed that the presence of in-frame stop codons and frameshift mutations and even the absence of some exon segments led to the pseudogenization of both of these duplicated FD6 genes in Tgi during the subsequent evolutionary process (Fig. 3). Another independent duplication was detected on the FD8 ortholog in Cro (Fig. 2H). Similar to the duplication of FD6 in Tgi, the expanded FD8 in Cro was also not adjacent distributed with the original ortholog (Fig. 1A). Transposable replication may mediate the expansion of these FD orthologs. Surprisingly, the segments encoding the FD8 in Ttr were distributed in an inverted region (Fig. 2I). Moreover, the FD8 in Ttr was also pseudogenized with several in-frame stop codons (data not shown). It is clear that the pseudogenization of FD8 leads to the contraction of FD orthologs in Ttr.

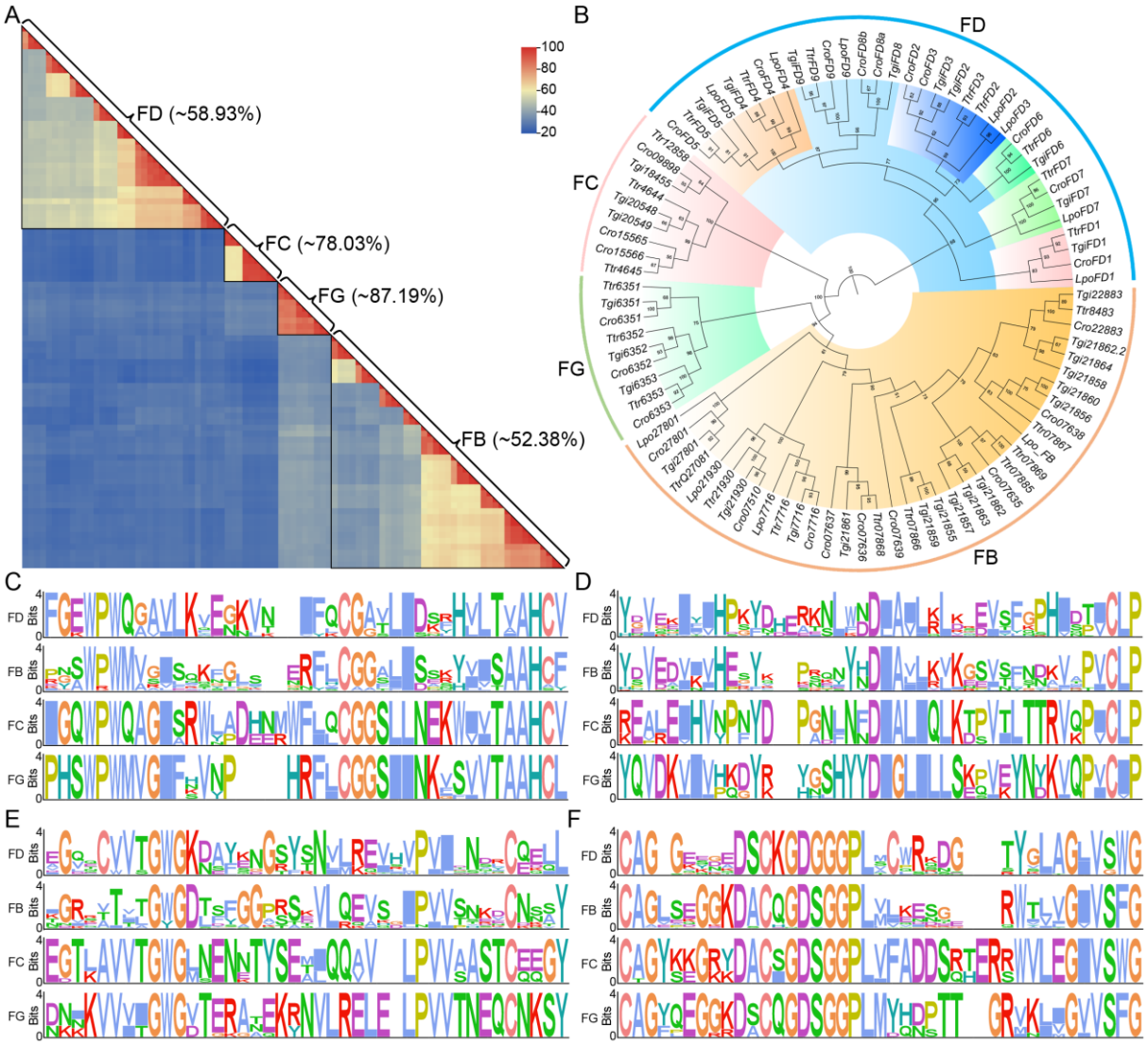

Fig. 4 Phylogenetic analysis of the identified serine proteases in horseshoe crabs. (A) Heatmap illustrating the identities between complete set of the identified serine proteases in horseshoe crabs. (B) Molecular phylogeny of the serine proteases in horseshoe crabs. (C-F) Alignments of the four conserved regions among the identified serine proteases amino acid sequences in horseshoe crabs

Diversity of the FD orthologs in horseshoe crabs

