RESEARCH NOTE

Characterization of duck (*Anas platyrhynchos*) *MHC class I* gene in two duck lines

LIN ZHANG[1], WEI-JIE LIU[1,2], JIA-QIANG WU[1]*, MIN-LI XU[1], ZHENG-JIE KONG[1,2], YAN-YAN HUANG[1] and SHAO-HUA YANG[1]

[1]*Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Science/Shandong Key Lab of Animal Disease Control and Breeding, Jinan 250100, Shandong, People's Republic of China*
[2]*College of Animal Science and Technology, Shandong Agricultural University, Taian 271018, Shandong, People's Republic of China*

Abstract

To enrich gene polymorphism of DuMHC I and provide data for further studies on disease resistance, 14 *DuMHC I* genes from Weishan Ma duck and Cherry Valley duck were cloned, and their characterization were investigated. The overall conservation of the 14 alleles could be observed within the sequences, and relative conservation were also displayed in the peptide-binding domain and CD8 interaction sites. Based on full-length amino acid homology, MHC class I from different duck lines could be divided into 13 gene groups and three novel gene groups existed. Moreover, 14 key variable residues corresponding to gene groups division were exhibited on the homology modelling constructed based on the resolved protein structure of DuMHC I. This study explicit the characteristics of DuMHC I in the two duck lines and could contribute to design effective diagnostics and vaccines for the species against various infections.

[Zhang L., Liu W.-J., Wu J.-Q., Xu M.-L., Kong Z.-J., Huang Y.-Y. and Yang S.-H. 2017 Characterization of duck (*Anas platyrhynchos*) *MHC class I* gene in two duck lines. *J. Genet*. **96**, 371–375]

Introduction

Major histocompatibility complex (MHC) class I molecules play a critical role in the immune defences against pathogens. Main function of the proteins is the presentation of endogenously derived peptides to specific T-cell receptors (TCRs) on $CD8^+$ T cells, resulting in the activation of cytotoxic lymphocytes (CTL) and the subsequent lysis of target cells (Bjorkman and Parham 1990; Garboczi *et al.* 1996). The MHC I complex contains a heavy chain and a light chain (also called β2-microglobulin; β2m). The heavy chain is comprised of α1, α2 and α3 domains, in which α1 and α2 performed polymorphism. Polygeny and polymorphism of *MHC class I* alleles contribute to the breadth of the immune response (Moon *et al.* 2005). Compared with mammal and chicken MHC class I, little was known about duck MHC class I molecule (DuMHC I), until its cDNA sequence was reported in 2004 (Xia *et al.* 2004). Further studies have demonstrated that DuMHC I contains five differentially expressed genes but predominantly express one (Moon *et al.* 2005). These reports also indicate that polymorphism exists in *DuMHC I* and the polymorphism results in the presentation of different sets of antigenic peptides by each DuMHC I allomorph. In light of the role of duck as a reservoir of influenza and agent of transmitting the virus to other avian and mammalian hosts, including human, further investigation into MHC I of ducks is needed to determine their capacity for defense against virus. However, only two duck lines (Peking duck and White Peking duck (also called Cheery Valley duck)) and 11 complete mRNAs of *DuMHC I* have been investigated by now, and other lines as well as more number of genes have not been identified. Therefore, we considered to investigate *DuMHC I* genes in main layer and meat lines in China to find out the gene polymorphism and provide data for further studies on disease resistance.

Weishan Ma duck (WS) is the eugenic endemic breed, one of the four famous ducks in China, as well as Cherry Valley duck (CV) are the largest number of breeding variety. WS is egg strain and CV is meat type. The two duck

*For correspondence. E-mail: 443794629@qq.com.

Keywords. duck; major histocompatibility class I; molecular characteristic; allelic group; amino acid variability.

lines mainly support the Chinese waterfowl industry. In this study, *MHC class I* genes of WS and CV were cloned, followed by analysis of the molecular characteristics. The results revealed that there were three novel *DuMHC I* allelic groups in the two duck lines. This type of discovery provided insight into the characterization of *DuMHC I* and contribute to design effective diagnostics and vaccines for the species against various infections.

