Molecular Immunology 65 (2015) 277–286

Contents lists available at ScienceDirect

Molecular Immunology

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

Annotation and genetic diversity of the chicken collagenous lectins

Edin Hamzić[a,b,c], Marie-Hélène Pinard-van der Laan[b,a], Bertrand Bed'Hom[b,a], Helle Risdahl Juul-Madsen[d,*]

[a] AgroParisTech, UMR1313 Animal Genetics and Integrative Biology, 16 rue Claude Bernard, 75231 Paris 05, France
[b] INRA, UMR1313 Animal Genetics and Integrative Biology, Domaine de Vilvert, 78350 Jouy-en-Josas, France
[c] Department of Molecular Biology and Genetics, Aarhus University, Blichers Allé 20, Building K23, 8830 Tjele, Denmark
[d] Department of Animal Science, Aarhus University, Blichers Allé 20, Building P25, 8830 Tjele, Denmark

ARTICLE INFO

Article history:
Received 10 November 2014
Received in revised form 26 January 2015
Accepted 30 January 2015
Available online 24 February 2015

Keywords:
Chicken
Collectins
Ficolins
Genotyping
Immunogenetics
Innate immunity

ABSTRACT

Collectins and ficolins are multimeric proteins present in various tissues and are actively involved in innate immune responses. In chickens, six different collagenous lectins have been characterized so far: mannose-binding lectin (MBL), surfactant protein A (SP-A), collectin 10 (COLEC10), collectin 11 (COLEC11), collectin 12 (COLEC12), lung lectin (LL) and one ficolin (FCN). However, the structural and functional features of the chicken collectins and ficolin are still not fully understood. Therefore, the aims of this study were: (i) to make an overview of the genetic structure and function of chicken collectins and the ficolin, (ii) to investigate the variation in the chicken collectins and the ficolin gene in different chicken populations, and (iii) to assess the presence of *MBL* gene variants in different chicken populations. We performed comparative genomic analysis using publically available data. The obtained results showed that collectins and ficolins have conserved protein sequences and gene structure across all vertebrate groups and this is especially notable for COLEC10, COLEC11 and COLEC12. For the purpose of studying the genetic variation, 179 animals from 14 populations were genotyped using 31 SNPs covering five genomic regions. The obtained results revealed low level of heterozygosity in the collagenous lectins except for the *COLEC12* gene and the *LL–SPA–MBL* region compared to heterozygosity at neutral microsatellite markers. In addition, the *MBL* gene variants were assessed in different chicken populations based on the polymorphisms in the promoter region. We observed 10 previously identified *MBL* variants with A2/A8 and A4 as the most frequent alleles.

1. Introduction

The innate immune system represents the first line of defense against invading pathogens. Main defining features of the innate immune response are: phagocytic and lytic functions, immune modulation and pattern recognition (Medzhitov and Janeway, 2000). Pattern recognition is an evolutionary conserved part of the innate immunity where pattern recognition receptors (PRRs) identify pathogen-associated molecular patterns (PAMPs) which are associated with microbial pathogens or cellular stress (Janeway, 1989). Collectins and ficolins belong to the group of soluble PRRs and they are present in various tissues in animals interacting with different microbial surface polysaccharides. In contrast to humans, structural and functional aspects of chicken collectins and ficolin are still not fully understood. Up to now, six different collectins have been described in chickens: mannose-binding lectin (MBL), surfactant protein A (SP-A), collectin 10 (COLEC10), collectin 11 (COLEC11), collectin 12 (COLEC12), lung lectin (LL), and one ficolin (FCN).