Considering the expansion of the FD genes in horseshoe crabs, sequence alignment was conducted to detect the evolutionary conservation and characteristic motifs in the nine orthologs. As a whole, the similarity of the amino acid sequences between the FD genes and the other serine proteases were significantly lower than that among the nine FD orthologs (Fig. 4A). Phylogenic analysis revealed that the diversified FD orthologs were also distinctly clustered together (Fig. 4B). Additionally, the conserved domain in FD also showed a distinct pattern when compared with that in the others (Fig. 4C-F). In summary, the average similarity among the amino acid sequences of the nine FD orthologs was approximately 50% (Fig. 4A). Intriguingly, the aligned sequences could be clearly distinguished into two independent parts according to the conservation of the corresponding amino acids. The first one-third of the sequence at the N-terminus showed significantly lower conservation in both identity and segment length (approximately 20%) (Fig. 5A). While the other two-thirds of the sequence at the C-terminus showed higher conservation (approximately 75%) among the nine orthologs (Fig. 5A). Pfam scan analysis revealed that the region at the C-terminus overlapped with the trypsin domain (Fig. 5A). Additionally, three peptidase motifs were also detected in the conserved trypsin domain (Mitchell *et al*., 2014).

It should be noted that each FD ortholog in the different horseshoe crab species was highly conserved, although the speciation of these species occurred more than 200 million years ago.