Materials and methods

Twelve-week-old ducks used in this study were from two lines: WS and CV, collected from three and two different farms in Shandong province, respectively. Total RNA of fresh spleen tissue from five WS and five CV were extracted using TRIzol reagent (Invitrogen, Carlsbad, USA). The isolated RNA samples were stored at −80°C until use. Then the first-strand cDNAs were synthesized using random primer from RNA samples. To amplify the *DuMHC I* gene, a pair of primers was designed according to the Peking and White Peking *DuMHC I* (AB115246 and AY294416). The forward prime P1 was 5′-ATGGGCGGGGCCCTGGGCCT-3′ and reverse prime P2 was 5′-TTAGACACTGGGGTTGCTCCCTGCG-3′. PCR was carried out in a final volume of 50 μL which includes 0.5 μg cDNA template, P1/P2 (100 μM) 1 μL, respectively, 0.5 U LA *Taq* polymerase (TaKaRa Biotechnology, Dalian, China) and amplified 98°C for 5 min, followed by 32 cycles (94°C 1 min, 65°C 1 min, 72°C 2 min), and ending with 72°C 10 min. PCR products of about 1.1 kb were separated by agarose gel electrophoresis and purified using DNA recovery kit (Tiangen Biotech, Beijing, China). Then, the purified fragments were inserted into pMD18-T easy vector (TaKaRa Biotechnology) according to the manufacturer's recommendations and transformed competent *E. coli* DH5α (Transgen Biotech) coated on LB plates containing ampicillin (100 μg/mL), IPTG (40 μg/mL) and X-gal (20 μg/mL). After incubation overnight at 37°C, three white spots on each LB plates were identified by restriction enzyme analysis using *Eco*R I and *Hin*d III (TaKaRa Biotechnology) and positive clones were sequenced by Wuhan Genecreate Biological Engineering, China. According to the sequence results, consensus sequences of the different clones for each allele from one animal were selected and submitted to GenBank of national center for biotechnology information (NCBI).

The GenBank accession numbers of the *DuMHC I* genes belonging to the two duck lines are listed in table 1. Alignments were performed using CLUSTALW. Comparison of deduced amino acid sequences were carried out using the search similarity and multiple alignment programmes of GENETYX9.0 computer software (Software Development, Tokyo, Japan) and DNAMAN demo software (Lynnon BioSoft, Quebec, Canada). To calculate amino acid variability, the protein variability server (calculating Wu-Kabat index, http://imed.med.ucm.es/PVS/) were used. The 3D structure of the extracellular domains of DuMHC I were predicted based on the known protein structure by SWISS-MODLE (https://swissmodel.expasy.org/). The amino acids corresponding to gene group division were analysed in the 3D structure which were made by pymol (DeLano Scientific, https://www.pymol.org).

Table 1. Gene groups of *DuMHC I* of the two duck lines and Peking duck.

Sequence name	Allelic groups	Alignment with the know sequences	GenBank accession no.
WS ducks			
UAA01*WS	11	New	KX118673
UAA03*WS	11	New	KX118675
UAA04*WS	11	New	KX118676
UAA05*WS	12	New	KX118677
UAA07*WS	8		KX118679
UAA08*WS	5		KX118680
UAA09*WS	12	New	KX118681
CV ducks			
UAA01*CV	8		KX118683
UAA02*CV	1		KX118684
UAA03*CV	1		KX118685
UAA04*CV	1		KX118686
UAA05*CV	1		KX118687
UAA06*CV	8		KX118688
UAA07*CV	13	New	KX118689
White Peking ducks			
Anpl-U02	1		AY294416
Anpl-U03	2		AY294417
Anpl-U04	3		AY294418
Anpl-U05	4		AY294419
Peking ducks			
Anpl-UAA01	5		AB115242
Anpl-UAA02	6		AB115241
Anpl-UBA01	7		AB115244
Anpl-UBA02	8		AB115243
Anpl-UCA01	9		AB115245
Anpl-UDA01	10		AB115246

