Journal of Microbiology and Biotechnology

Molecular Characterization and Expression of Cadherin-2 and Cadherin-4 in Nile tilapia (*Oreochromis niloticus*) in response to *Streptococcus agalactiae* Stimulus

Xueying Liang[1#], Yusi Zheng[1#], Zemiao Zhang[1], Yinhui Peng[1,2], Honglin Chen[3], Peng Xu[1], Xinzhong Wu[1], Xiaohui Cai[1*]

[1]Guangxi Key Laboratory of Beibu Gulf Marine Biodiversity Conservation, Beibu Gulf University, Qinzhou, 536011, China

[2]College of Fishery, Guangdong Provincial Key Laboratory of Aquatic Animal Disease Control and Healthy culture, Guangdong Ocean University, Zhanjiang, 524088, China

[3]Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Ministry of Education, Nanning Normal University, Nanning, 530001, China

Research Article

Received: 03-Jun-2024,
Manuscript No. JMB-24-137971;
Editor assigned: 05-Jun-2024,
PreQC No. JMB-24-137971 (PQ);
Reviewed: 19-Jun-2024,
QC No. JMB-24-137971;
Revised: 26-Jun-2024,
Manuscript No. JMB-24-137971 (R);
Published: 03-Jul-2024,
DOI: 10.4172/2320-3528.13.2.004

***For Correspondence:**
Dr. Xiaohui Cai, Ph D, Associate Researcher, Guangxi Key Laboratory of Beibu Gulf Marine Biodiversity Conservation, Beibu Gulf University, No.12, Binhai Avenue, Qinzhou, Guangxi 536011, PR China
Tel/fax: +86 777 2807959
E-mail: caixiaohui66@163.com

Citation: Liang X, et al. Molecular Characterization and Expression of Cadherin-2 and Cadherin-4 in Nile tilapia (*Oreochromis niloticus*) in response to *Streptococcus agalactiae* Stimulus. J Microbiol Biotechnol. 2024;13:004

ABSTRACT

Cadherins are a molecular family that is essential for the Ca^{2+} dependent process of cell-cell adhesion. Moreover, cadherins can also act as a receptor to mediate bacteria entering into non-phagocytic cells. In the present study, members of the cadherins family of cytokines, cadherin-2 (OnCdh2) and Cadherin-4 (OnCdh4), were successfully cloned and characterized from the Nile tilapia *(Oreochromis niloticus)*. Their tissue distribution and expression patterns following bacterial were also investigated. The full-length cDNA sequences of OnCdh2 and OnCdh4 contained an open reading frame of 2721 bp and 2802 bp, encoding 906 amino acids and 933 amino acids with a theoretical isoelectric point of 4.74 and 4.73, respectively. Prediction of protein domains showed that OnCdh2 and OnCdh4 both consisted of one cadherin prodomain super family, one cadherin repeat-like domain, three cadherin tandem repeat domain, one cadherin domain, one cadherin cytoplasmic region and a transmembrane domain. Homology comparisons indicated that OnCdh2 and OnCdh4 showed 94.10% and 99.25% identity to the *Astatotilapia calliptera* and a relative low identity of 75.63%-75.97% and 70.35%-74.09% with its mammalian counterparts. Moreover, the residue 16 of cadherin repeat-like domain of OnCdh2 and OnCdh4 is proline, which indicated that it may play a role in mediating bacterial invasion into intestinal epithelial cells. Phylogenetic tree analysis showed that OnCdh2 and OnCdh4 cluster together with other fish OnCdh2 and OnCdh4 molecules. The results of tissue distribution showed that OnCdh2 and OnCdh4 were both ubiquitous in all tissues examined of healthy tilapia with the highest level of expression in heart and brain, respectively. The expression level of OnCdh2 and OnCdh4 were rapidly activated at 1 hour in brain, intestine and spleen after challenged by *Streptococcus agalactiae*. Taken together, the results indicated that OnCdh2 and OnCdh4 might be involved in the process of *Streptococcus agalactiae* invading into Nile tilapia and the immune response of Nile tilapia against bacterial infection.

Keywords: Nile tilapia; Cadherin-2; Cadherin-4; *Streptococcus agalactiae*; Immune response; Bacterial infection; *Astatotilapia calliptera*

INTRODUCTION

Cell-cell adhesion molecule (CAM) is a general term for many molecules that mediate contact and binding between cells or between cells and extracellular matrix (ECM) and is a complex system, in which various mechanisms and factors are involved [1]. Cell-cell adhesion plays essential roles in organogenesis, physical transport, signal transmission, and immunological function in multicellular organisms [2]. According to their structure, these have been classified into at least three major molecular families, the immunoglobulin (Ig) superfamily, the integrin superfamily, and the cadherin family [1]. Cadherins are a molecular family that is essential for the Ca^{2+} dependent process of cell-cell adhesion [3]. Through their homophilic binding interactions, cadherins play a role in cell-sorting mechanisms, conferring adhesion specificities on cells. Moreover, the regulated expression of cadherins also controls cell polarity and tissue morphology [4].

According to the structure and function, cadherin can be divided into the classical cadherins, the protocadherins and desmosomal cadherins [5]. Among them, more than 20 members of the classical cadherin have been identified in human and mediate cell-cell adhesion through their extracellular domain and their cytosolic domains connect to the actin cytoskeleton by binding to catenins [6]. Recently, our lab have screened total of 18 classical gene family members from Nile tilapia genome data by using a series of bioinformatics methods, but their functions have not been studied [7].

The members of the classic cadherin subfamily, cadherin-2 (also called N-cadherin) and cadherin-4 (also called R-cadherin), have been shown to be involved in development of a variety of tissues and organs [8]. Cadherin-2, the first cadherin discovered in the vertebrate nervous system, has been shown to be of critical importance in the early differentiation of the some vertebrate central and peripheral nervous structures [9]. Recent studies confirmed that cadherin-2 plays an essential role in zebrafish cardiovascular development [10].

Cadherin-4, most similar to cadherin-2 in amino acid sequence among the classic cadherins, has been shown to play a role in the formation of specific brain circuits [11,12]. More importantly, several studies have confirmed that cadherin-2 can be as a receptor to mediate bacteria entering into cells. For example, cadherin-2 can bind to adhesion encoded by *Als3* gene of *Candida albicans* to inducing endocytosis in endothelial cells [13]. Cadherin-2 can also be a receptor for adhesion and endocytosis of *Aspergillus fumigatus* by Human Umbilical Vein Endothelial Cells (HUVEC) [14]. Up to now, the key sites of interaction between cadherin-2 and bacterial specific proteins has not been reported.

Fortunately, the residue 16th prolin of cadherin repeat-like domain of cadherin-1 is the key site for binding with internalin of *Listeria monocytogenes* to mediate bacteria entering into intestine epithelial cells [15]. Unfortunately, it was found that the 16th amino acid of cadherin-1 in Nile tilapia (*Oreochromis niloticus*) was not proline through bioinformatics analysis. Whether cadherin-2 and cadherin-4 of Nile tilapia (OnCdh2 and OnCdh4) with a 16th proline have the same function as E-cadherin in the process of bacterial invasion have not been studied.

Nile tilapia is a hardy, most cultured freshwater fish in the world [16]. Due to their hardiness, fecundity, feed efficiency, and rapid growth, tilapia and their hybrids are among the most important group of farm-raised fish for the 21st century in the world [15]. Data from the 2019 China fisheries statistical year book show that China's tilapia production reached 1.62 million tons in 2018, an annual increase of 2.52% [17]. Although resistant to unfavourable water quality, Nile tilapia is susceptible to *Streptococcus agalactiae* [16] and can cause huge economic loss.

The mortality rate of Thai Nile tilapia [18] and Malaysian red tilapia [19] caused by *Streptococcus agalactiae* infection reached up to 30% and 60%~70% [20] in 2013, and there is no effective control [21]. It has been shown that the main entry site of *Streptococcus agalactiae* is the gastrointestinal epithelium [22] and whether the members of the classic cadherin with 16th poline play a role in process of bacterial invasion is unknown. In this study, the cadherin-2 and cadherin-4 was cloned and identified from Nile tilapia.

The mRNA expression pattern of OnCdh2 and OnCdh4 of Nile tilapia different tissues of healthy fish were performed. In addition, the expression level of OnCdh2 and OnCdh4 were also detected in brain, intestine, spleen and head kidney after challenged by *Streptococcus agalactiae* through oral route. The results will help us to understand the process of *Streptococcus agalactiae* invading into Nile tilapia and the immune response of Nile tilapia against bacterial infection.

MATERIALS AND METHODS

Fish and bacterial challenge

Healthy Nile tilapia (50 g ± 7 g body weight) were obtained from a commercial Nile tilapia farm in Qinzhou, Guangxi, China. Fish were maintained in an aquarium at 28ºC~30ºC and subjected to 12 hours' light: 12 hours' dark cycles at 28ºC for 7 days. Fish were fed twice daily with commercial fish expanded pellets and waste removed daily.