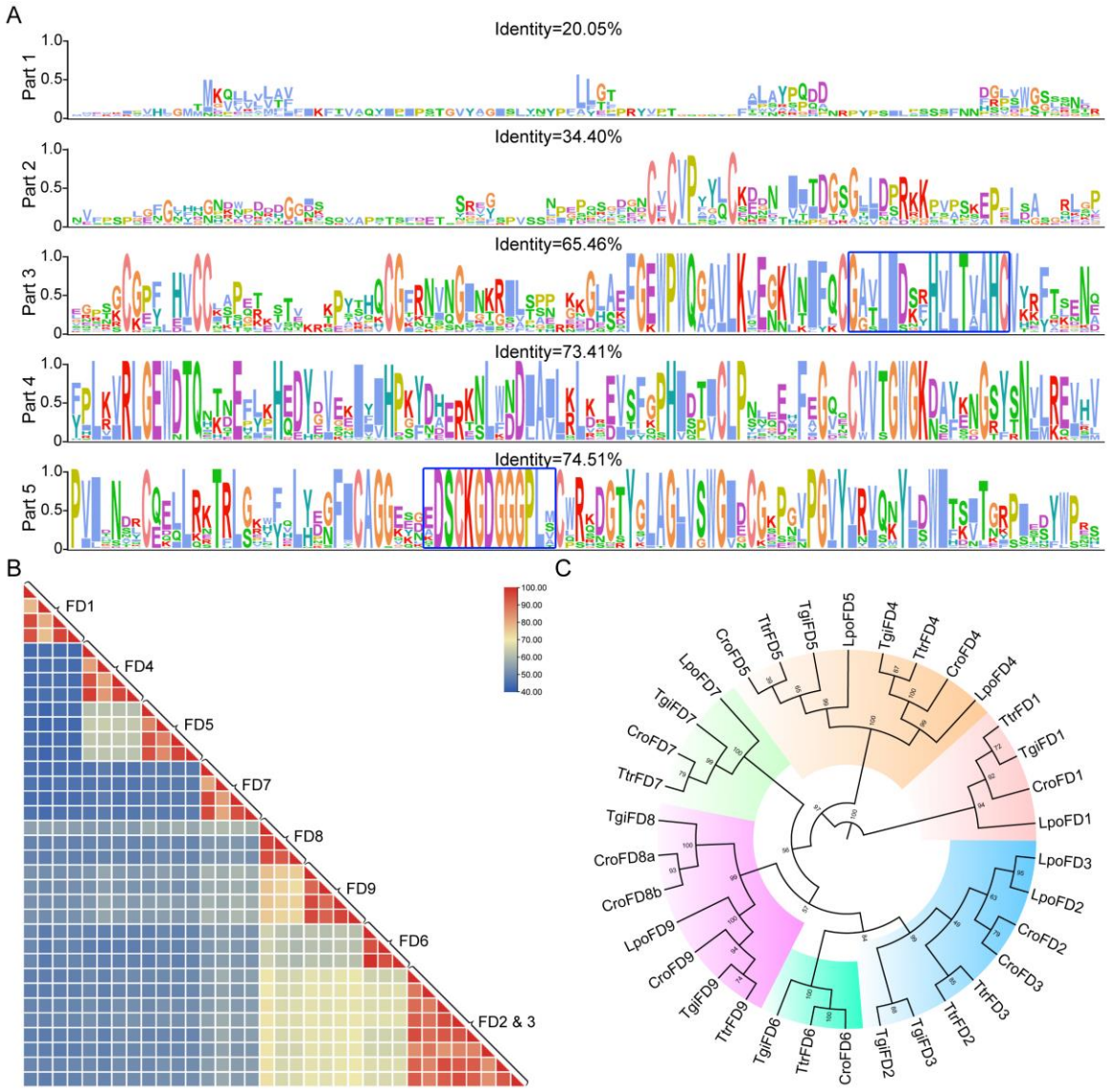

Fig. 5 Conservation of the nine FD orthologs in horseshoe crabs. (A) The conserved amino acids in the sequences of the nine FD orthologs. (B) Similarities between each of nine FD orthologs in the four extant horseshoe crabs. (C) Phylogenetic analysis of the nine FD orthologs in the four different horseshoe crabs

Remarkably, the two pairs of tandemly duplicated FD orthologs showed distinguishing similarities. For instance, FD2 and FD3 were highly conserved with each other (Fig. 5C). Phylogenetic analysis indicated that these two adjacent distributed orthologs in each horseshoe crab species were also clustered into independent subbranches (Fig. 5C). However, the similarity between the conservatively evolved FD6 and FD7 orthologs in these three species was much lower than that between the FD2 and FD3 (Fig. 5C). Phylogenetic analysis demonstrated that the branches including FD6 and FD7 were also not clustered together (Fig. 5C). These data suggested that the tandem duplication event that occurred between FD6 and FD7 was more ancient than the one that occurred between the others.

It is noteworthy that the similarities between FD4 and FD5, two orthologs that are separately distributed in the genome, were higher when compared with the others (Fig. 5C). A similar phenomenon was also detected between the FD8 and FD9 orthologs (Fig. 5C). Phylogenetic analysis showed that the two pairs of FD orthologs were also independently clustered into the same subbranch (Fig. 5C). Considering the three rounds of WGD events that occurred during the evolutionary history of horseshoe crabs, the high conservation between these two pairs of separately distributed FD orthologs suggests that the origin of these genes may be derived from the expansion of WGD.