Results and discussion

Fourteen *DuMHC I* sequences from the two duck lines covering an open reading frame (ORF) contained leader peptides, α1, α2, α3 and TM/CY domains were obtained. All the 14 alleles belonged to UAA locus, seven of which were obtained from WS (named UAA01*WS, UAA03-05*WS, UAA07-09*WS) and seven from CV (named UAA01-07*CV). The UAA leader peptides, α1, α2, α3 domains in the two lines were composed of 21, 88, 92 and 91 amino acids, respectively. The striking amino acids length variations existed in TM/CY domains, most of which encoding 62 amino acids, while two alleles (UAA05*WS and UAA09*WS) encoding 61 amino acids and UAA04*CV encoding 51 without a part of exons 7 and 8 (Moon *et al.* 2005). As shown in figure 1, four

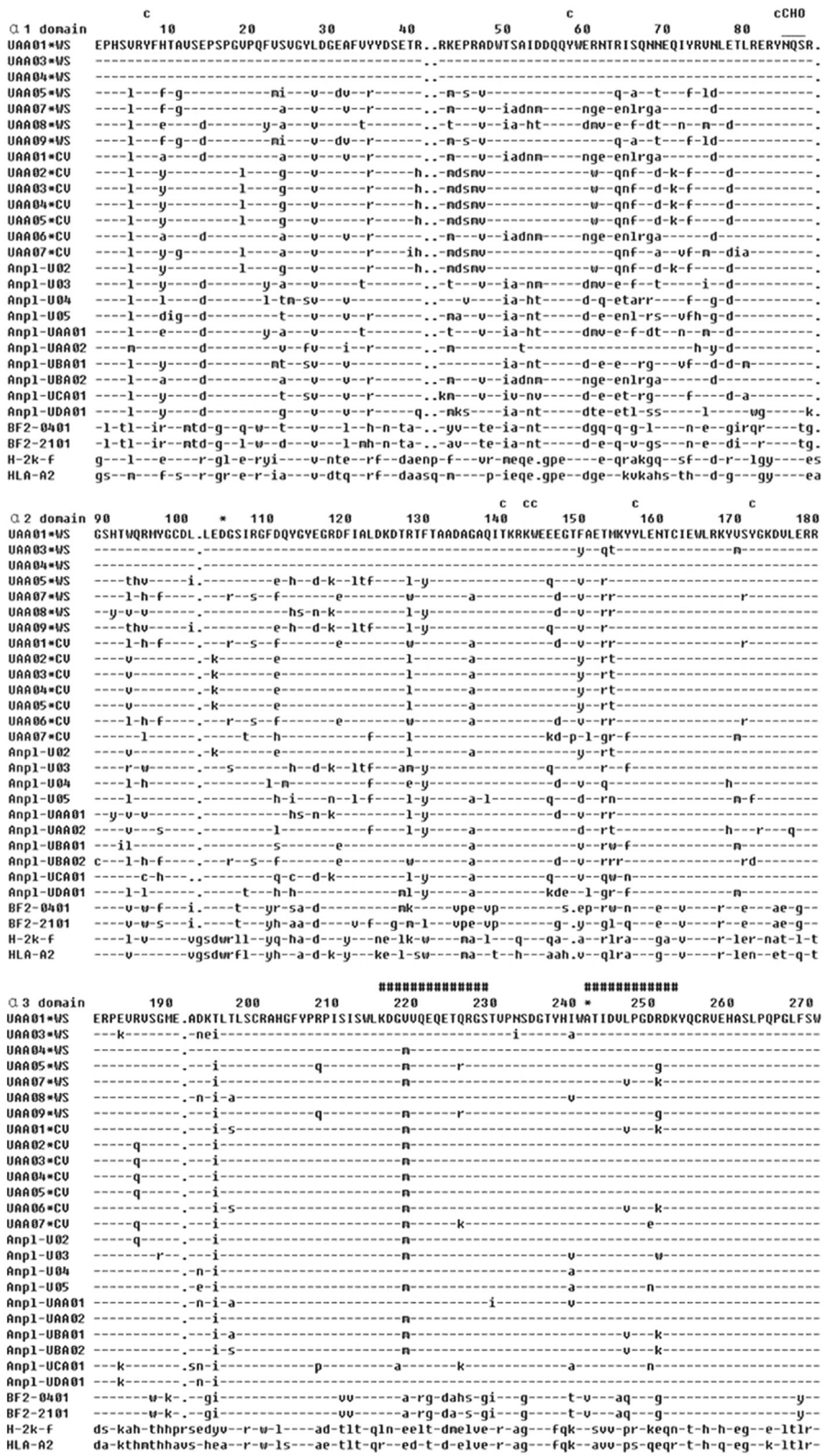

Figure 1. The α*pg*l, α2 and α3 domain amino acids of *MHC I* in WS, CV, Peking duck, mouse and human are aligned. The symbols are used to indicate the following: '.' gaps in the sequence; '-' the same amino acid as those in UAA01*WS; 'c' the conserved eight key residues interacting with peptide presentation in the extracellular domain; 'CHO' glycosylation site; '_' conserved residues involved in binding the peptide-terminal in UAA*WS and UAA*CV; '*' conserved amino acid binding with CD8 in HLA-A; '#' negatively charged residues with CD8 binding.

cysteines (positions 99, 162, 220 and 256) were present, likely to form two sets of an intrachain disulphide bridge in α2, α3 domains and the potential N-glycosylation site existed at positions 85–87.