To detect the gene expression level of different tissue, samples of ten tissues (eyes, brain, liver, heart, muscle, stomach, spleen, gill, head kidney and intestine) were sampled from healthy Nile tilapia, and immediately frozen in RNA hold for storage at -80ºC until use. All experiments were conducted under a protocol approved by the Institutional Animal Care and Use Committee (IACUC, China).

Streptococcus agalactiae TOS01, isolated from golden pompano with a high virulence, was used in this study [23]. The *Streptococcus agalactiae* was diluted by Phosphate-Buffered Saline (PBS) with a final concentration of 1×10^7 CFU mL^{-1}. The challenge group was inoculated intragastrically with 0.1 mL of *Streptococcus agalactiae* [22].

Tissue samples (brain, gills, liver, spleen and head kidney) were collected at nine time points (0 hour, 1 hour, 3 hours, 6 hours, 12 hours, 24 hours, 48 hours, 72 hours, 96 hours) after stimulation with 0 hour as control. All the primers used in this study were showed in Table 1.

Table 1. Primers used for OnCdh2 and OnCdh4 cloning and expression detection.

Primer name	Sequence 5'-3'	Purpose
Cadherin-2F	TACATGAGGAATGTGCTCCCCGTCG	ORF cloning
Cadherin-2R	GTACATGCCACCTTGCACTGCTGATC	ORF cloning
Cadherin-4F	ATGAGAACTGACTTCGCTCTGGTC	ORF cloning
Cadherin-4R	TTAATCATCATCTCCTCCGTACATG	ORF cloning
Cadherin-2qF	ATGTACTCCTTACACGAGGGCAGC	ORF cloning
Cadherin-2qR	CTAGTCGTCACTTCCACCGTACATG	Real-time PCR
Cadherin-4qF	GCCGGTAGACTTTGAGATGAAC	Real-time PCR
Cadherin-4qR	GTTTTGGGTTTGTGGGGAAGTA	Real-time PCR

Cloning of OnCdh2 and OnCdh4

Total RNA was extracted from the gill of healthy Nile tilapia with Trizol reagent (TransGen, Beijing, China) and then reverse-transcribed into first-strand cDNA using the TransScript® One-Step gDNA removal and cDNA synthesis super mix kit (TransGen, Beijing, China) according to the manufacturers protocol. The specific primers for the full-length Open Reading Frame (ORF) of OnCdh2 and OnCdh4 were designed based on the genome data of Nile tilapia (https://www.ncbi.nlm.nih.gov/genome/197?genome_assembly_id=391053) in GenBank (Table 1). The PCR products were purified and ligated into a pMD18-T Vector (TaKaRa, Japan) and then transformed into competent *Escherichia coli* DH5α cells. Then, the positive clones were sequence by ThermoFisher (Guangzhou) Co. Ltd. The full length cDNA sequences were obtained by assembling the sequences using the software DNAMAN V6 (Lynnon Biosoft, USA).

Bioinformatics analysis of OnCdh2 and OnCdh4

The homology searches of the full length cDNA sequence of OnCdh2 and OnCdh4 gene was analyzed using Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) program. The molecular weights and pI of OnCdh2 and OnCdh4 was calculated with the Expasy compute pI/MW tool (https://web.expasy.org/protparam
//), and the signal peptides of OnCdh2 and OnCdh4 were predicted by SignalP 5.0 Server (http://www.cbs.dtu.dk/s
ervices/SignalP/). The functional conserved domains of OnCdh2 and OnCdh4 were presumed using CDD program (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Multiple sequence alignments were performed using Clustal X 2.0 program and visualized using GeneDoc software. Phylogenetic analyses were performed using the neighbor-joining method in the MEGA 7.0 with a minimum of 1000 bootstraps replications.

Quantitative analysis of OnCdh2 and OnCdh4 mRNA

The quantitative real-time PCR analysis of OnCdh2 and OnCdh4 mRNA expression in pre-stimulation and post-stimulation tissues were performed on the QuantStudio™ 6 Flex Real-Time PCR System (Thermo Fisher, USA) first-strand cDNA was synthesized from total RNA as described in section 2.2. Real-time PCR was performed with the specific primers designed based on the full-length sequence of OnCdh2 and OnCdh4, and the β-actin gene was used as internal control in all RT-qPCR experiments (Table 1). Each reaction contains 1 μL of reverse and forward primers, 10 μL of SYBR Green Mix (TransGen, Beijing, China), and 2 μL of 1:8 diluted cDNA and RNase-free water to a final reaction volume of 20 μL. The program was performed as follows: 1 cycle of 10 minutes at 95ºC, 40 cycles of 10 seconds at 95ºC, 15 seconds at 55ºC and 15 seconds at 72ºC. The relative expression level of OnCdh2 and OnCdh4 mRNA were calculated using $2^{-\Delta\Delta Ct}$ method [24]. Each experiment was performed in triplicate.

Statistical analysis

All data in this study were displayed as means ± standard deviation (SDs). Statistical analysis performed by the LSD (Least Significant Difference) test using SPSS 17.0 software. Combined with the information of significance level and mean value, the significant difference was marked by the method of "a b c". There was exactly the same amount that is, a level of significant difference between the two groups.

RESULTS

Sequence analysis of OnCdh2 and OnCdh4

The ORF of OnCdh2 (GenBank: MN641108) was 2721 bp, encoding a putative protein of 906 amino acids with a predicated molecular mass of 100.06 kDa and theoretical isoelectric point of 4.74 (Figure 1A), A putative signal peptide of 23 amino acids at the N-terminus was predicted using the SignalP 5.0 program. A transmembrane domain was found between 724 amino acids and 746 amino acids by using TMHMM Server 2.0. Analysis of physicochemical properties showed that the content of proline (Pro) in the amino acid sequence of OnCdh2 was the highest

(8.3%), followed by aspartic acid (Asp) (8.1%), and the lowest was cysteine (Cys) and tryptophan (Trp) (0.9%). Prediction of protein domains showed that OnCdh2 consisted of one cadherin prodomain super family (30 aa-121 aa), one cadherin repeat-like domain (163 aa-263 aa), three cadherin tandem repeat domain (276 aa-378aa, 386 aa-493 aa, 503 aa-601 aa), one cadherin domain (610 aa-701 aa) and one cadherin cytoplasmic region (751 aa-904 aa) (Figure 1A).

Figure 1A. The predicted amino acid sequences of OnCdh2.

Note: ■ Signal peptides; — Cadherin prodomain super family; — Cadherin repeat-like domain; ═ Cadherin tandem repeat domain; ≈ Cadherin domain; ■ Cadherin cytoplasmic region; ∿ Cadherin transmembrane domain.