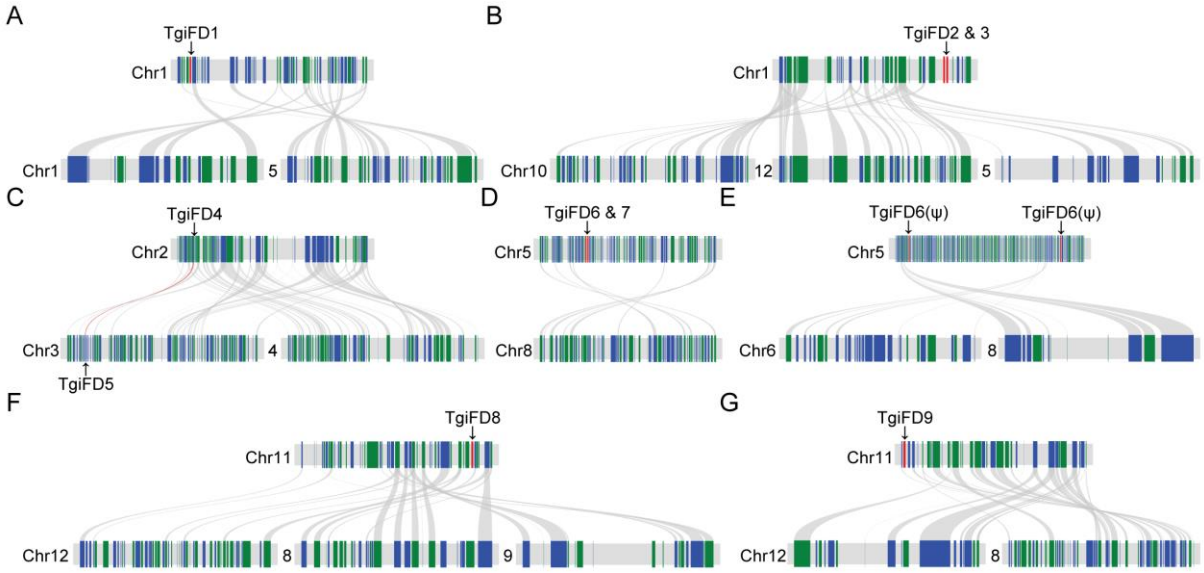

Fig. 6 Intraspecies comparative analysis of the nine FD orthologs in the genome of *T. gigas*. The rectangles colored in green and blue represent the genes that encoded in different reading directions. The FD genes were marked with red rectangles. The orthologs in different chromosomes were linked with grey curves. Collinearities between the duplicated FD orthologs were linked with red curves

Duplication and deletion of the FD orthologs after the WGD events

In order to determine the effect of WGD events on the expansion of the FD orthologs and the subsequent deletion of the duplicated genes, intraspecies comparative genomic analysis was performed based on the genome assembly of Tgi. As we know that the genome of horseshoe crabs has experienced several rounds of WGD events, while rare syntenic collinearity has been detected between the FD orthologs through micro-synteny analysis (Fig. 6). As an exception, a significant collinearity relationship was detected between the region encoding the FD4 and FD5 orthologs (Fig. 6C). These data were consistent with the phylogenetic analysis (Fig. 5C). Unexpectedly, no syntenic collinearity was detected between the region encoding the FD8 and FD9 orthologs, which showed higher similarity and clustered into the same branch (Fig. 5C and Fig. 6). One possibility is that a large number of expanded genes tend to be lost in the subsequent evolutionary process due to functional redundancy.