By alignment of the *DuMHC I* genes in WS and CV with that in Peking duck and BF2 in chickens, eight key residues (7Y, 58Y, 84Y (84R in nonmammalian vertebrates), 140T, 143K, 144W, 157Y, 172Y) interacting with peptide presentation in the extracellular domain were conserved. In agreement with other avian sequences and many other species, 120F in all 14 *DuMHC I* genes replaced 123Y, which also involved in peptide anchoring (Shum *et al.* 1999). Conserved residues NQSR involved in binding the peptide-terminal main chain atoms also existed in UAA*WS and UAA*CV. The negatively charged residues in two regions of α3 domain (positions 216–229 and 242–253) implicated in CD8 binding were conserved except 251K mutated in UAA07*WS, UAA01*CV and UAA06*CV. The key residue 245A, which affected CD8 binding in HLA-A (Salter *et al.* 1989) also conserved in the *DuMHC I* sequences. In addition, 104D forming hydrogen bond with CD8 were existed in all alignment sequences (Liu *et al.* 2014). The result followed that the overall conservation involved in PBD (α1 and α2 domains) and CD8 interaction sites were observed in *DuMHC I* alleles.

Based on full-length amino acid homology, *MHC class I* from different duck lines could be divided into 13 gene groups (table 1). Although, the similarity of amino acid among the different lines of WS, CV, White Peking ducks and Peking ducks was 83.6–99.4%, 82.5–99.9%, 84.3–85.8% and 84.2–87.8%, respectively, but within the same gene group, the members shared >91% amino acid homology. Among these, 11, 12 and 13 were identified as novel allelic groups. Division of *DuMHC I* gene groups might imply that alleles within one group may be with the same disease resistance and be helpful to conform resistant lines to certain diseases more reasonable.