A

```
1     ATGTACTCCTTACACGAGGGCAGCGGTCTGCTTCTGCTGCTGGTCTTGGCAGGACTGCAGATTGCAGCAGAGAGCAGGGGTGCCCCACCATGCAGGCCAGGC
1     M  Y  S  L  H  E  G  S  G  L  L  L  L  L  V  L  A  G  L  Q  I  A  A  E  S  R  G  A  P  P  C  R  P  G
103   TTCTCTGAAGATGTATACAAGGTTGTGGTGCCAGATATCACCGAGAAAGGACAGCCCCTTTTTAACGTGAGGTTTGTCAGCTGTGGCCGTGGTGGTCTGATG
35    F  S  E  D  V  Y  K  V  V  V  P  D  I  T  E  K  G  Q  P  L  F  N  V  R  F  V  S  C  G  R  G  G  L  M
205   CGGTTTGAATGCAGCAACCCATCTGACTTTCGGATAGAGGCCGACGGAGTCCTTTATGCTGCCAGGACTCTTCAGCGCTCCACACTCCCTGTGTCACCTTTG
69    R  F  E  C  S  N  P  S  D  F  R  I  E  A  D  G  V  L  Y  A  A  R  T  L  Q  R  S  T  L  P  V  S  P  L
307   CTGATCACGGCGAGTGACACCGCCACTCAGCAGCAATGGGTGGCCCAAGTCAGACTGGTGCCCAGGCAGCAGGTGCAGGAAGGCGCTGTCCCTCCAGTCATG
103   L  I  T  A  S  D  T  A  T  Q  Q  Q  W  V  A  Q  V  R  L  V  P  R  Q  Q  V  Q  E  G  A  V  P  P  V  M
409   TCCAAGGATTTAAAGGAAATAGCATTTCCACCACGAAACATGGACAGCCAGCTGAAACGCATGAAGAGAGACTGGGTCATCCCCCCAATCAATGTCCCAGAG
137   S  K  D  L  K  E  I  A  F  P  P  R  N  M  D  S  Q  L  K  R  M  K  R  D  W  V  I  P  P  I  N  V  P  E
511   AACTCACGAGGCCCATTCCCACAAGAGCTTGTCCGGATCCGATCAGATCATGATAAAAACCGATCCCTGCGATACAGCGTGACAGGCCCTGGTGCTGACCAG
171   N  S  R  G  P  F  P  Q  E  L  V  R  I  R  S  D  H  D  K  N  R  S  L  R  Y  S  V  T  G  P  G  A  D  Q
613   CCTCCCACTGGAATCTTCATCATCGACCCAATTTCCGGAGAGCTGTCTGTCAACAAGCCCCTTGATCGGGAGCACATCCCCAACTTCCATCTGAGAGCTCAT
205   P  P  T  G  I  F  I  I  D  P  I  S  G  E  L  S  V  N  K  P  L  D  R  E  H  I  P  N  F  H  L  R  A  H
715   GCAGTAGACCTGAATGGAAACCAAGTTGAAAACCCTATTGATATCGTGATCAATGTGATCGACATGAATGACAACAGACCCGAGTTCACCCATCAGATCTGG
239   A  V  D  L  N  G  N  Q  V  E  N  P  I  D  I  V  I  N  V  I  D  M  N  D  N  R  P  E  F  T  H  Q  I  W
817   AATGGCACCGTTCCAGAGGGTTCCAAGCCAGGATCATTTGTTATGACAGTAACTGCTGTTGACAAAGATGACCCCAAAACAGCTAATGGGATGCTACGTTAC
273   N  G  T  V  P  E  G  S  K  P  G  S  F  V  M  T  V  T  A  V  D  K  D  D  P  K  T  A  N  G  M  L  R  Y
919   AAGATCTTGTCTCAGAATCCCGAGAGTCCAACCTCCAATATGTTCACCATCAACAACAAAACCGGAGGCATCATTACAGTGGCAGCAGGCCTGGACAGAGAG
307   K  I  L  S  Q  N  P  E  S  P  T  S  N  M  F  T  I  N  N  K  T  G  G  I  I  T  V  A  A  G  L  D  R  E
1021  AAAGTGCCACAGTACACACTGATAATCCAGGCCAGTGACATGGAGGGCAATCCCATGTATGGCCTCTCTAACACGGCCACGGCTGTCATCAGAGTCACTGAC
341   K  V  P  Q  Y  T  L  I  I  Q  A  S  D  M  E  G  N  P  M  Y  G  L  S  N  T  A  T  A  V  I  R  V  T  D
1123  ATTAATGACAACCCGCCAGAGTTTACTGCTGAGACGTTTTTCGGCGAGGTGCCTGAAAACCGCGTGAATGTCATTGTGGCTAACCTGACAGTGACTGACAAG
375   I  N  D  N  P  P  E  F  T  A  E  T  F  F  G  E  V  P  E  N  R  V  N  V  I  V  A  N  L  T  V  T  D  K
1225  GACCAGCCCAACACGCTGGCCTGGAATGCTGTCTACAAAATCACCGGGGGCGACCCCACCGGCAGGTTCTCCGTGCCCACTGATCCAACCACTAATGAGGGC
409   D  Q  P  N  T  L  A  W  N  A  V  Y  K  I  T  G  G  D  P  T  G  R  F  S  V  P  T  D  P  T  T  N  E  G
1327  TTGGTAACAGTCGTCAAGCCTATTGATTATGAAACCAGCAGGACTTATGTGCTGACAGTGGAAGCGAGGAATGAAGTTCCCCTGGCCAAAGGCATCCACACT
443   L  V  T  V  V  K  P  I  D  Y  E  T  S  R  T  Y  V  L  T  V  E  A  R  N  E  V  P  L  A  K  G  I  H  T
1429  CCCCGACAGTCCACTGCCACCGTCTCCATCAGAGTGATAGATGTCAATGAGAGTCCCTACTTTGAGCCCAACCCCAAATTCATTAGGCTTGAAGAGGGAATG
477   P  R  Q  S  T  A  T  V  S  I  R  V  I  D  V  N  E  S  P  Y  F  E  P  N  P  K  F  I  R  L  E  E  G  M
1531  ATGCCAGACTCAACACTGACCACTTTCAGTGCACAGGATCCTGATCGCTTCATGCAGCAAACCATCAGGTACTCCAAACTCTCAGATCCAGCCAACTGGTTA
511   M  P  D  S  T  L  T  T  F  S  A  Q  D  P  D  R  F  M  Q  Q  T  I  R  Y  S  K  L  S  D  P  A  N  W  L
1633  AGGATCGACACAAACACCGGCCGGATCACAACGATTGCCATTTTGGACCGAGAATCACCTTATGTGAAGAACAATCTGTACAATGCAACATTCCTGGCTACT
545   R  I  D  T  N  T  G  R  I  T  T  I  A  I  L  D  R  E  S  P  Y  V  K  N  N  L  Y  N  A  T  F  L  A  T
1735  GACAATGGTATACCGCCAGCAAGTGGCACAGGAACTCTCCAAATCTTCCTACTGGATATCAATGACAATGCGCCCCGGGTGTCCCCTCAGGAGGTGGAAATA
579   D  N  G  I  P  P  A  S  G  T  G  T  L  Q  I  F  L  L  D  I  N  D  N  A  P  R  V  S  P  Q  E  V  E  I
1837  TGTGAAAAGCCAGACCCGAATGCCATCAACATCACTGCATTTGATGGAGACCTGAGCCCAAATGCTGGGCCATTTGCCTTTGAGTTAGCCAACAGGCCCGCT
613   C  E  K  P  D  P  N  A  I  N  I  T  A  F  D  G  D  L  S  P  N  A  G  P  F  A  F  E  L  A  N  R  P  A
1939  GATGTTCGGAGAAACTGGACCATAACAAGAATCAGCGGTGACTATGCTCAGGTGAGCCTGAAGATTGGCTTCCTTGAAAGTGGTATCTATGAGATCCCCATC
647   D  V  R  R  N  W  T  I  T  R  I  S  G  D  Y  A  Q  V  S  L  K  I  G  F  L  E  S  G  I  Y  E  I  P  I
2041  ATAATCACGGACTCAGGCAACCTGCCCATGTCTAACACCTCCTACCTGCGGATTAAAGTGTGCCAGTGCGACATCAACGGAGACTGCACTGACCAGCAGCAC
681   I  I  T  D  S  G  N  L  P  M  S  N  T  S  Y  L  R  I  K  V  C  Q  C  D  I  N  G  D  C  T  D  Q  Q  H
2143  ATCATGGCCGCCGGCCTAGGCACAGGCGCCATCATATCCATTCTGCTCTGCATCATCATCCTCCTCATTCTCGTGCTGATGTTTGTGGTGTGGATGAAGCGT
715   I  M  A  A  G  L  G  T  G  A  I  I  S  I  L  L  C  I  I  I  L  L  I  L  V  L  M  F  V  V  W  M  K  R
2245  CGTGATAAAGAACGTCAGGCTAAGCAGCTCCTTATCGATCCTGAGGATGACGTGAGAGACAACATCCTGAAGTATGATGAAGAAGGTGGAGGAGAAGAGGAT
749   R  D  K  E  R  Q  A  K  Q  L  L  I  D  P  E  D  D  V  R  D  N  I  L  K  Y  D  E  E  G  G  G  E  E  D
2347  CAGGATTACGACCTGAGTCAGCTCCAACAGCCGGACACGATAGAGCCCGATGCCATCAAGCCGGTGGGAATCCGCCGTCTAGACGAGAGGCCCCTCCACCCT
783   Q  D  Y  D  L  S  Q  L  Q  Q  P  D  T  I  E  P  D  A  I  K  P  V  G  I  R  R  L  D  E  R  P  L  H  P
2449  GAGCCACAGTACCCAATGAGGTCGGCAGCACCCCACCCCGGAGACATCGGAGACTTCATCAATGAGGGACTCAAGGCAGCAGACAATGACCCCACCGCTCCT
817   E  P  Q  Y  P  M  R  S  A  A  P  H  P  G  D  I  G  D  F  I  N  E  G  L  K  A  A  D  N  D  P  T  A  P
2551  CCCTACGACTCCCTGCTAGTGTTCGACTACGAGGGCAGCGGCTCCACAGCTGGCTCCCTAAGCTCCCTCAACTCCTCCAGCAGCGGGGGTGACCAGGACTAC
851   P  Y  D  S  L  L  V  F  D  Y  E  G  S  G  S  T  A  G  S  L  S  S  L  N  S  S  S  S  G  G  D  Q  D  Y
2653  GATTACCTCAACGACTGGGGCCCTCGCTTTAGAAAACTGGCAGACATGTACGGTGGAAGTGACGACTAG
885   D  Y  L  N  D  W  G  P  R  F  R  K  L  A  D  M  Y  G  G  S  D  D  *
```

Figure 1B. The predicted amino acid sequences of OnCdh4.

Note: Signal peptides; Cadherin prodomain super family; Cadherin repeat-like domain; Cadherin tandem repeat domain; Cadherin domain; Cadherin cytoplasmic region; Cadherin transmembrane domain.