Discussion

As a group of ancient invertebrates, horseshoe crabs are characterized by stable morphology and contracted species. Additionally, the discovery of horseshoe crabs' highly sensitive immune in response to bacterial endotoxins has led to the use of amebocyte lysate for the assay of endotoxin contaminants in injectable drugs and medical equipment (Kawabata and Iwanaga, 1999). Consequently, overharvesting and bleeding for medical use have dramatically decreased the population of horseshoe crabs (Chen *et al*., 2004). Recombinant antimicrobial proteins/peptides developed through artificial synthesis or genetic engineering is a vital way to solve the current dilemma.

Research on recombinant bioactive substances in horseshoe crab hemolymph through bioengineering started in the 1990s (Matsuzaki *et al*., 1991). However, there are still no available commercial products to date. The majority of recombinant expression studies have focused on tachyplesin, an antimicrobial peptide with 17 amino acids (Matsuzaki *et al*., 1991; Hong *et al*., 2016). Unexpectedly, recombinant expression analysis demonstrated that the expression levels and antibacterial effect of each duplicated tachyplesin homolog were significantly varied (Marggraf *et al*., 2018; Li *et al*., 2019). The genomes of the existing horseshoe crab species are featured by the multiple rounds of WGD events that anciently occurred before the speciation of these species (Shingate *et al*., 2020a). Therefore, selecting an appropriate target gene from the duplicated orthologs for recombinant expression is essential for later experimentation.

Previous studies revealed that most expanded homologs in horseshoe crabs were significantly contracted after WGD events (Shingate *et al*., 2020a). In this study, a total of nine FD orthologs were identified according to the four reference genomes of horseshoe crabs. Unexpectedly, intraspecies collinearity analysis showed that only one pair of FD orthologs (FD4 and FD5) had significant WGD replication characteristics (Fig. 6C). These data suggested that most originally duplicated FD homologs may be subsequently

absent after WGD events. Comparative genomic analysis clearly revealed that the composition of the FD orthologs was extremely conserved among different horseshoe crab species (Fig. 5A). These data demonstrated that the contraction of the FD orthologs were anciently occurred before the speciation of these species (Fig. 5C). Additionally, rare duplications or deletions have been detected during the long evolutionary process after the differentiation of these species.

Previous analysis indicated that massive tandem duplications have been detected in a large number of gene families that are involved in innate immunity in horseshoe crabs (Shingate *et al*., 2020a). Surprisingly, only two pairs of the identified FD orthologs in horseshoe crabs showed tandem replication relationships. It should be noted that some FD orthologs were independently duplicated in different species; however, the duplicated genes were not adjacent to the original genes (Fig. 1). In addition, comparative genomics analysis showed that no inversions were detected around these regions. These data suggested that the duplications mediated by transposons may significantly expand the FD orthologs in horseshoe crabs. Unexpectedly, the orthologs, including FD6, FD8 and FD9, that were separately distributed in the genome showed high similarity with FD2 and FD3, two orthologs that were obtained by tandem replication (Fig. 5B). Similar to the independently duplicated orthologs after speciation, the duplication of these orthologs may also commonly originate from different transposition replication events.

Although the duplication and functional differentiation of the FD orthologs in horseshoe crabs have undergone a long evolutionary process, the amino acids in the C-terminus of the FD orthologs are still extremely conserved (Fig. 5A). The amino acids in these segments compose the essential catalytic unit and share high conservation with homologs in the chymotrypsin subfamily (Rawlings and Barrett, 1994). Previous studies revealed that most members in this subfamily function extracellularly and play important roles in food digestion, fibrinolysis, and complement activation (Rawlings and Barrett, 1994). However, no experiments have been conducted on the correlation between sequence diversification and functional differentiation of the duplicated orthologs. To date, nothing is known about the function of the extended sequences in the N-terminus with low conservation.

In summary, nine FD orthologs were detected in horseshoe crabs. Phylogenetic analysis revealed that the expanded FD orthologs were highly diversified (Fig. 5C). In addition, the diversification of these orthologs anciently occurred before the speciation of these species. As a group of species that have experienced several rounds of WGD events, rare syntenic correlations have been detected between the expanded FD orthologs. Intriguingly, the anciently diversified FD orthologs showed high conservation at the C-terminus (Fig. 5A). However, detailed functional differentiation of these orthologs still needs further investigation.

