Cloning and Bioinformatics Analysis of *hcp* Gene in *Aeromonas hydrophila*

Fan LI[1,2△], Xingjun FU[1,2△], Haiyun FENG[1,2], Shi WANG[1,2], Weijie ZHANG[1,2], Huanying PANG[1,2*], Na WANG[3]

1. Fisheries College, Guangdong Ocean University, Zhanjiang 524025, China; 2. Guangdong Provincial Key Laboratory of Aquatic Animal Disease Control and Healthy Culture & Key Laboratory of Control for Diseases of Aquatic Economic Animals of Guangdong Higher Education Institutes, Zhanjiang 524025, China; 3. Chinese Academy of Inspection and Quarantine, Beijing 100176, China

Abstract [**Objectives**] To explore the function of *hcp* gene in *Aeromonas hydrophila*. [**Methods**] A pair of specific primers was designed referring to the *hcp* gene sequence of *A. hydrophila*. The *hcp* gene was amplified by PCR, and performed bioinformatics analysis. [**Results**] The *hcp* gene had a total length of 1 650 bp and encoded 549 amino acids. The theoretical molecular weight of the protein predicted was about 59 476.44 kDa. After predicting the N-terminal signal peptide structure of the amino acid sequence, neither obvious signal peptide cleavage site nor signal peptide was found, and the protein had no transmembrane region. The amino acid sequence had a N-glycosylation site, 4 protein kinase C phosphorylation sites, 7 casein kinase II phosphorylation sites, 9 N-myristoylation sites, 4 isoprene binding sites, 10 microbody C-terminal target signal sites, and an ATP/GTP binding site motif A (P-ring). The amino acid sequence of *hcp* gene of *A. hydrophila* was performed homology analysis with other *Aeromonas* strains, and it showed higher homology with *A. veronii*. In the secondary structure, the α-helix, β-sheet, random coil and extended strand accounted for 45.36%, 6.01 %, 37.52% and 11.11%, respectively. The tertiary structure model consisted of 18 α-helix and 22 β-sheet. Analysis of protein-protein network interaction demonstrated that the proteins interacting with Hcp protein were AHA_3407, nrfA, nirB-1, nirB-2 and AHA_1112. [**Conclusions**] Through the bioinformatics prediction results, the basic information of *hcp* gene of *A. hydrophila* is preliminarily understood, and the possible function of this protein is predicted, in order to provide guidance for subsequent vaccine research.

Key words *Aeromonas hydrophila*, Gene amplification, *hcp* gene, Bioinformatics analysis

1 Introduction

Aeromonas hydrophila is a gram-negative brevibacterium (G-) belonging to *Aeromonas*, Vibrionaceae. This bacterium, obtuse at both ends and extremely monomastigote, is a common opportunistic pathogen without capsule or spore, and can cause infection individually or simultaneously with other pathogens. *A. hydrophila* is widely found in aquatic environment and can cause outbreaks of hemorrhagic septicemia in a wealth of freshwater fish, as well as human food poisoning, infectious diarrhea and septicemia[1], being an important pathogen that restricts the development of freshwater aquaculture. Since 1989, hemorrhagic septicemia induced by *A. hydrophila* has been violently prevalent in Jiangsu, Hubei and other major aquaculture provinces, seriously affecting local economy and ecology[2]. Therefore, research on the pathogenic mechanism of *A. hydrophila* is of great practical significance for preventing and controlling the occurrence and spread of infections and diseases caused by this pathogen[3]. At present, the pathogenic mechanism of *A. hydrophila* is mainly studied from factors such as motility[4-5], adhesion factors[6-7], toxins[8], enzymes[9], and secretory system[1]. However, little is shown about the type 6 secretory system (T6SS) of *A. hydrophila*. Hemolysin co-regulated protein (Hcp), the core component of T6SS that constitutes its inner tube structure, forms a hexagonal ring to allow effectors to pass through, and protects effectors from being degraded[10]. Hcp can form the puncture device of T6SS with VgrG to inject effectors into target cells[11], and can also play a role as effector protein or molecular companion[12], being a key factor for the normal function of T6SS. Chen Yun *et al.*[13] used Hcp protein as antigen to construct the recombinant plasmid pET-30a-Hcp-scFv containing co-regulatory single-chain antibodies (scFv) against *Pseudomonas plecoglossicida* haemolysin, which was then transformed into the expression strain of *Escherichia coli*, offering a simple method for rapid preparation of antibodies and providing a reference for the development of vaccines.