B

```
1    ATGAGAACTGACTTCGCTCTGGTCCTGACAGCAGCCCTTCTTCTTGACACATTCTGGAGGAATAATGGGGAGCACGCTTCCAAAGGTCCCTGCCAACCTGGC
1    M  R  T  D  F  A  L  V  L  T  A  A  L  L  L  D  T  F  W  R  N  N  G  E  H  A  S  K  G  P  C  Q  P  G
103  TTCTCACAAAGCTTCTACTCTGTGCTGATTTCAAGGGATGTACTGCAAGGTCAGGGCATACTTAAAGATACTACTGTCTCATCCCAGAAAGTCAAGTTTGAT
35   F  S  Q  S  F  Y  S  V  L  I  S  R  D  V  L  Q  G  Q  G  I  L  K  D  T  T  V  S  S  Q  K  V  K  F  D
205  GACTGCATGGGAAATGAAGGTGTGGTCTTCCAGAGCAGCAACCCCCTCTTTGGCGTGCGGACAGATGGATCCATCTTTGCCCAACGAGAGGGAGCCAACCTA
69   D  C  M  G  N  E  G  V  V  F  Q  S  S  N  P  L  F  G  V  R  T  D  G  S  I  F  A  Q  R  E  G  A  N  L
307  GATGAACCTGTCCAGTTGAAACTGACAGCAAGTGGTCCGCACACACATGTCTGGGAGACTATGGTGCAGCTGGCACTGATTGACCTACCATCACCCCAACAA
103  D  E  P  V  Q  L  K  L  T  A  S  G  P  H  T  H  V  W  E  T  M  V  Q  L  A  L  I  D  L  P  S  P  Q  Q
409  ACTGAAAATGAGGTCACAGAGAATGAAAAGAAAGTTGCTGGGAGAAAAGAGAAGAAGGAGGCGGTACAAATACCACCTCCTGTCATCATGTTCCCAGGGTCC
137  T  E  N  E  V  T  E  N  E  K  K  V  A  G  R  K  E  K  K  E  A  V  Q  I  P  P  P  V  I  M  F  P  G  S
511  GGTCACAACATCACCAAGGGTCTACGAAGACAGAAGAGGGACTGGGTAATCCCACCAATCAATGTGGCTGAAAATTCTCGAGGACCGTTCCCTCAGATGCTT
171  G  H  N  I  T  K  G  L  R  R  Q  K  R  D  W  V  I  P  P  I  N  V  A  E  N  S  R  G  P  F  P  Q  M  L
613  GTCAGCATTCGTTCCGACCAGGACCAGGACATCTCAATCCGCTATAGCATCACTGGAGTGGGAGCAGATCAACCACCGAATGAAGTATTCAGTATCGACCCT
205  V  S  I  R  S  D  Q  D  Q  D  I  S  I  R  Y  S  I  T  G  V  G  A  D  Q  P  P  N  E  V  F  S  I  D  P
715  GTGCTTGGAAAGATGTTTGTCACCAAACCTCTTGATCGAGAACATCGCTCGTCCTATCACCTGCGAGCCCATGCTGTGGACATGAACGGCAACCAGGTTGAG
239  V  L  G  K  M  F  V  T  K  P  L  D  R  E  H  R  S  S  Y  H  L  R  A  H  A  V  D  M  N  G  N  Q  V  E
817  AACCCTATCGATCTGTACATCTTTGTCATCGATATGAATGACAACAGACCCGAGTTCAAAAACCAGGTCTACAATGGCTCTGTGGATGAAGGCTCGAAACCA
273  N  P  I  D  L  Y  I  F  V  I  D  M  N  D  N  R  P  E  F  K  N  Q  V  Y  N  G  S  V  D  E  G  S  K  P
919  GGAACATATGTGATGCATCTGTCAGCATTTGATGCAGATGACAACACCACAGCTAATGGAATGGTGCGCTACCGGATCCTGTCCCAGACACCGCACAGCCCG
307  G  T  Y  V  M  H  L  S  A  F  D  A  D  D  N  T  T  A  N  G  M  V  R  Y  R  I  L  S  Q  T  P  H  S  P
1021 ATCCCTAACATGTTTACCATCAACAGCGAGACTGGAGACATTGTGACAGTGGCGCCTGGTCTTGATCGAGAGAAAGTGTCCCAGTATACCATTATCGTGCAG
341  I  P  N  M  F  T  I  N  S  E  T  G  D  I  V  T  V  A  P  G  L  D  R  E  K  V  S  Q  Y  T  I  I  V  Q
1123 GCCACGGACATGGAGGGCAACCTGAACTTTGGCCTTTCCAACACAGCCACTGCCATCATCACCGTCACGGACATCAATGACAACCCTCCCATGCTGACCTCC
375  A  T  D  M  E  G  N  L  N  F  G  L  S  N  T  A  T  A  I  I  T  V  T  D  I  N  D  N  P  P  M  L  T  S
1225 AGAACTTTTTCTGGAGAGGTCCCCGAGAACCGAGTAGATGTCGTTGTGGCTAACCTGACGGTTATAGATGCTGATCAACCGCACTCCCCTAATTGGAATGCC
409  R  T  F  S  G  E  V  P  E  N  R  V  D  V  V  V  A  N  L  T  V  I  D  A  D  Q  P  H  S  P  N  W  N  A
1327 ATCTACAGGATCATCAGCGGAGACCCATCTGGGCACTTTACCATCCGCACCGACCCAATTTCAAATGATGGCATGGTCACCGTGGTGAAGCCGGTAGACTTT
443  I  Y  R  I  I  S  G  D  P  S  G  H  F  T  I  R  T  D  P  I  S  N  D  G  M  V  T  V  V  K  P  V  D  F
1429 GAGATGAACCGTGCCTTCATGCTGACGGTAGTGGTGTCCAACCAGGCCCACCTGGCCACTGGCATCCAGTCCTCACTCCAGTCCACGGCTGGGGTCACCATA
477  E  M  N  R  A  F  M  L  T  V  V  V  S  N  Q  A  H  L  A  T  G  I  Q  S  S  L  Q  S  T  A  G  V  T  I
1531 TCTGTGCAGGATGTAAACGAACCTCCTTACTTCCCCACAAACCCAAAACACATCCGCTTTGAGGAAGGTGTCCCTGCTGGCACCAGCCTGACGACGTTCTCG
511  S  V  Q  D  V  N  E  P  P  Y  F  P  T  N  P  K  H  I  R  F  E  E  G  V  P  A  G  T  S  L  T  T  F  S
1633 GCGTCAGATCCCGACCGACTTCAGATGCAGCAGATAATCAGGTATTCAAAGCTGAACGACCCAGCAGACTGGCTGTCCATCAATAGAAGCAACGGTCAGATA
545  A  S  D  P  D  R  L  Q  M  Q  Q  I  I  R  Y  S  K  L  N  D  P  A  D  W  L  S  I  N  R  S  N  G  Q  I
1735 GTAACCACAGCAATGCTCGACCGGGAGTCTGTCTATGTAAAAAACAACATATATGAAGCCACATTTCTGGCAACAGACAATGGCTCTCCTGCAGCCAGCGGC
579  V  T  T  A  M  L  D  R  E  S  V  Y  V  K  N  N  I  Y  E  A  T  F  L  A  T  D  N  G  S  P  A  A  S  G
1837 ACAGGCACACTGCAGATCTATCTGATAGATATTAATGACAACCCCCCAGCACTCATACCCAAAGAGTCTCAGATCTGTGAAAGGACGTACAGGAGCATCAAC
613  T  G  T  L  Q  I  Y  L  I  D  I  N  D  N  P  P  A  L  I  P  K  E  S  Q  I  C  E  R  T  Y  R  S  I  N
1939 GGTGTCAACATCACTGCCGGTGATGCCGACACAGACCCCAACGCTGGACCTTTTGTCTTCGAGTTGCCCACGTTTCCCACCACCATCCGACGTAACTGGACA
647  G  V  N  I  T  A  G  D  A  D  T  D  P  N  A  G  P  F  V  F  E  L  P  T  F  P  T  T  I  R  R  N  W  T
2041 ATTTCTAGAATCAGCGGTGATTTTGCCCGTTTGAGTCTGCGGTATCAAATATATTTAGAGCCTGGAATGTATGAAGTCCCTGTAATCGTGTCAGACTCCGGG
681  I  S  R  I  S  G  D  F  A  R  L  S  L  R  Y  Q  I  Y  L  E  P  G  M  Y  E  V  P  V  I  V  S  D  S  G
2143 AACCCTCCATTGTCCAACAGGTCAGTCATCAAAGTTAAGGTGTGTCCCTGCGATGAAAACGGAGATTGCACCACCCTCGGGGCCGTGGCAGCTGCTGGCCTC
715  N  P  P  L  S  N  R  S  V  I  K  V  K  V  C  P  C  D  E  N  G  D  C  T  T  L  G  A  V  A  A  A  G  L
2245 GGAACTGGAGCCATCATCTCCATCCTCATCTGCATCATCATACTACTCAGCATGGTGCTGTTGTTTGTGGTGTGGATTAAGCGCAGAGAGAAGGAGAGGCAG
749  G  T  G  A  I  I  S  I  L  I  C  I  I  I  I  L  L  S  M  V  L  L  F  V  V  W  I  K  R  R  E  K  E  R  Q
2347 ACAAAGCAACTGCTCATTGACCCTGAAGATGATGTCAGGGACAACATCCTCAAGTATGATGAAGAGGGAGGAGGAGAAGAGGACCAGGACTATGACTTAAGC
783  T  K  Q  L  L  I  D  P  E  D  D  V  R  D  N  I  L  K  Y  D  E  E  G  G  G  E  E  D  Q  D  Y  D  L  S
2449 CAGCTGCAGCAGCCCGACTCCTTGGAGCACATCATCAGCAAGCCGCCTGGTGTCCGCAGGGTGGATGAGCGTCCGATCATCCCAGAGTCCCAGTATCCAATC
817  Q  L  Q  Q  P  D  S  L  E  H  I  I  S  K  P  P  G  V  R  R  V  D  E  R  P  I  I  P  E  S  Q  Y  P  I
2551 AGACCTGTTGTTCCCCATCCCGGAGACATTGGGGACTTCATCCATGAGGGACTGAGAGCAGCAGACAACGATCCCACAGCTCCTCCATACGACTCCTTGCTG
851  R  P  V  V  P  H  P  G  D  I  G  D  F  I  H  E  G  L  R  A  A  D  N  D  P  T  A  P  P  Y  D  S  L  L
2653 GTGTTTGACTACGAGGGTAGCGGCTCCACCGCTGGCTCTGTCAGCTCCCTCAACTCCTCCAGCACAGGAGACCAGGACTATGACTACCTCAATGACTGGGGC
885  V  F  D  Y  E  G  S  G  S  T  A  G  S  V  S  S  L  N  S  S  S  T  G  D  Q  D  Y  D  Y  L  N  D  W  G
2755 CCCCGATTTAAGAAGCTGGCTGACATGTACGGAGGAGATGATGATTAA
919  P  R  F  K  K  L  A  D  M  Y  G  G  D  D  D  *
```

