Comparative Biochemistry and Physiology, Part D 13 (2015) 60–69

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology, Part D

journal homepage: www.elsevier.com/locate/cbpd

Claudin multigene family in channel catfish and their expression profiles in response to bacterial infection and hypoxia as revealed by meta-analysis of RNA-Seq datasets

Luyang Sun, Shikai Liu, Lisui Bao, Yun Li, Jianbin Feng, Zhanjiang Liu *

The Fish Molecular Genetics and Biotechnology Laboratory, Aquatic Genomics Unit, School of Fisheries, Aquaculture and Aquatic Sciences, Program of Cell and Molecular Biosciences, Auburn University, Auburn, AL 36849, USA

ARTICLE INFO

Article history:
Received 26 October 2014
Received in revised form 14 January 2015
Accepted 14 January 2015
Available online 21 January 2015

Keywords:
Claudin
Tight junction
Disease infection
Hypoxia
Fish

ABSTRACT

Claudins are one of the major groups of transmembrane proteins that play crucial roles in tight junctions. In addition to their function in the regulation of paracellular permeability, claudins are also involved in a number of biological processes related to pathogen infection, embryonic development, organ development and hypoxia response. Despite its importance, analyses of claudin genes in channel catfish have not been systematically performed. In this study, a total of 52 claudin genes were identified and characterized in channel catfish. Phylogenetic analyses were conducted to determine their identities and identify a number of lineage-specific claudin gene duplications in channel catfish. Expression profiles of catfish claudin genes in response to enteric septicemia of catfish (ESC) disease and hypoxia stress were determined by analyzing existing RNA-Seq datasets. Claudin genes were significantly down-regulated in the intestine at 3 h post-infection, indicating that pathogens may disrupt the mucosal barrier by suppressing the expression of claudin genes. A total of six claudin genes were significantly regulated in the gill after hypoxia stress. Among them, the expressions of *cldn-11b* and *cldn-10d* were dramatically altered when comparing hypoxia tolerant fish with intolerant fish, though their specific roles involved in response to hypoxia stress remained unknown.

1. Introduction

Tight junctions are structures that appear in the areas of cell–cell contact where plasma membranes join together. The variability of tight junction tightness makes it possible to form tissue barriers and pores, which can prevent uncontrolled passage of molecules and ions through the intercellular space (Krause et al., 2008). They also function as a fence to separate the apical part from other parts of membrane proteins of epithelial cells. They are involved in the configuration of membrane proteins that are essential for directional transcellular transport (Krause et al., 2008). Four primary groups of tansmembrane proteins are present in epithelial tight junctions: occludin, claudins, the junction-adhesion-molecules (JAMs), and the Coxsackievirus and adenovirus Receptor (CAR) proteins.

Claudins are important for the formation and function of tight junctions (Tsukita and Furuse, 2000). They were first discovered from chicken liver in 1998 as a family of novel integral membrane proteins of tight junctions (Furuse et al., 1998). Since then, many claudin genes have been identified (Kollmar et al., 2001). In mammals, a total of 27 claudin genes have been identified (Mineta et al., 2011; for review see Günzel and Alan, 2013). Among teleosts, 56 claudin genes were identified in Fugu (*Takifugu rubripes*) and at least 43 claudin genes were identified in zebrafish (*Danio rerio*) (Loh et al., 2004; Baltzegar et al., 2013; Kolosov et al., 2013). The number of claudin genes in teleost fish is apparently larger than that in mammals, as a result of fish-specific whole genome duplication in teleosts (Hoegg et al., 2004; Hurley et al., 2007). In contrast to the high numbers of claudin genes in vertebrates, only a few claudin homologues have been identified in invertebrates such as *Drosophila melanogaster* (Wu et al., 2004) and *Caenorhabditis elegans* (Asano et al., 2003).