Acknowledgements

The related bioinformatic analysis in present research were supported by Ocean Science Data Center, IOCAS. This research was financially supported in part by grants from the Natural Science Foundation of China (32301408 and 41676152) and the Key Laboratory of Mariculture of Ministry of Education, Ocean University of China.

References

Bicknell RDC, Kimmig J, Budd GE, Legg DA, Bader KS, Haug C, *et al*. Habitat and developmental constraints drove 330 million years of horseshoe crab evolution. Biol J Linn Soc. 136: 155-172, 2022.

Bolden J, Knutsen C, Levin J, Milne C, Morris T, Mozier N, *et al*. Currently available recombinant alternatives to horseshoe crab blood lysates: are they comparable for the detection of environmental bacterial endotoxins? A review. PDA J Pharm Sci Technol. 74: 602-611, 2020.

Chen CP, Yeh HY, Lin PF. Conservation of the horseshoe crab at Kinmen, Taiwan: strategies and practices. Biodiversity & Conservation. 13: 1889-1904, 2004.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, *et al*. TBtools: An Integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 13: 1194-1202, 2020.

Chen J, Xu XM, Underhill CB, Yang S, Wang L, Chen Y, *et al*. Tachyplesin activates the classic complement pathway to kill tumor cells. Cancer Res. 65: 4614-4622, 2005.

Dai JG, Xie HW, Jin G, Wang WG, Zhang Y, Guo Y. Preliminary study on high-level expression of tandem-arranged tachyplesin-encoding gene in *Bacillus subtilis* Wb800 and its antibacterial activity. Mar Biotechnol. 11: 109-117, 2009.

Faurby S, King TL, Obst M, Hallerman EM, Pertoldi C, Funch P. Population dynamics of American horseshoe crabs - historic climatic events and recent anthropogenic pressures. Mol Ecol. 19: 3088-3100, 2010.

Gong L, Fan G, Ren Y, Chen Y, Qiu Q, Liu L, *et al*. Chromosomal level reference genome of *Tachypleus tridentatus* provides insights into evolution and adaptation of horseshoe crabs. Mol Ecol Resour. 19: 744-756, 2019.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, *et al*. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol. 29: 644-652, 2011.

Hong J, Hu J, Ke F. Experimental induction of bacterial resistance to the antimicrobial peptide tachyplesin I and investigation of the resistance mechanisms. Antimicrob Agents Ch. 60: 6067-6075, 2016.

Iwanaga S. The *Limulus* clotting reaction. Curr Opin Immunol. 5: 74-82, 1993.

Iwanaga S. The molecular basis of innate immunity in the horseshoe crab. Curr Opin Immunol. 14: 87-95, 2002.

Iwanaga S. Biochemical principle of *Limulus* test for detecting bacterial endotoxins. Proc Jpn Acad Ser B Phys Biol Sci. 83: 110-119, 2007.

Kawabata SI, Tokunaga F, Kugi Y, Motoyama S, Miura Y, Hirata M, *et al*. *Limulus* factor D, a 43-kDa protein isolated from horseshoe crab hemocytes, is a serine protease homologue with antimicrobial activity. FEBS Lett. 398: 146-150, 1996.

Kawabata S, Iwanaga S. Role of lectins in the innate immunity of horseshoe crab. Dev Comp Immunol. 23: 391-400, 1999.

Kendrick MR, Brunson JF, Sasson DA, Hamilton KL, Gooding EL, Pound SL, *et al*. Assessing the viability of American horseshoe crab (*Limulus polyphemus*) embryos in salt marsh and sandy beach habitats. Biol Bull-Us. 240: 145-156, 2021.

Li H, Ali Z, Liu X, Jiang L, Tang Y, Dai J. Expression of recombinant tachyplesin I in *Pichia pastoris*. Protein Expres Purif. 157: 50-56, 2019.