The ORF of OnCdh4 (GenBank: MN641109) was 2802 bp, encoding a putative protein of 933 amino acids with a predicated molecular mass of 102.80 kDa and theoretical isoelectric point of 4.73 (Figure 1B). Similar with OnCdh2, a signal peptide of 23 amino acids at the N-terminus and a transmembrane domain between 752 amino acids and 774 amino acids was identified using the SignalP 5.0 program and TMHMM Server 2.0 (Figure 1B). Analysis of physicochemical properties showed that the content of aspartic acid (Asp) in the amino acid sequence of OnCdh4 was the highest (8.1%), followed by serine (Ser) and valine (Val) (7.8%), and the lowest was cysteine (Cys) (0.8%). Prediction of protein domains showed that OnCdh2 also contained one cadherin pro-domain super family (30 aa-117 aa), one cadherin repeat-like domain (187 aa-287 aa), three cadherin tandem repeat domain (295 aa-402 aa, 410 aa-517 aa, 532 aa-626 aa), one cadherin domain (636 aa-729 aa) and one cadherin cytoplasmic region (775 aa-931 aa) (Figure 1B).

Homology alignment and phylogenetic analysis

Homology comparisons indicated that OnCdh2 shared the highest sequence identity of 94.10% with *Astatotilapia calliptera Cdh2*, followed by 75.97% identity with *Bos taurus* and a relative low identity of 75.63% with *Home sapiens* and *Mus musculus Cdh2* (Figure 2A). OnCdh4 showed 99.25% identity to the *Maylandia zebra Cdh4* and a relative low identity of 70.35%-74.09% with its mammalian counterparts (Figure 2B). In addition, the residue 16 of cadherin repeat-like domain is proline, which is as a residue critical for host specificity towards the human pathogen *Listeria* monocytogene [25].

Figure 2A. Multiple alignment of OnCdh2 with other *Cdh2* in other species. Sequences are aligned using clustalx 2.0 and GenoDoc program.

Note: ■ Same amino acid; ■ Similar amino acid; | Residue 16 of cadherin repeat-like domain; *Oreochromis niloticus* (MN641108); *Maylandia zebra* (XP_004561628); *Homo sapiens* (NP_001783); *Mus musculus* (EDK96930); *Bos taurus* (NP_001159964).

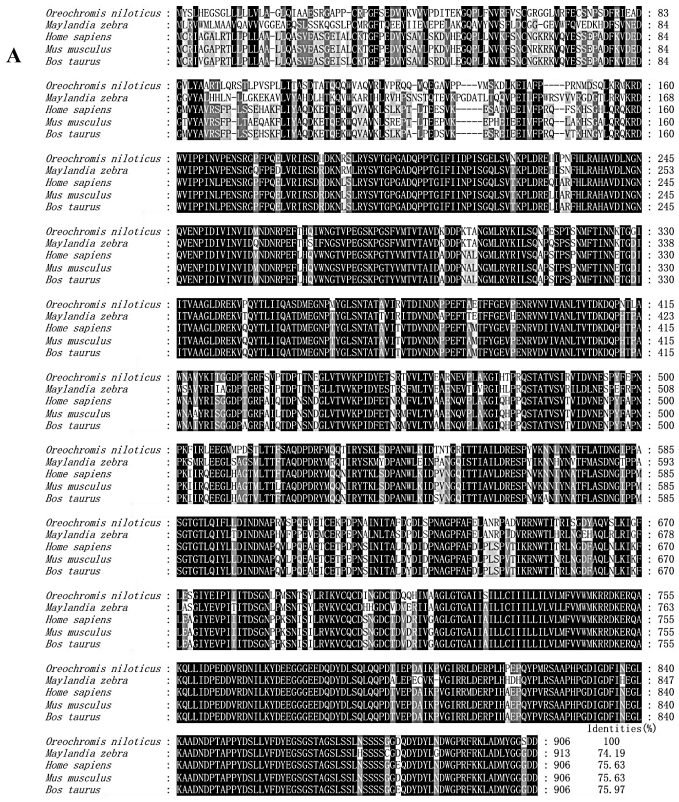
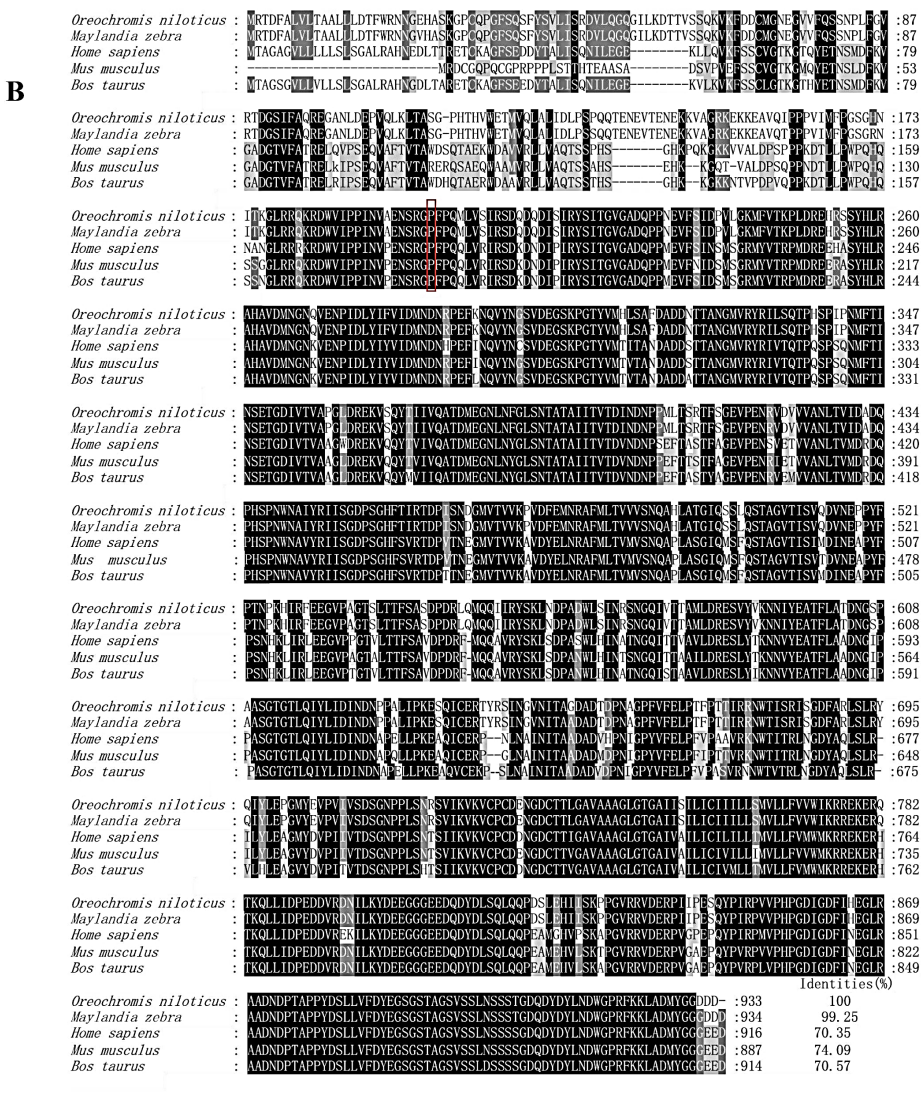


Figure 2B. Multiple alignment of OnCdh4 with other *Cdh4* in other species. Sequences are aligned using clustalx 2.0 and GenoDoc program.