In recent years, many efforts have been dedicated to *hcp* gene of *Bacillus* and *Vibrio* at home and abroad, such as *Acinetobacter baumannii*[14], *Vibrio harveyi*[15], and *Vibrio anguillarum*[16], while there have been few studies on *hcp* gene of *Aeromonas*, especially *A. hydrophila*. Therefore, in order to explore the function of *hcp* gene of *A. hydrophila*, this paper cloned the *hcp* gene of *A. hydrophila* and conducted bioinformatics analysis of its sequence, laying a foundation for further research on it as a vaccine candidate protein.

Received: November 12, 2023 Accepted: February 28, 2024
Supported by Outstanding Graduate Entering Laboratory Project of College of Fisheries, Guangdong Ocean University; National Natural Science Foundation of China (32073015); Undergraduate Innovation and Entrepreneurship Training Program of Guangdong Ocean University (CXXL2023008); Undergraduate Innovation Team of Guangdong Ocean University (CCTD201802).

△These authors contributed equally to this work.
* Corresponding author. Huanying PANG, PhD., associate professor, research fields: aquatic veterinary medicine.

2 Materials and methods

2.1 Materials

2.1.1 Strains. *A. hydrophila* LP-2 strain was isolated and preserved in Guangdong Provincial Key Laboratory of Aquatic Animal Disease Control and Healthy Culture & Key Laboratory of Control for Diseases of Aquatic Economic Animals of Guangdong Higher Education Institutes[17].

2.1.2 Reagents. (i) Ex *Taq* DNA polymerase was derived from Takara; (ii) bacterial genome DNA extraction kit and DNA glue recovery kit were purchased from Tiangen Biotech Co., Ltd.; (iii) other reagents were imported or domestic analytical pure; (iv) PCR primer synthesis and sequencing were completed by Sangon Biotech (Shanghai) Co., Ltd.; (v) antibiotic ampicillin (Amp^+) was applied at a concentration of 100 μg/mL.

2.1.3 Instruments. Electrophoresis apparatus and PCR apparatus (Bio-Rad); ultra pure water instrument (RightLeder); refrigerated high-speed centrifuge (Eppendorf); HVE-50 series autoclave (Hirayama); gel imaging equipment (Priotein Simple); ultraviolet spectrophotometer (Shimadzu); ultra-low temperature refrigerator (ThermoFisher Scientific).

2.2 Methods

2.2.1 Extraction of *hcp* genome from *A. hydrophila*. *A. hydrophila* was taken out from the ultra-low temperature refrigerator and streaked on the LB plate. Single colonies were picked and inoculated onto LB medium at a ratio of 1∶100, and oscillated in a shaker at 28 ℃, 180 r/min for over 14 h. Small amount of bacterial solution was loaded into a centrifuge tube and centrifuged at 10 000 rpm/min for 2 – 3 min. After the supernatant was discarded, *A. hydrophila* was collected. According to the bacterial genome extraction kit, the total DNA of *A. hydrophila* was extracted, and stored at −20 ℃ for later use.

2.2.2 Cloning of *hcp* gene. Based on the *hcp* gene sequence of *A. hydrophila* LP-2, a pair of primers was designed: forward primer F1: 5'-ATGTTTTGTGTGCAATGTGAACAGA-3'; reverse primer R1: 5'-TCAGGCCGCCAGGATCTCGGCGAGA-3'. Using the total DNA extracted from *A. hydrophila* as a template, PCR was performed in the following procedures: pre-denaturating at 95 ℃ for 3 min; denaturating at 95 ℃ for 30 sec, annealing at 63.3 ℃ for 30 sec, extension at 72 ℃ for 60 sec, 32 cycles; extension at 72 ℃ for 10 min. PCR products were detected by electrophoresis on 1% agarose gel and purified by DNA gel cutting kit.