MBL binds to a wide range of pathogens leading to activation of the complement system or phagocytosis (Ip et al., 2009). An extensive number of studies documented the presence or absence of clinical impact of MBL deficiency in serum and association between the common MBL genetic variants and various diseases in humans (Heitzeneder et al., 2012). Most mammalian species have two different MBL proteins; MBL-A encoded by *MBL1* and MBL-C encoded by *MBL2* (Agah et al., 2001; Gjerstorff et al., 2004; Sastry et al., 1991). In chickens only one MBL gene was identified and characterized by Laursen et al. (1995). Serum concentration of chicken MBL has also been associated with the severity of infections caused by

* Corresponding author at: Department of Animal Science, Aarhus University, Blichers Allé 20, Building P25, Room 3354, 8830 Tjele, Denmark. Tel.: +45 87157837.
E-mail addresses: edin.hamzic@jouy.inra.fr, edin.hamzic@mbg.au.dk (E. Hamzić), marie-helene.pinard@jouy.inra.fr (M.-H. Pinard-van der Laan), bertrand.bedhom@jouy.inra.fr (B. Bed'Hom), helle.juulmadsen@anis.au.dk (H.R. Juul-Madsen).

http://dx.doi.org/10.1016/j.molimm.2015.01.027
0161-5890/

infectious bronchitis virus (IBV) (Juul-Madsen et al., 2007; Kjærup et al., 2014a). Schou et al. (2010) reported an association between serum concentration of chicken MBL and the disease severity of *Pasteurella multocida* infection. In addition, chickens with low chicken MBL serum concentrations had a lower weight gain after being infected with *Escherichia coli* (Norup et al., 2009). Recently, Kjærup et al. (2014a observed that a higher antibody response may be obtained in chickens with high concentration of MBL in serum after addition of an MBL ligand to vaccine, indicating that chicken MBL is involved in the adaptive immune response to IBV vaccination as previously suggested by Juul-Madsen et al. (2011). These findings verify that, similar to mammals, chicken MBL plays an important role in innate immune responses as well as in adaptive.

Humans possess three surfactant proteins with pathogen recognition properties: SP-A1, SP-A2 and SP-D, encoded by respectively *SFTPA1*, *SFTPA2* and *SFTPD*. SP-A is capable of binding to a wide range of microbes leading to agglutination and/or aggregation which enhances their attachment to phagocytic cells leading to killing and clearance (Nayak et al., 2012). DiAngelo et al. (1999) identified several genetic variants for human *SFTPA1* and *SFTPA2* genes. Furthermore, numerous studies have reported associations between *SFTPA1*, *SFTPA2* and *SFTPD* variants and pulmonary diseases in different human populations and study groups (Silveyra and Floros, 2012). Chickens possess only one surfactant protein A (SP-A) and no SP-D. However, chickens also possess lung lectin protein (LL) which is a C-type lectin lacking a collagen-like domain whose sequence strongly resembles the SP-A sequence. No human homolog of chicken lung lectin has been discovered. LL forms trimer structures which inhibits hemagglutination activity of isolates of human influenza A virus (Hogenkamp et al., 2008). Additionally, Reemers et al. (2010) have reported that mRNA expression of chicken *LL* and *SFTPA* were up regulated in the trachea and down regulated in the lung after an avian influenza A virus infection indicating that LL and SP-A might play an important role in the innate immune system.

Human collectin liver 1 (CL-L1) and collectin kidney 1 (CL-K1) are serum proteins and belong to the group of more recently characterized collectins. Henriksen et al. (2013a) reported that CL-L1 and CL-K1 form heteromeric complexes which interact with mannose binding lectin associated serine proteases (MASP) mediating activation of the complement pathway, like MBL. CL-K1 is as a collectin with Ca^{2+}-dependent lectin activity and binds to different microorganisms (Hansen et al., 2010; Keshi et al., 2006). Additionally, Henriksen et al. (2013b) have reported that CL-K1 also interacts with extracellular DNA and might have a role in response to particles and surfaces presenting extracellular DNA. The chicken orthologous of CL-L1 and CL-K1 are collectin 10 (COLEC10) and collectin 11 (COLEC11), respectively. Hogenkamp et al. (2006) have originally characterized chicken COLEC10 and COLEC11 naming them chicken collectin 1 (cCL-1) and chicken collectin 2 (cCL-2), respectively. As for *LL* and *SFTPA*, Reemers et al. (2010) have observed that expression of chicken *COLEC10* and *COLEC11* was up regulated in trachea and down regulated in lung after an avian influenza A virus infection indicating potentially important role of these genes in chicken innate immunity.