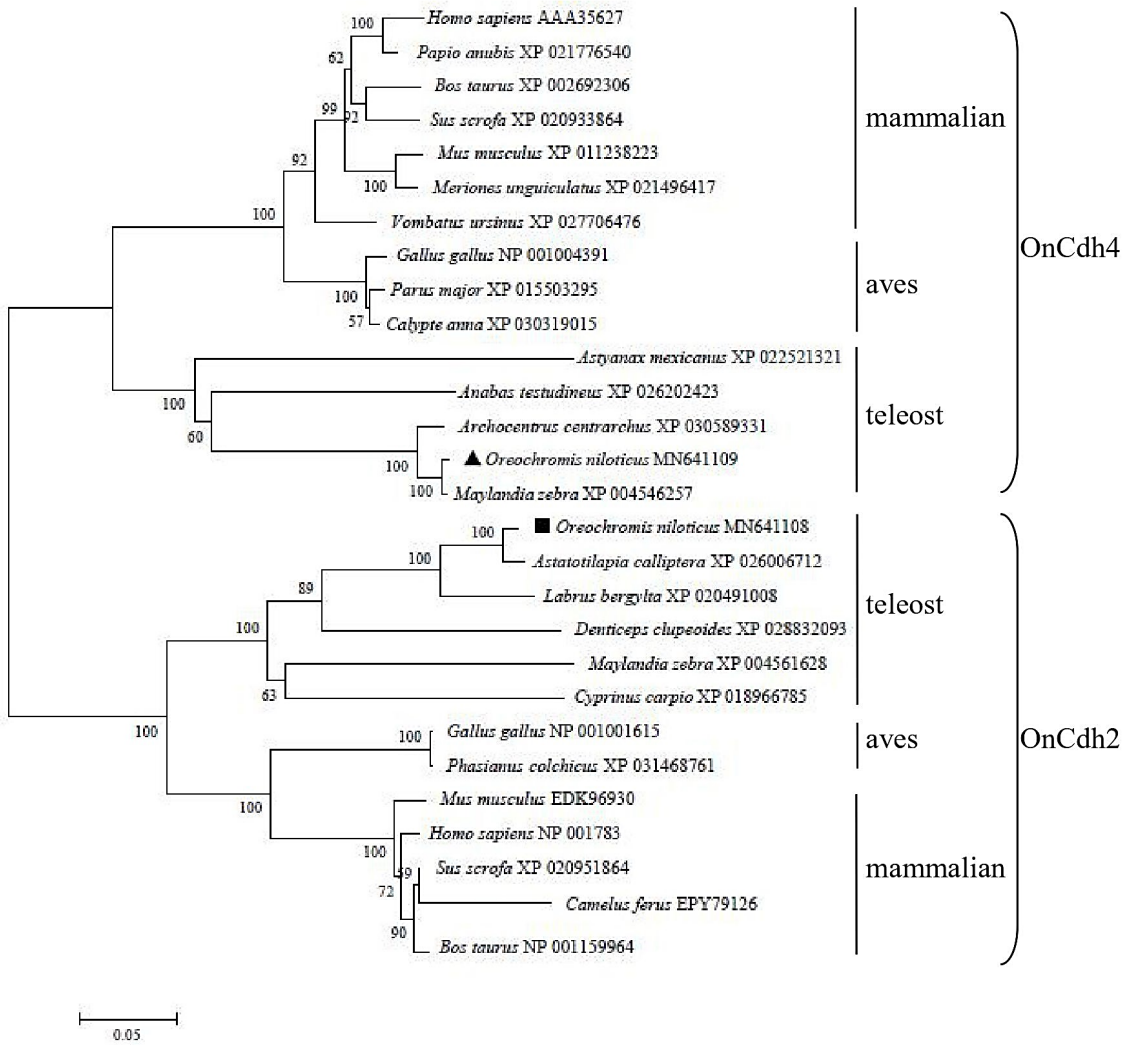
Note: ■ Same amino acid; ■ Similar amino acid; | Residue 16 of cadherin repeat-like domain; *Oreochromis niloticus* (MN641108); *Maylandia zebra* (XP_004561628); *Homo sapiens* (NP_001783); *Mus musculus* (EDK96930); *Bos taurus* (NP_001159964).



To determine the evolutionary relationship between Nile tilapia OnCdh2 and OnCdh4 and that of other species, a phylogenetic tree was constructed using neighbor-joining method with a bootstrap of 1000 replications. As shown in Figure 3, all the Cdh2 and Cdh4 proteins used in this study were grouped into the *Aves*, *Mammalia* and *Teleostei* clade. OnCdh2 and OnCdh4 was clustered with *Astatotilapia calliptera Cdh2* and *Maylandia zebra Cdh4* together, and forming a separate teleost clade with that of other sorts of fish, respectively.

Figure 3. Phylogenetic tree analysis of the full-length amino acid sequences of Oncdh2 and OnCdh4 from various species. A phylogenetic tree is constructed by neighbor-joining method using MEGA 7.0 with the bootstrap values of 1,000 replicates. Based on the protein sequence, the relations of different organisms are shown by dendrogram graphically. The scale bar and the branch lengths in terms of genetic distance is denoted above the tree.

Note: ■ OnChd2 molecules; ▲ OnCdh4 molecules.

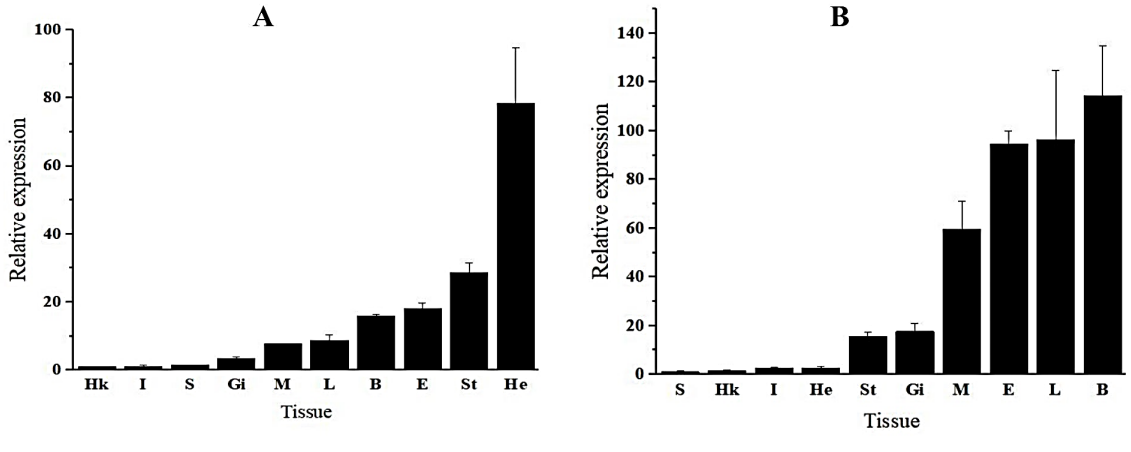


Tissue distribution of OnCdh2 and OnCdh4 in healthy Nile tilapia

To determine the tissue distribution of OnCdh2 and OnCdh4, real-time quantitative PCR (RT-qPCR) was used to detect the expression levels of OnCdh2 and OnCdh4 in healthy Nile tilapia tissues. The results showed that OnCdh2 and OnCdh4 ubiquitously expressed in all detected tissues, with the highest level of expression in heart and brain, respectively (Figure 4A). Moreover, *OnCdh2* gene was highly expressed in stomach, eyes and brain displayed lower expression head kidney and intestine, while high expression of *OnCdh4* gene was observed in liver, eyes, muscle with the lowest levels detected in spleen (Figure 4B).

Figure 4. Tissue distribution of (A) OnCdh2 and (B) OnCdh4 mRNA in healthy Nile tilapia. The ratio refers to the gene expression in different tissues relative to that in head kidney, spleen and target gene expression is normalized to β-actin. Data are presented as means ± standard deviation.