2.2.3 Connection and sequencing of target fragment and vector. The *hcp* and pMD 18-T vector were connected overnight at 4 ℃. The conjugated products were transferred into *E. coli* DH5α competent cells, and incubated at 37 ℃. After colony PCR detection, the positive clones were sent for testing.

2.2.4 Web sites for bioinformatics analysis. Referring to the methods of Pang Huanying *et al.*[18], bioinformatics websites were utilized for analysis. (i) ExPASy Proteomics Server was employed to predict physicochemical properties of target proteins online, such as amino acid sequence, theoretical isoelectric point (pI), and molecular mass; (ii) SignalP 5.0 Server was applied to predict whether the target protein contained signal peptides; (iii) TMHMM Server 2.0 was utilized to predict whether the target protein contained transmembrane structure; (iv) SoftBerry-Psite was applied to analyze the functional sites of protein amino acid sequence; (v) PSORT Prediction was employed to predict protein subcellular localization; (vi) NCBI was a tool for sequence homology comparison and similarity analysis; (vii) DNAMAN Version 6.0 was applied for amino acid homology analysis; (viii) Clastal 2.0 and MEGA 5.0 software were adopted to construct the phylogenetic tree by neighbor-joining method; (ix) SMART program was utilized to predict protein functional domain; (x) SOPMA software was employed to predict the secondary structure of Hcp; (xi) SWISS-MODEL was applied to predict and construct the tertiary structure of proteins online; (xii) STRING database was employed to search protein-protein network interaction.

3 Results and analysis

3.1 Full-length cloning of *hcp* gene The *hcp* gene was amplified by PCR. The amplification product was analyzed by agarose gel electrophoresis, and specific bands were amplified (Fig. 1). Sequencing of the amplification product and cloning vector pMD18-T showed that the *hcp* gene contained an open reading frame of 1 650 bp and encoded 549 amino acids.

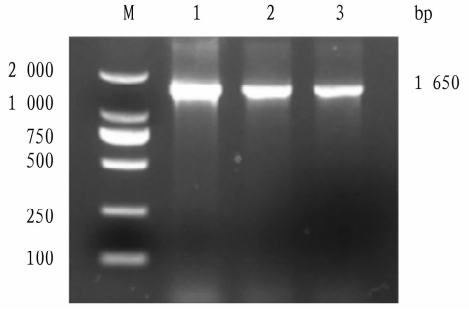


Fig. 1 Cloning of *hcp* gene

3.2 Physicochemical properties The Hcp protein of *A. hydrophila* was analyzed by ExPASy software. The physicochemical properties of the protein were: total number of atoms 8 380, molecular formula $C_{2\,681}H_{4\,196}N_{696}O_{783}S_{24}$, number of amino acids 549, theoretical molecular weight 59 476.44 kDa, theoretical pI 5.00, fat coefficient 98.01, and total mean hydrophilicity 0.151. The instability coefficient was 20.50, indicating the protein was stable. There were 62 acid amino acid residues (Asp + Glu) and 42 basic amino acid residues (Arg + Lys) in the protein, with Met locating at the N terminus. The estimated half-life was 30 h in mammalian reticulocytes (*in vitro*), longer than 10 h in *E. coli* (*in vivo*), and longer than 20 h in yeast (*in vivo*).

3.3 Sequence analysis The N-terminal signal peptide structure of Hcp protein was predicted via SignalP 5.0 Server program, and neither obvious signal peptide cleavage site nor signal peptide

was found, indicating that it was not a secretory protein. Prediction made by TMHMM Server 2.0 program showed that the protein had no transmembrane region. Prediction by SoftBerry-Psite demonstrated that the amino acid sequence contained a N-glycosylation site, 4 protein kinase C phosphorylation sites, 7 casein kinase II phosphorylation sites, 9 N-myristoylation sites, 4 isoprene binding sites, 10 microbody C-terminal target signal sites, and an ATP/GTP binding site motif A (P-ring) (Fig. 2). The prediction results of protein subcellular localization by website wolfpsort suggested that most of them existed in the cytoplasm and mitochondria.