The primary function of claudins is to regulate the paracellular permeability properties of vertebrate epithelia. Different claudins can either enhance the "barrier" properties of tight junctions or enhance the "leak or pore" forming properties of tight junctions (Kolosov et al., 2013). For example, claudin-2 and claudin-4 were involved in the permeability of intercellular Na^+ and K^+, and claudin-16 is essential for renal resorption of Ca^{2+} and Mg^{2+} (Van Itallie et al., 2001; Müller et al., 2003).

Claudins were also reported to be involved in immune and stress responses. For instance, claudin-1 and claudin-2 were reported to be involved in the early stage of neoplastic transformation, and were up-regulated in the associated inflammatory bowel disease (Weber et al.,

* Corresponding author at: The Fish Molecular Genetics and Biotechnology Laboratory, Aquatic Genomics Unit, School of Fisheries, Aquaculture and Aquatic Sciences, Program of Cell and Molecular Biosciences, Auburn University, Auburn, AL 36849, USA. Tel.: +1 334 844 4784; fax: +1 334 844 2937.
E-mail address: liuzhan@auburn.edu (Z. Liu).

http://dx.doi.org/10.1016/j.cbd.2015.01.002
1744-117X/

2008). The barrier formed by claudins and other transmembrane proteins can protect the organisms by preventing the pathogens from penetrating through the epithelial layers. However, some pathogens can affect the structure of tight junctions and influence the synthesis of claudins. Studies on lung diseases revealed that asthma could lead to the fragmentation and thinning of the tight junction belt and down-regulation of claudin genes (Godfrey, 1997). The synthesis and organization of claudins as well as other tight junction proteins could be affected by several integrins and growth factors produced by inflammation (Mazzon and Cuzzocrea, 2007). The down-regulation of claudin-5 in the cells suffering from respiratory distress led to local hypoxia of the lung (Soini, 2011). Furthermore, a recent study reported that the expression of a number of claudins was altered in cancer cells, suggesting their new roles in addition to the regulation of paracellular permeability (Singh et al., 2010). In fish, it has been reported that claudins were involved in responses to growth hormone, cortisol and salinity (Tipsmark et al., 2009; Bui et al., 2010; Clelland et al., 2010; Tipsmark and Madsen, 2012).

With the interest of understanding the claudin genes in the channel catfish genome and their involvement in response to disease infection and hypoxia stress, herein, we report the identification and characterization of 52 claudin genes in the channel catfish (*Ictalurus punctatus*) genome. The identities of catfish claudins were determined by phylogenetic analyses, and their mRNA expression abundance after enteric septicemia of catfish (ESC) disease infection and hypoxia challenges were analyzed by meta-analysis of RNA-Seq datasets, respectively.

2. Materials and methods

2.1. Gene identification and sequence analysis

The initial catfish claudin sequences were identified by searching the channel catfish transcriptome database using TBLASTN, with human, zebrafish (*D. rerio*) and Fugu (*T. rubripes*) claudins as queries. *E-value* was set equal to $1e^{-5}$ to ensure the quality of the results. The transcriptome database was generated by RNA-Seq assembly of a doubled haploid channel catfish (Liu et al., 2012), which has been used as the main resource for the identification of full-length gene transcripts in various catfish gene family studies (Rajendran et al., 2012a, 2012b; Liu et al., 2013a; Wang et al., 2013a, 2014; Zhang et al., 2013; Sun et al., 2014). Duplicates in the initial sequence pool were eliminated by using ClustalW2 (https://www.ebi.ac.uk/Tools/msa/clustalw2/) and a unique set of sequences was subject to further analysis. The unique sequences from the RNA-Seq database were used to blast against catfish draft whole genome sequence database (unpublished) with a cutoff *E-value* of $1e^{-10}$ for further confirmation. Open reading frame (ORF) was identified by using ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The identities of the predicted ORFs were then verified by BLASTP against NCBI non-redundant (nr) protein database with a cutoff *E-value* of $1e^{-5}$. The gene identities were determined by performing phylogenetic analysis with the claudin amino acid sequences from multiple species (see below).