The amino acid position variability in the PBD of *DuMHC I* was shown in table 2. A total of 13 amino acid positions of UAA*WS and five amino acid positions of UAA*CV showed high scores (≥4) by Wu–Kabat index analysis. Compared with Peking duck *MHC class I* gene (designated UAA*bj in this paper) and chicken BF2, the Wu–Kabat index of amino acid variability revealed that each line not only possessed unique characteristics but also had some common characteristics (Yan *et al.* 2005). For instance, the highest and higher indices of variability were at different positions in different lines: UAA*bj, positions 9, 61, 68, 69, 95 and 111; UAA*WS, positions 9, 24, 66, 93; for UAA*CV, positions 61, 93, 95, 111, 150 had the highest indices and the values were all 4.25. Positions 64 and 128 were the notable positions which were variable only in UAA*WS but invariable in UAA*bj, UAA*CV and BF2. Although, the Wu–Kabat index of position 43 was only five in UAA*WS but appeared higher for that of UAA*bj and UAA*CV. Position 93 was variable in all the three duck lines but the indices were significant difference: nine in UAA*WS, six in UAA*bj and only 4.25 in UAA*CV, respectively. These characteristics of specific locus may be useful for different lines to maintain the species' disease resistance during their evolution. Substitutions with high Wu–Kabat index in the two duck lines were also observed as well as at amino acid positions corresponding to those that can interact with antigenic peptides in chicken and other mammals (Bjorkman and Parham 1990; Pullen *et al.* 1992; Yan *et al.* 2005). At position 9, UAA*WS and BF2 showed the highest indices of variability, meanwhile

Table 2. Wu–Kabat index of the PBD for *BF2* genes, *DuMHC I* genes in the Peking duck and the two duck lines.

Position	BF2	UAA*bj	UAA-WS	UAA-CV
9	10.9	12.1	9	
24	7.4	8	7.5	
27		4.5		
32		6		
34	4.7			
43	4.7		5	
51		4.5		
52		4.5		
53		6		
54		8		
60		4.5		
61	4.2	12.1		4.25
62		6		
64			6.5	
65		6		
66	7.4	9	9	
68		12.1		
69		12.1	5	
73	5.9			
75		6		
76		4.5		
93		6	9	4.25
95		15	5	4.25
97		8		
111		15	5	4.25
113		10		
116		6	5	
121	4.7			
128			5	
130				
146		4.5		
148	5.3			
150[149]		4.5	4	4.25
153[152]		4.5	4	
154[153]		4.5		
156[155]		4.5		

The Wu–Kabat index of BF2 and UAA*bj are calculated based on the alignment of 19 and six cDNA sequences, respectively, that have been submitted to GenBank. The accession numbers are as follows, BF2: AF013491–AF013496, AF231500, AF231502, AY234768, AY234769, L28958, X12780, Z54315–Z54317, Z54321, Z54323, Z54326, Z54330; UAA-bj: AB115241–AB115246.

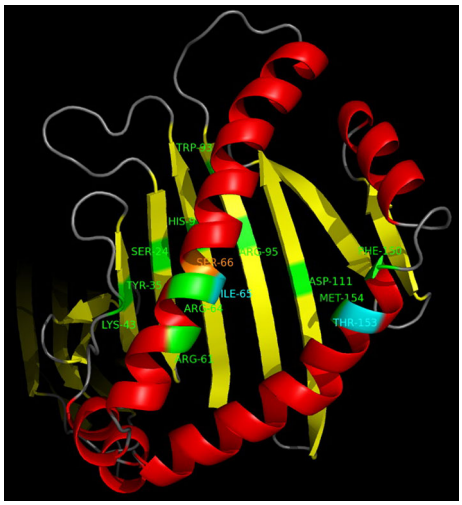

Figure 2. Three-dimensional structure of DuMHC I extracellular domain. The α-helix and β-sheet are indicated by red and yellow, respectively; the turn and radom coli are coloured by gray.

UAA*bj exhibited higher index. The Wu–Kabat index of position 66 was not only high (9) in UAA*WS but also in UAA*bj (9) and BF2 (7.4). There were five variable positions in UAA*CV, the indices of these positions were also higher in the rest of the three loci. The result concluded that the variability of amino acids in the PBD would be the result of immense environmental selection pressures during their evolution and propitious to present more antigenic peptides.