NOTE N-myristoylation site: (15-20 aa, 17-22 aa, 56-61 aa, 157-162 aa, 190-195 aa, 290-295 aa, 402-407 aa, 532-537 aa, 535-540 aa); microbody C-terminal targeting signal (82-84 aa, 100-102 aa, 116-118 aa, 168-170 aa, 341-343 aa, 353-355 aa, 396-398 aa, 418-420 aa, 433-435 aa, 438-440 aa). : Protein kinase C phosphorylation site; ____: N-glycosylation site; : ATP/GTP binding site motif A (P ring); ____: Casein kinase II phosphorylation site; : Isoprene binding site; * : terminator.

Fig. 2 Nucleotides of *hcp* gene and its encoded amino acid sequences

3.4 Homology and evolutionary analysis Homology analysis was performed using DNAMAN software. The amino acid sequences of *hcp* gene from *A. veronii*, *A. sanarellii*, *A. caviae* and *A. dhakensis* were aligned with that from *A. hydrophila*. The results showed that the *hcp* gene of *A. hydrophila* had high homology with that of *A. veronii*, with a similarity of 99.09% (Fig. 3).

The phylogenetic tree was constructed by deduced amino acid sequences of *hcp* gene from *A. hydrophila* and other *Aeromonas* strains via Neighbor-joining method of MEGA 5.0. The results showed that *A. hydrophila* and *A. veronii* clustered into the same subfamily (Fig. 4).

3.5 Prediction of functional domain, secondary and tertiary structure of Hcp protein Prediction by SMART program showed that Hcp protein had a prismane functional domain (1-544 aa) (Fig. 5), with the E value of 3.2e-162. SOPMA software was used to predict the secondary structure of Hcp protein online, and the results showed that the secondary structure of Hcp protein consisted of 45.36% α-helix, 6.1% β-sheet, 37.52% random coil and 11.11% extended strand (Fig. 6).

The amino acid sequence of Hcp protein was submitted to SWISS-MODEL program, and homologous proteins were automatically searched as templates to obtain the tertiary structure model of Hcp protein. The results showed that the protein mainly had 18 α-helix and 22 β-sheet (Fig. 7).

3.6 Protein-protein network interaction of Hcp protein In protein-protein network interaction, it can be found that the proteins adjacent to this protein were AHA_3407, nrfA, nirB-1, nirB-2, and AHA_1112 (Fig. 8).

NOTE *Aeromonas veronii* (WP_201995662.1); *A. sanarellii* (WP_209793950.1); *A. caviae* (WP_257713735.1); *A. dhakensis* (WP_201988758.1).

Fig. 3 Homology comparison of amino acid sequences deduced by *hcp* gene

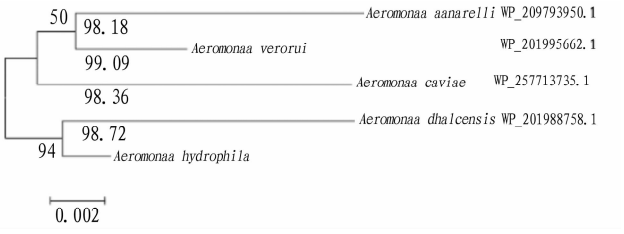


Fig. 4 Phylogenetic tree of *hcp* gene constructed by neighbor-joining method

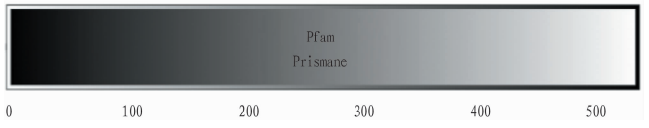
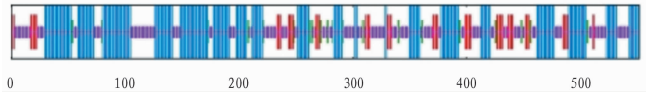


Fig. 5 Functional domain of Hcp protein



NOTE Blue: α-helix; Green: β-sheet; Purple: random coil; Red: extended strand.