Note: E: Eyes; B: Brain; G: Gill; He: Heart; L: Liver; S: Spleen; St: Stomach; HK: Head kiney; M: Muscle, I: Intestine.



Real-time PCR analysis of the response of *OnGal-2* mRNA expression to bacterial infection

In order to explore the expression pattern of OnCdh2 and OnCdh4 post infection with pathogens, Nile tilapia was challenged with *Streptococcus agalactiae* by oral route, and the expression level of OnCdh2 and OnCdh4 in brain, intestine, spleen, and head kidney were tested by RT-qPCR. The 0 hour tissues were used as control. In the brain, the expression of OnCdh2 and OnCdh4 increased significantly and peaked at 1 hour. Moreover, the expression of OnCdh2 was still at the peak at 3 hours, and then slowly decreased to the normal level. However, the expression level of OnCdh4 decreased sharply to the level below the control group at 3 hours, and then slowly increased, but its expression level was still lower than the control group. In the intestine, the expression level of OnCdh2 and OnCdh4 showed a time-dependent trend, rising slowly, peaking at 6 hours and 12 hours, respectively, and then slowly decreased to normal level.

In the spleen, the expression of OnCdh2 showed the similar time-dependent trend with in the intestine, peaking at 1 hour after stimulation, then decreasing to normal level and reaching the lowest level at 12 hours. The expression level of OnCdh4 also raised sharply at 1 hour (~35.68 folds), suddenly decreased below the normal level at 3 hours, and increased to summit (~75 folds) at 6 hours again, then rapidly decreased to normal level. In the head kidney, OnCdh2 and OnCdh4 showed different expression patterns. During the entire process of the experiment, the expression level of OnCdh2 was significantly lower than that of the control group. The expression pattern of OnCdh4 was higher at 6 hours and 96 hours and lower at other times than at 0 hour.

DISCUSSION

The present study focused on the characterization and expression profiles of two cadherins, OnCdh2 and OnCdh4 of Nile tilapia. OnCdh2 and OnCdh4 are 2721 bp and 2802 bp in length, encoding 906 amino acids and 933 amino acids. OnCdh2 and OnCdh4 both had one cadherin prodomain super family, one cadherin repeat-like domain, three cadherin tandem repeat domain, one cadherin domain and one cadherin cytoplasmic region, which were consistent with structural characteristics of other classical cadherins. Typically, cadherins mediate Ca^{2+} dependent homophilic adhesion, thereby promoting association of cells expressing the same cadherin family members to form adherens junctions [26]. It has been confirmed that classical cadherins have a large extracellular domain that contains four homologous subdomains (one cadherin repeat-like domain and three cadherin tandem repeat domain), which are involved in adhesive interactions [27].

The cadherin pro-domain contains a binding site that is necessary for specific homoadhesion [15,25]. As the other classical cadherins, OnCdh2 and OnCdh4 both had a single transmembrane domain, and a highly conserved cytoplasmic domain that mediated linkage to the actin cytoskeleton. Moreover, a cell aggregation activity of the cytoplasmic domain of cadherin-4 was also characterized through interacting with β-catenin and its partners [28]. The conserved domains of OnCdh2 and OnCdh4 indicated that they may have similar functions to that of human classical cadherins.

Multiple homologous comparison showed that OnCdh2 and OnCdh4 display higher homology (94.10% and 99.25%) to their isoforms in fish than other species (75.63%-75.97% and 70.35%-74.09%), indicating high conservation of *Cdh2* and *Cdh4* in fish. Phylogenetic tree analysis showed that OnCdh2 and OnCdh4 cluster together with other teleost isoforms, indicating that they were closely related, but far away from other species. Therefore, we speculate that the conserved domain of cadherin-2 and cadherin-4 is highly conserved in the process of evolution, and there are different differences among different species.

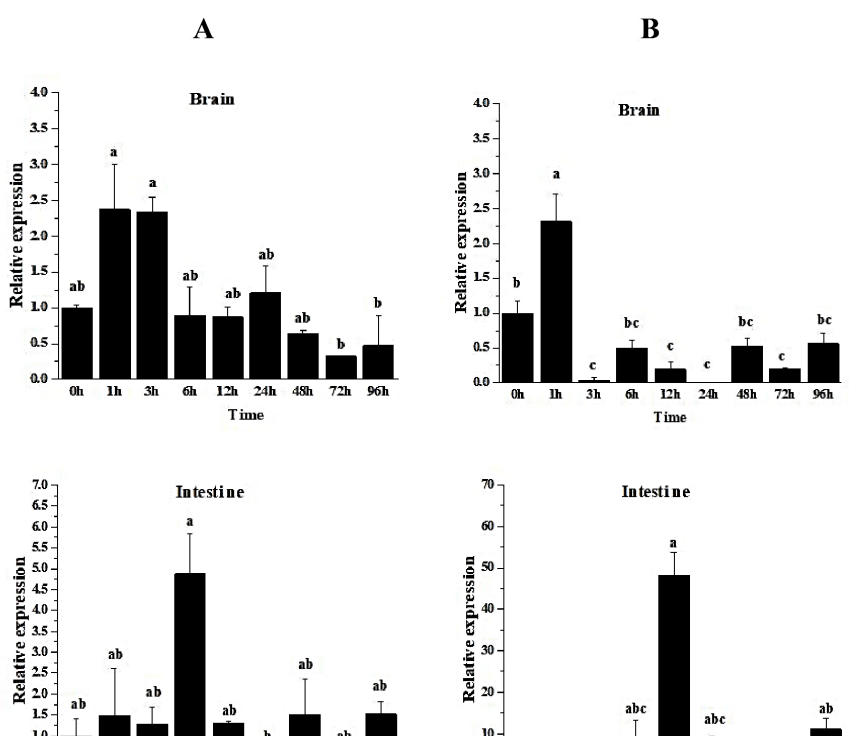
It has been reported that the 16th amino acid of the first cadherin repeat-like domain was proline, which indicated that OnCdh2 and OnCdh4 may be involved in the invasion of *Streptococcus agalactiae* into Nile tilapia (Figure 5). Furthermore, we observed whether the expression level of OnCdh2 and OnCdh4 can be induced by *Streptococcus agalactiae*. First, we found that OnCdh2 and OnCdh4 ubiquitously expressed in all detected tissues, with the high level of expression in heart, stomach, eyes and brain of OnCdh2, while in brain, liver, eyes, muscle of OnCdh4, weakly in intestine, head kidney and skin of OnCdh2 and OnCdh4. We speculate that the high expression of OnCdh2 and OnCdh4 in brain is mainly related to their function mediating cell-cell adhesion.

Obst-Pernberg et al. found that the *Cdh2* and *Cdh4* genes are expressed in specific and restricted patterns in numerous brain nuclei, gray matter areas and cortical layers that are widely distributed throughout the mouse forebrain at postnatal one day [29]. Bagatto et al. found that *Cdh2* also plays an essential role in zebrafish cardiovascular development [10]. In the chicken, mouse and zebrafish, *Cdh4* is considered to play an important role in brain segmentation and neuronal growth, and plays a certain role in the development of kidney and muscle [30-33]. Second, the expression level of OnCdh2 and OnCdh4 showed a significant change in brain, intestine, spleen and head kidney after simulation by *Streptococcus agalactiae*. It is worth noting that the expression of OnCdh2 and OnCdh4 show a similar up-regulated trends in the brain, intestine and spleen show a different expression patterns with up-regulating of OnCdh4 and down-regulating of OnCdh2.

We speculate that this may be related to their function. On the one hand, due to its essential role in several developmental process [34], *Cdh2* and *Cdh4* were specifically up-regulated in regenerating retina and/or the optic pathway in the visual system of adult zebrafish after eye or optic nerve lesions [35] and in regenerating peripheral nervous tissues (eg, sciatic nerve) [36-39]. In cancer, *Cdh2*, as an invasion promoter, is frequently

upregulated [40-42]. Expression of cadherin-2 in epithelial cells induces changes in morphology to a fibroblastic phenotype, rendering the cells more motile and invasive. On the other hand, in some cancers, like osteosarcoma, *Cdh2* may behave as a tumour suppressor [43]. Therefore, we speculate that OnCdh2 and OnCdh4 may be involved in the invasion of *Streptococcus agalactiae*.

Figure 5. Temporal expression of (A) OnCdh2 and (B) OnCdh4 mRNA in brain, intestine, spleen and head kidney after challenge with *Streptococcus agalactiae*. All data were normalized to the expression of β-actin. The results were expressed as mean ± SD (n=3). Significant difference was indicated with "abc", starting from the maximum value, starting from "a" and gradually marking "b", "c", etc., from the maximum value to the small value.