As in figure 2, the homology modelling of all 14 DuMHC I showed very similar structure that contained two α-helix and eight β-sheet in αl and α2 domains, constituted the antigenic peptide groove, and seven β-sheet in α3 domain which was constant in sequence. According to amino acid alignment of *DuMHC I* gene groups, the 14 key variable residues were all found in the α-helix and β-sheet of PBD, which were similar to the key residues interacting with peptide presentation. Among them, 9H, 24S, 35Y, 43K, 93W, 95R and 111D were in the β-sheet, whereas 61R, 64R, 65I, 66S, 150F, 153T, 154M were in the α-helix. The result suggested that different *DuMHC I* allelic groups may possess different spectrums of antigenic peptide presentation.

Although, some researches had verified polymorphism of BF2 were association with disease resistant (Aeed *et al.* 1993; Wallny *et al.* 2006; Dalgaard *et al.* 2009), but no information in duck. Thus, further studies should be carried out to confirm the relationship between polymorphism of *DuMHC I* and disease resistance.

Acknowledgements

This study was cosupported by National Science and Technology Pillar Programme during the 12th period Five-year plan (2015BAD12B03) and the Shandong Academy of Agricultural Science Funds for Young Scholar (2015YQN55).

References

Aeed P. A., Collins W. M., Briles W. E. and Zsigray R. M. 1993 Influence of different B-complex recombinants on the outcome of Rous sarcomas in chickens. *Anim. Genet.* **24**, 177–181.

Bjorkman P. J. and Parham P. 1990 Structure, function, and diversity of class I major histocompatibility complex molecules. *Annu. Rev. Biochem*. **59**, 253–288.

Dalgaard T., Boving M. K., Handberg K., Jensen K. H., Norup L. R. and Juul-Madsen H. R. 2009 MHC expression on spleen lymphocyte subsets in genetically resistant and susceptible chickens infected with Marek's disease virus. *Viral Immunol.* **22**, 321–327.

Garboczi D. N., Ghosh P., Utz U., Fan Q. R., Biddison W. E. and Wiley D. C. 1996 Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**, 134–141.

Liu Y. J., Chen R., Tariq M. and Xia C. 2014 Complex assembly, crystallization and preliminary X-ray crystallographic analysis of the chicken CD8αα-BF2*0401 complex. *Acta Crystallogr. Sect. F struct. Biol. Commun.* **70**, 1264–1267.

Moon D. A., Veniamin S. M., Parks-Dely J. A. and Magor K. E. 2005 The MHC of the duck (*Anas platyrhynchos*) contains five differentially expressed MHC genes. *J. Immunol.* **175**, 6702–6712.

Pullen J. K., Horton R. M., Cai Z. L. and Pease L. R. 1992 Structural diversity of the classical H-2 genes: K, D, and L. *J. Immunol.* **148**, 953–967.

Shum B. P., Rajalingam R., Magor K. E., Azumi K., Carr W. H., Dixon B. *et al.* 1999 A divergent non-classical class I gene conserved in salmonids. *Immunogenetics* **49**, 479–490.

Salter B. D., Norment A. M., Chen B. P., Clayberger C., Krensky A. M., Littman D. R. *et al.* 1989 Polymorphism in the α3 domain of HLA-A molecules affects binding to CD8. *Nature* **338**, 345–347.

Wallny H. J., Avila D., Hunt L. G., Powell T. J., Riegert P., Salomonsen J. *et al.* 2006 Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. *Proc. Natl. Acad. Sci. USA***103**, 1434–1439.

Xia C., Lin C. Y., Xu G. X., Hu T. J. and Yang T. Y. 2004 cDNA cloning and genomic structure of the duck (*Anas platyrhynchos*) MHC class I gene. *Immunogenetics* **56**, 304–309.

Yan R. Q., Li X. S., Yang T. Y. and Xia C. 2005 Characterization of BF2 and β_2m in three Chinese chicken lines. *Vet. Immunol. Immunopathol.* **108**, 417–425.

Received 27 April 2016; revised 18 July 2016; accepted 26 August 2016
Unedited version published online: 29 August 2016
Final version published online: 17 June 2017

Corresponding editor: N. G. Prasad