Fig. 6 Secondary structure of Hcp protein

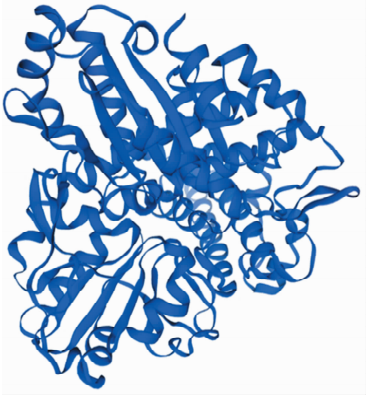


Fig. 7 Three-dimensional structure model of Hcp protein subunit of *Aeromonas hydrophila*

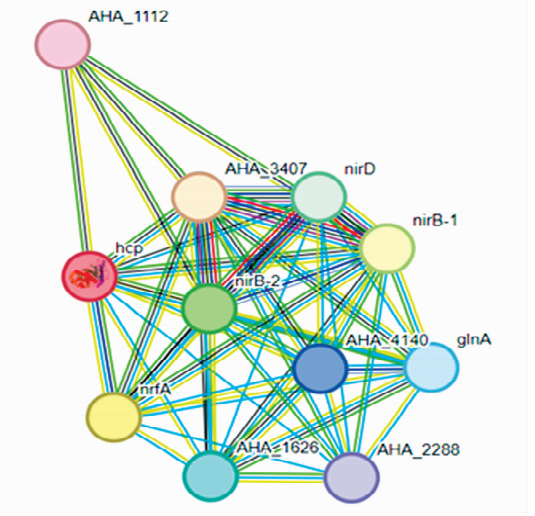


Fig. 8 Protein-protein network interaction of Hcp protein

4 Discussion and conclusions

4.1 Discussion *A. hydrophila* is greatly harmful to aquaculture industry due to different levels of pathogenicity to many kinds of aquatic economic animals, causing great economic losses. At present, chemical drugs and antibiotics are mainly used for the prevention and control of this disease, resulting in increasingly serious problems such as drug residues and increased bacterial resistance[19]. Therefore, seeking excellent protective antigens and developing novel vaccines have gradually become a hot spot in the research of aquatic diseases.

Bioinformatics analysis can accurately predict the evolutionary relationship, physicochemical properties, secondary structure and tertiary structure of target proteins, serving as an important technology for analyzing protein structure and function[20]. Wang Nannan *et al.*[1] conducted bioinformatics analysis on 14 strains of *A. hydrophila* that had completed whole genome sequencing, and found that the recombinant Hcp protein was expected to be a candidate vaccine target protein to prevent *A. hydrophila* infection. Wei Chang *et al.*[21] conducted coning and recombinant expression of pathogenic bacterium *Edwardsiella ictaluri hcp* gene, obtained the fusion protein with high purity after purification, and formulated high-titer polyclonal antibody. In this study, we cloned the *hcp* gene of *A. hydrophila* and performed bioinformatic analysis. The sequence analysis results showed that neither obvious signal peptide cleavage site nor signal peptide was found at the N terminal of Hcp protein, indicating that Hcp protein was not a secretory protein. Prediction made by TMHMM Server 2.0 program demonstrated that the protein had no transmembrane region. The amino acid sequence had a N-glycosylation site, 4 protein kinase C phosphorylation sites, 7 casein kinase II phosphorylation sites, 9 N-myristoylation sites, 4 isoprene binding sites, 10 microbody C-terminal target signal sites, and an ATP/GTP binding site motif A (P-ring).

Studies have proved that Hcp protein plays a good role of immunoprotection[22]. We used bioinformatics analysis websites to analyze the characteristics of Hcp protein and explore the potential of Hcp protein as a candidate vaccine against *A. hydrophila* infec-

tion. In the future, we will prepare recombinant Hcp protein of *A. hydrophila* to further study its immunogenicity.