In conclusion, cadherin-2 (OnCdh2) and cadherin-4 (OnCdh4) were successfully cloned and characterized from the Nile tilapia *(Oreochromis niloticus)*. Prediction of protein domains showed that OnCdh2 and OnCdh4 both consisted of one cadherin prodomain super family, one cadherin repeat-like domain, three cadherin tandem repeat domain, one cadherin domain, one cadherin cytoplasmic region and a transmembrane domain. Homology comparisons and phylogenetic tree analysis indicated that OnCdh2 and OnCdh4 showed a higher identity (94.10% and 99.25%) and cluster together with the teleost counterparts. The results of tissue distribution showed that OnCdh2 and OnCdh4 were both ubiquitous in all tissues examined of healthy tilapia. The expression level of OnCdh2 and OnCdh4 were rapidly activated at 1 hour in brain, intestine and spleen after challenged by *Streptococcus agalactiae*. Taken together, the results indicated that OnCdh2 and OnCdh4 might be involved in the process of *Streptococcus agalactiae* invading into Nile tilapia.

ACKNOWLEDGEMENTS

This work was financially supported by Guangxi Natural Science Foundation [Grant No. 2023GXNSFAA026335 and 2018GXNSFAA281208], Guangdong Provincial Key Laboratory of Aquatic Animal Disease Control and Healthy culture [Grant No. PBEA2022ZD04], Guangxi Key Laboratory of Beibu Gulf Marine Biodiversity Conservation [Grant No. 2023ZB08], Marine Science Guangxi First-Class Subject, Beibu Gulf University [Grant No. DRC004].

COMPETING INTERESTS

The authors have declared that no competing interest exists.

AUTHOR CONTRIBUTIONS

Xueying Liang prepared conceptualization, methodology, software, validation, formal analysis, investigation and wrote the original draft preparation. Yusi Zheng took the responsibility of validation, software, and original draft preparation. Zemiao Zhang curated the data, formal analysis and visualization. Yinhui Peng for resources and supervision. Honglin Chen, Peng Xu and Xinzhong Wu, supervised the entire work. Xiaohui Cai involved in writing-review and editing and project administration.

REFERENCES

1. Takeichi M. Cadherins: A molecular family important in selective cell-cell adhesion. J Annu Rev Biochem. 1990;59:237-252.
2. Shimoyama Y, et al. Cadherin cell-adhesion molecules in human epithelial tissues and carcinomas. Cancer Res. 1989;8:2128-2133.
3. Takeichi M. The cadherins: Cell-cell adhesion molecules controlling animal morphogenesis. Development. 1988;4:639-655.
4. Takeichi M. Cadherin cell adhesion receptors as a morphogenetic regulator. Science. 1991;251:1451-1455.
5. Nollet F, et al. Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. J Mol Biol. 2000;299:551-572.
6. Brieher WM, et al. Cadherin junctions and their cytoskeleton(s). Curr Opin Cell Biol. 2013;25:39-46.
7. Yang S, et al. Genome-wide screening of the classical cadherin gene family and cadherin-1 expression response infected with *Streptococcus agalactiae* in Nile tilapia (*Oreochromis niloticus*). Aquacul Rep. 2020;17:100393.
8. Redies C. Cadherin expression in the developing vertebrate CNS: From neuromeres to brain nuclei and neural circuits. Exp Cell Res. 1995;220:243-256.
9. Hatta K, et al. Expression of N-cadherin adhesion molecules associated with early morphogenetic events in chick development. Nature. 1986;320:447-455.
10. Bagatto B, et al. Cadherin2 (N-cadherin) plays an essential role in zebrafish cardiovascular development. BMC Dev Biol. 2006;6:23-36.
11. Redies C. Cadherins and the formation of neural circuitry in the vertebrate CNS. Cell Tissue Res. 1997;290:405-413.
12. Treubert-Zimmermann U, et al. Targeting axons to specific fiber tracts *in vivo* by altering cadherin expression. J Neurosci. 2002;22:7617-7626.
13. Phan QT, et al. Als3 is a *Candida albicans* invasin that binds to cadherins and induces endocytosis by host cells. Plos Biol. 2007;5:e64.
14. Xu XY, et al. N-cadherin as a receptor for adhesion and endocytosis of *Aspergillus fumigatus* by human umbilical vein endothelial cells. Chin J Tuberc Resp Dis. 2010;33:734-737.
15. Bonazzi M, et al. *Listeria* monocytogenes internalin and E-cadherin: From structure to pathogenesis. Cell Microbiol. 2009;11:693-702.
16. Liu G, et al. Development of *Streptococcus agalactiae* vaccines for tilapia. Dis Aquat Organ. 2016;122:163-170.
17. World food and agriculture. Statistical Year book. 2020:1-366.
18. Suanyuk N, et al. Occurrence of rare genotypes of *Streptococcus agalactiae* in cultured red tilapia *Oreochromis* sp. and Nile tilapia *O. niloticus* in Thailand-relationship to human isolates. Aquaculture. 2008;284:35-40.

19. Musa N, et al. Streptococcosis in red hybrid tilapia (*Oreochromis niloticus*) commercial farms in Malaysia. Aquac Res. 2009;40:630-632.

20. Siti A, et al. Multiple streptococcal species infection in cage-cultured red tilapia but showing similar clinical signs. Dis Asian Aqua. 2008:313-320.

21. Chideroli RT, et al. Emergence of a new multidrug-resistant and highly virulent serotype of *Streptococcus agalactiae* in fish farms from Brazil. Aquaculture. 2017;479:45-51.

22. Iregui CA, et al. Experimental early pathogenesis of *Streptococcus agalactiae* infection in red tilapia *Oreochromis* spp. J Fish Dis. 2016;39:205-215.

23. Cai XH, et al. Characterization and identification of *streptococci* from golden pompano in China. Dis Aquat Organ. 2016;119:207-217.

24. Livak KJ, et al. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. Methods. 2001;25:402-408.

25. Lecuit M, et al. A single amino acid in E-cadherin responsible for host specificity towards the human pathogen *Listeria* monocytogenes. EMBO J. 1999;18:3956-3963.

26. Wheelock MJ, et al. Cadherins as modulators of cellular phenotype. Annu Rev Cell Dev Biol. 2003;19:207-235.

27. Liu Q, et al. Differential expression of cadherin-2 and cadherin-4 in the developing and adult zebrafish visual system. Visual Neurosci. 2001;18:923-933.

28. Tanihara H, et al. Characterization of cadherin-4 and cadherin-5 reveals new aspects of cadherins. J Cell Sci. 1994;107:1697-1704.

29. Obst-Pernberg K, et al. Expression of R-cadherin and N-cadherin by cell groups and fiber tracts in the developing mouse forebrain: Relation to the formation of functional circuits. Neuroscience. 2001;106:505-533.

30. Akins MR, et al. Cadherin expression in the developing mouse olfactory system. J Comp Neurol. 2010;501:483-497.

31. Wilson AL, et al. Cadherin-4 plays a role in the development of zebrafish cranial ganglia and lateral line system. Dev Dyn. 2010;236:893-902.

32. Lin JT, et al. Restricted expression of classic cadherins in the spinal cord of the chicken embryo. Front Neuroanat. 2014;8:1-18.

33. Stoya G, et al. Inversion of layer-specific cadherin expression profiles and maintenance of cytoarchitectonic areas in the allocortex of the reeler mutant mouse. J Comp Neurol. 2014;522:3106-3119.

34. Radice GL. N-cadherin-mediated adhesion and signaling from development to disease: Lessons from mice. Prog Mol Biol Transl Ence. 2013;116:263-289.

35. Liu Q, et al. Up-regulation of cadherin-2 and cadherin-4 in regenerating visual structures of adult zebrafish. Exp Neurol. 2002;177:396-406.

36. Bates CA, et al. Expression of polysialylated NCAM but not L1 or N-cadherin by regenerating adult mouse optic fibers *in vitro*. Exp Neurol. 1999;155:128-139.

37. Martini R, et al. Immunoelectron microscopic localization of neural cell adhesion molecules (L1, N-CAM, and myelin-associated glycoprotein) in regenerating adult mouse sciatic nerve. J Cell Biol. 1988;106:1735-1746.

38. Shibuya Y, et al. Localization of N-cadherin in the normal and regenerating nerve fibers of the chicken peripheral nervous system. Neuroscience. 1995;67:253-261.

39. Squitti R, et al. Effects of axotomy on the expression and ultrastructural localization of N-cadherin and neural cell adhesion molecule in the quail ciliary ganglion: An *in vivo* model of neuroplasticity. Neuroscience. 1999;91:707-722.

40. Hazan RB, et al. Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis. J Cell Biol. 2000;148:779-790.

41. Nieman MT, et al. N-Cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression. J Cell Biol. 1999;147:631-644.

42. Yu TH, et al. Viral hepatitis is associated with intrahepatic cholangiocarcinoma with cholangiolar differentiation and N-cadherin expression. Modern Pathol. 2011;24:810-819.

43. Derycke LDM, et al. N-cadherin in the spotlight of cell-cell adhesion, differentiation, embryogenesis, invasion and signaling. Int Dev Biol. 2004;48:463-476.