Marggraf MB, Panteleev PV, Emelianova AA, Sorokin MI, Bolosov IA, Buzdin AA, *et al*. Cytotoxic potential of the novel horseshoe crab peptide polyphemusin III. Mar Drugs. 16: 466, 2018.

Matsuzaki K, Fukui M, Fujii N, Miyajima K. Interactions of an antimicrobial peptide, tachyplesin I, with lipid membranes. Biochim Biophys Acta. 1070: 259-264, 1991.

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, *et al*. Pfam: The protein families database in 2021. Nucleic Acids Res. 49: D412-D419, 2021.

Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, *et al*. The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res. 43: D213-D221, 2014.

Miyata T, Tokunaga F, Yoneya T, Yoshikawa K, Iwanaga S, Niwa M, *et al*. Antimicrobial peptides, isolated from horseshoe crab hemocytes, tachyplesin II, and polyphemusins I and II: chemical structures and biological activity. J Biochem. 106: 663-668, 1989.

Muta T, Fujimoto T, Nakajima H, Iwanaga S. Tachyplesins isolated from hemocytes of Southeast Asian horseshoe crabs (*Carcinoscorpius rotundicauda* and *Tachypleus gigas*): identification of a new tachyplesin, tachyplesin III, and a processing intermediate of its precursor. J Biochem. 108: 261-266, 1990.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32: 268-274, 2014.

Obst M, Faurby S, Bussarawit S, Funch P. Molecular phylogeny of extant horseshoe crabs (Xiphosura, Limulidae) indicates Paleogene diversification of Asian species. Mol Phylogenet Evol. 62: 21-26, 2012.

Owings M, Chabot C, Watson W. Effects of the biomedical bleeding process on the behavior of the American horseshoe crab, *Limulus polyphemus*, in Its natural habitat. Biol Bull-Us. 236: 207-223, 2019.

Rawlings ND, Barrett AJ. Families of serine peptidases. Method Enzymol. 244: 19-61, 1994.

Rudkin DM, Young GA, Nowlan GS. The oldest horseshoe crab: a new xiphosurid from Late Ordovician Konservat-Lagerstätten deposits, Manitoba, Canada. Palaeontology. 51: 1-9, 2008.

Shingate P, Ravi V, Prasad A, Tay BH, Garg KM, Chattopadhyay B, *et al*. Chromosome-level assembly of the horseshoe crab genome provides insights into its genome evolution. Nat Commun. 11: 2322, 2020.

Shingate P, Ravi V, Prasad A, Tay BH, Venkatesh B. Chromosome-level genome assembly of the coastal horseshoe crab (*Tachypleus gigas*). Mol Ecol Resour. 20: 1748-1760, 2020.

Tarazona OA, Slota LA, Lopez DH, Zhang G, Cohn MJ. The genetic program for cartilage development has deep homology within Bilateria. Nature. 533: 86-89, 2016.

Toh Y, Mizutani A, Tokunaga F, Muta T, Iwanaga S. Morphology of the granular hemocytes of the Japanese horseshoe crab *Tachypleus tridentatus* and immunocytochemical localization of clotting factors and antimicrobial substances. Cell Tissue Res. 266: 137-147, 1991.

Wang WF, Xie XY, Chen K, Chen XL, Zhu WL, Wang HL. Immune responses to gram-negative bacteria in hemolymph of the Chinese horseshoe crab, *Tachypleus tridentatus*. Front Immunol. 11: 584808, 2020.

Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, *et al*. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 40: e49, 2012.

Xu F, Meng K, Wang Y, Luo H, Yang P, Wu N, *et al*. Eukaryotic expression and antimicrobial spectrum determination of the peptide tachyplesin II. Protein Expres Purif. 58: 175-183, 2008.

Yang H, Thompson JR, Flower RJ. Save horseshoe crabs and coastal ecosystems. Science. 366: 813-814, 2019. DOI: 10.1126/science.aaz8654.