4.2 Conclusions In this study, the *hcp* gene was successfully cloned from *A. hydrophila* and its bioinformatics analysis was performed. The results showed that the hcp gene had a total length of 1 650 bp and encoded 549 amino acids, with molecular formula $C_{2\,681}H_{4\,196}N_{696}O_{783}S_{24}$, theoretical molecular weight 59 476.44 kDa, and theoretical pI 5.00. The instability coefficient was 20.50, indicating the protein was stable. Neither obvious signal peptide cleavage site nor signal peptide was found, and the protein had no transmembrane region. It might have close genetic relationship with *A. veronii* due to high homology. In the secondary structure, the α-helix, β-sheet, random coil and extended strand accounted for 45.36%, 6.01%, 37.52% and 11.11%, respectively. The tertiary structure model consisted of 18 α-helix and 22 β-sheet. Through the above bioinformatics prediction results, the basic information of *hcp* gene of *A. hydrophila* is preliminarily understood, and the possible function of this protein is predicted, in order to provide guidance for subsequent vaccine research.

References

[1] WANG NN. Functional characteristics of the type VI secretion system effector Hcp in *Aeromonas hydrophila*[D]. Nanjing: Nanjing Agricultural University, 2018. (in Chinese).

[2] BLEVES S, VIARRE V, SALACHA R, *et al.* Protein secretion systems in *Pseudomonas aeruginosa*: A wealth of pathogenic weapons. Journal of Medical Microbiology, 2010(300): 534-543.

[3] LUO GD. Cloning and bioinformatics analysis of T6SS chaperone protein ClpV from *Aeromonas hydrophila*[D]. Ji'nan: Shandong Agricultural University, 2018. (in Chinese).

[4] RUSSELL AB, PETERSON SB, MOUGOUS JD. Type VI secretion system effectors: Poisons with a purpose[J]. Nature Reviews Microbiology, 2014, 12(2): 137-148.

[5] ASCHTGEN MS, BERNARD CS, DE SB, *et al.* SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative *Escherichia coli*[J]. Journal of Bacteriology, 2008, 190(22): 7523-7531.

[6] FELISBERTORODRIGUES C, DURAND E, ASCHTGEN MS, *et al.* Towards a structural comprehension of bacterial type VI secretion systems: Characterization of the TssJ-TssM complex of an *Escherichia coli* pathovar[J]. PLoS Pathogens, 2011, 7(11): e1002386.

[7] MA LS, LIN JS, LAI EM. An IcmF family protein, Imp LM, is an integral inner membrane proteirinteracting with ImpKL, and its Walker A motif is required for type VI secretionsystem-mediated Hcp secretion in *Agrobacterium tumefaciens*[J]. Journal of Bacteriology, 2009, 191(13): 4316-4329.

[8] LEROUX M, LEON JA D, KUWADA N J, *et al.* Quantitative singlecell characterization of bacterialinteractions reveals type VI secretion is a double-edged sword[J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(48): 19804-19809.

[9] RUSSELL AB, SINGH P, BRITTNACHER M, *et al.* A widespread bacterial type VI secretion effectorsuperfamily identified using a heuristic approach[J]. Cell Host & Microbe, 2012, 11(5): 538-549.

[10] WANG P. Advances in hemolysin co-regulated protein of bacterial type VI secretion system[J]. Progress In Microbiology and Immunology, 2018, 46(3): 67-71. (in Chinese).

[11] HO BT, DONG TG, MEKALANOS JJ. A view to a kill: The bacterial type VI secretion system[J]. Cell Host & Microbe, 2014, 15(1): 9-21.

[12] WHITNEY JC, BECK CM, GOO YA, *et al.* Genetically distinct pathways guide effector export through the type VI secretion system[J]. Molecular Microbiology, 2014, 92(3): 529-542.

[13] CHEN Y, SHI H, XU WJ, *et al.* Preparation and characterization of single-chain antibodies against *Pseudomonas plecoglossicida* haemolysin co-regulatory protein[J]. Marine Fisheries, 2022, 44(6): 769-779. (in Chinese).