2.2. Phylogenetic analysis

The claudins from several representative vertebrates were selected for phylogenetic analysis, including those from human, mouse and Fugu. The amino acid sequences of claudin genes from these species were retrieved from the NCBI, Ensembl and UniProt databases. Multiple alignment of claudin amino acid sequences was conducted using the ClustalW2 program (Larkin et al., 2007). Phylogenetic analyses were performed using MEGA 5.2 with neighbor-joining method (Tamura et al., 2011). Bootstrapping with 1000 replications was conducted to evaluate the phylogenetic trees, only values greater than 50% were shown.

2.3. Expression of claudin genes in healthy tissues

Meta-analysis was conducted to compare the mRNA expression abundance among healthy tissues, including gill, intestine, liver, ovary and testis. Illumina-based RNA-Seq datasets of control fish (untreated healthy fish) were retrieved from previous and on-going RNA-Seq studies (Li et al., 2012; Sun et al., 2012; Wang et al., 2013b). Normalization was performed for the combined dataset using RUVSeq R package (Risso et al., 2014), which utilized the information of negative controls. Before normalization, a set of in-silico "empirical" negative controls were first identified using edgeR R package (Robinson et al., 2010). These negative control genes were the least significantly differentially expressed genes among the five healthy tissues. The gene read counts of these negative control genes were then used by the RUVg function of the RVUSeq package for claudin gene read count normalization to control the bias in the meta-analysis, such as batch effect, sequencing depth difference and bias generated during sequencing library preparation. After normalization, the mRNA read counts for each of the 52 claudin genes were used to measure the mRNA abundance of each claudin in the five tissues. Claudins with zero mRNA read count were regarded as "undetectable". Claudins with less than five read counts were regarded as "almost undetectable", which were removed in the following ESC disease infection analysis and low oxygen stress response analysis.

2.4. Expression of claudin genes after ESC disease infection

Illumina-based RNA-Seq datasets were retrieved from a previous RNA-Seq study of the entire intestine tissue in catfish after infection with the *Edwardsiella ictaluri* (Li et al., 2012). A detailed protocol of disease challenge, RNA extraction, library construction and sequencing was described in Li et al. (2012). Four time-points (0 h, 3 h, 24 h and 3 d) were included in the analysis. The expression of claudin genes at 3 h, 24 h, and 3 d post-infection was compared with that at 0 h to determine mRNA abundance changes, respectively. Read mapping was performed using CLC Genomics Workbench (version 4.0.2; CLC bio, Aarhus, Denmark). Before mapping, raw reads were trimmed to remove adaptor sequences, ambiguous nucleotides (N's), extremely short reads (<30 bp) and low quality sequences (Quality score < 20). The clean reads from each time-point were then aligned with all gene transcripts assembled previously (Li et al., 2012) in addition with all cDNA sequences of channel catfish claudin genes identified in this study. The mapping parameters were set as: mismatch cost of 2, deletion cost of 3 and insertion cost of 3. The highest scoring matches that shared ≥95% similarity with the reference sequence across ≥90% of their lengths were included in the alignment.

The number of total mapped reads for each claudin gene was determined and normalized to assess the expression level as denoted by RPKM (reads per kilobase of the transcript per million mapped reads). The expression fold change of each claudin gene was calculated based on normalized RPKM. Claudin genes with absolute fold change values ≥ 2, total read number ≥ 5 and p-value ≤ 0.05 were regarded as differentially expressed genes.

2.5. Expression of claudin genes after hypoxia stress

The expressions of claudin genes after hypoxia stress were determined by conducting similar meta-analysis. Illumina-based RNA-Seq datasets were retrieved from our RNA-Seq study in which the gill tissue of catfish after treatment of hypoxia stress was sequenced (NCBI SRA no. SRP039612). Three datasets were used in the study, including sequenced reads from fish tolerant to hypoxia, sequenced reads from fish susceptible to hypoxia, and sequenced reads from control fish without hypoxia stress. The differential expression analyses among these three groups were conducted the same as described above in ESC disease infection.