Ficolins are also serum PRRs, with a similar structure as MBL and SP-A. Three ficolins have been described in humans: ficolin-1 (M-ficolin, FCN1), ficolin-2 (L-ficolin, FCN2) and ficolin-3 (H-ficolin, FCN3) which are encoded by *FCN1*, *FCN2* and *FCN3*, respectively (Edgar, 1995; Endo et al., 1996; Harumiya et al., 1995; Matsushita et al., 1996; Sugimoto et al., 1998). In addition, Munthe-Fog et al. (2009) reported that a frame-shift variation in FCN3 gene is associated with immunodeficiency. Only one ficolin gene has been characterized in chicken by Lynch et al. (2005). No functional study on chicken ficolin has been reported up to now.

Based on the research conducted so far, it is evident that collectins and ficolins have important roles in innate immune responses. Innate immunity is focused on responses to a broad range of pathogens and therefore investigating the genetic basis of innate immunity is of great interest for improving chicken robustness. Therefore, the aims of this study were: (i) to make an overview of the genetic structure and function of chicken collectins and the ficolin (ii) to investigate the variation in the chicken collectins and the ficolin gene in different chicken populations (iii) to assess the presence of the *MBL* gene variants in different chicken populations.

2. Materials and methods

2.1. Comparative genomics

For comparison of human and chicken collectins and ficolins gene models we used publically available data from: Ensembl Genome Browser (http://www.ensembl.org/), NCBI Gene Database (http://www.ncbi.nlm.nih.gov/gene) and UCSC Genome Browser (http://genome.ucsc.edu/). Gene expression profiling was assessed using available literature and NCBI UniGene Profiles (http://www.ncbi.nlm.nih.gov/unigene/) (Table 1).

Synteny comparison was performed using Genomicus (Louis et al., 2013; Muffato et al., 2010). We used following chicken protein Ensembl IDs for phylogenetic analysis and synteny comparison: mannose-binding lectin (ENSGALP00000040351), surfactant protein A (ENSGALP00000042366), lung lectin (ENSGALP00000003921), collectin 10 (ENSGALP00000043336), collectin 11 (ENSGALP00000042407), collectin 12 (ENSGALP00000030293), and ficolin (ENSGALP00000014886).

The neighbor-joining tree was constructed based on the alignment obtained from ClustalW (Larkin et al., 2007). Chicken protein sequences used for the alignment were retrieved from Ensembl Genome Browser release 77 taking only the longest protein sequences. We used protein sequences of collectins and ficolins from chicken, cow, human, pig, zebra finch and zebra fish.

2.2. Origin of genomic DNA and population history

DNA sample from 179 animals originating from 14 different populations and lines of chicken (*Gallus gallus*) (Table 2) were used in the study. Samples for Brown egg layer, Fayoumi, Houdan, Nagoya, Orlov, Red jungle fowl and White egg layer originated from the European AvianDiv project described by Hillel et al. (2003). The other seven populations were: Skalborg (Sorensen, 2001), Old Danish White Cornish (Sorensen, 1984), Babcock, White Hisex, New Hampshire and White Leghorn which originate from Aarhus University.

2.3. Selection of single nucleotide polymorphisms (SNPs)

Selection of SNPs used in population analysis was performed in multiple steps. The first step was a collection of all available information on SNPs in the genomic regions covering the chicken collectins and ficolin genes taking into account 10 kb of flanking sequences for each genomic region. The six chicken collagenous lectins and the two ficolin genes are located in five genomic regions, because *LL*, *SFTPA* and *MBL* form a cluster on chromosome 6. The following sources of information were used for retrieval of SNP data:

(i) Publicly available data from NCBI SNP database (http://www.ncbi.nlm.nih.gov/snp).
(ii) Information on the genetic variation in divergently selected lines for residual feed intake. These lines were described in detail by Bordas et al. (1992).
(iii) Novel SNPs detected using sequence data from SRA (Short Read Archive, http://www.ncbi.nlm.nih.gov/sra) and EST (Expressed