[14] HU YY. Hemolysin-coregulated protein and its related pathogenicity in *Acinetobacter baumannii*[D]. Wenzhou Medical University, 2018. (in Chinese).

[15] BEI L. *RbsB*, *vgrG*, *hcp* gene of type VI secretion system (T6SS) in *Vibrio harveyi*: Expression and functional analysis[D]. Shanghai: Shanghai Ocean University, 2018. (in Chinese).

[16] WANG XL, LI J, LI GY, *et al.* The effect of rpoS on Hcp expression and bactericidal activity in *Vibrio anguillarum* MHK3[J]. Progress in Fishery Sciences, 2021, 42(6): 125-134. (in Chinese).

[17] WANG YQ, WANG XY, ALI F, *et al.* Comparative extracellular proteomics of *Aeromonas hydrophila* reveals iron-regulated secreted proteins as potential vaccine candidates[J]. Frontiers in Immunology, 2019(10): 256.

[18] PAN HY, ZHOU ZJ, DING J, *et al.* Molecular cloning and bioinformatics analysis of T3SS chaperone escort protein VscO from *Vibrio alginolyticus*[J]. Biotechnology Bulletin, 2014(6): 155-61. (in Chinese).

[19] YUAN SB, ZHU AY. Progress on pathogenicity research on *Vibrio alginolyticus* to aquatic products[J]. Journal of Zhejiang Ocean University (Natural Science Edition), 2012, 31(3): 256-64. (in Chinese).

[20] WU NN, KANG C, RONG N, *et al.* Bioinformatics analysis of *Vibrio alginolyticus* TolB protein[J]. Journal of Henan Agricultural Sciences, 2018, 47(11): 134-141. (in Chinese).

[21] WEI C, LI H, YE SG, *et al.* Cloning and recombinant expression of pathogenic bacterium *Edwardsiella ictaluri hcp* gene[J]. Journal of Dalian Ocean University, 2013(5): 424-430. (in Chinese).

[22] SUAREZ G, SIERRA JC, SHA J, *et al.* Molecular characterization of a functional type VI secretion system from a clinical isolate of *Aeromonas hydrophila*[J]. Microbial Pathogenesis, 2008, 44(4): 344-361.

(From page 35)

[20] SHAO C, HU JP, YAN JC, *et al.* Effect of atractyl on proliferation of human colon cancer LS174T cells[J]. Journal of Jiangsu University (Medical Edition), 2016, 26(6): 480-483. (in Chinese).

[21] ZHANG T, LI SM, LI YN, *et al.* Atractylodin induces apoptosis and institutes the migration of A549 lung cancer cells by regulating ROS-mediated signaling pathways[J]. Molecules, 2022, 27(9): 2946.

[22] KOTAWONG K, CHAIJAROENKUL W, MUHAMAD P, *et al.* Cytotoxic activities and effects of atractylodin and β-eudesmol on the cell cycle arrest and apoptosis on cholangiocarcinoma cell line[J]. Journal of Pharmacological Sciences, 2018, 136(2): 51-56.

[23] QU L. LIN X. LIU C, *et al.* Atractylodin attenuates dextran sulfate sodium-induced colitisby alleviating gut microbiota dysbiosis and incubating influmatory response through the MAPK pathway. Frontiers in Pharmacology, 2021(12): 665376.

[24] LIU Z, JI X, CHEN G, *et al.* Atractylodin ameliorates lipopolysaccharide and d-galactosamine-induced acute liver failure via the support of propagation and oxidative stress. International Immunopharmacology, 2019(72): 348-357.

[25] GAO L, ZHANG WH, WANG YW, *et al.* Effect of atractyl on positive inotropic effect of rat heart and its mechanism[J]. Chinese Journal of Applied Physiology, 2020, 36(5): 408-413. (in Chinese).

[26] ZHU HJ, HONG Y, MA YL, *et al.* Protective effect of n-butanol extract of Atractylodis Rhizoma on myocardial ischemia and ischemia-reperfusion injury in rats[J]. Chinese Science and Technology of Traditional Chinese Medicine, 2000, 7(3): 173-174. (in Chinese).