MDPI

Article

Molecular Characterization of MHC Class I Alpha 1 and 2 Domains in Asian Seabass (*Lates calcarifer*)

Zhixuan Loh [1,2] **, Xuelu Huan** [1] **, Sunita Awate** [3] **, Markus Schrittwieser** [3] **, Laurent Renia** [1,2,4,5]
and Ee Chee Ren [1,6,*]

1 Singapore Immunology Network, Agency for Science, Technology and Research (A*STAR),
   Singapore 138648, Singapore
2 A*STAR Infectious Diseases Labs (A*STAR ID Labs), Agency for Science, Technology and Research (A*STAR),
   Singapore 138648, Singapore
3 UVAXX Pte Ltd., Singapore 159546, Singapore
4 Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 636921, Singapore
5 School of Biological Sciences, Nanyang Technological University, Singapore 637551, Singapore
6 Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore
* Correspondence: ren_ee_chee@immunol.a-star.edu.sg; Tel.: +65-64070004

Abstract: The Asian seabass is of importance both as a farmed and wild animal. With the emergence of infectious diseases, there is a need to understand and characterize the immune system. In humans, the highly polymorphic MHC class I (MHC-I) molecules play an important role in antigen presentation for the adaptive immune system. In the present study, we characterized a single MHC-I gene in Asian seabass (*Lates calcarifer*) by amplifying and sequencing the MHC-I alpha 1 and alpha 2 domains, followed by multi-sequence alignment analyses. The results indicated that the Asian seabass MHC-I α1 and α2 domain sequences showed an overall similarity within Asian seabass and retained the majority of the conserved binding residues of human leukocyte antigen-A2 (HLA-A2). Phylogenetic tree analysis revealed that the sequences belonged to the U lineage. Mapping the conserved binding residue positions on human HLA-A2 and grass carp crystal structure showed a high degree of similarity. In conclusion, the availability of MHC-I α1 and α2 sequences enhances the quality of MHC class I genetic information in Asian seabass, providing new tools to analyze fish immune responses to pathogen infections, and will be applicable in the study of the phylogeny and the evolution of antigen-specific receptors.

Keywords: Asian seabass; *Lates calcarifer*; MHC class I; sequence variability; polymorphism

1. Introduction

The Asian seabass (*Lates calcarifer,* Bloch 1790), also known as barramundi, is a commercially important fish species that is widely distributed in the Indo-Pacific Asia region, including Singapore and Northern Australia [1,2]. As the Asian seabass industry expands and production increases, commercial farmers have been battling emerging infectious diseases that affect the survival of farmed livestock [3–8]. Therefore, it is important to understand and characterize the immune system in order to combat this over-arching problem.

Classical major histocompatibility complex class I (MHC-I) molecules play pivotal roles in initiating an adaptive immune response against pathogens [9,10]. Structurally, the classical MHC-I molecules consist of three extracellular domains, where the α1 and α2 domains are involved in peptide binding, and the membrane α3 domain interacts with β2-microglobulin and provides molecular stability [11]. A major feature of these classical MHC-I molecules is the extensive alleles polymorphism predominantly found within the peptide binding α1 and α2 domains to help ensure the survival of the population by displaying a broad range of peptides to T cell receptors [12,13].

Currently, six MHC-I lineages, U, Z, S, L, P, and H, have been described in teleost fish [14–19]. These lineages are defined by polymorphisms, expression patterns, and peptide binding ability. Among the lineages, only the U lineage has been well studied at the molecular and functional levels due to its binding abilities [20]. While little is known about the nonclassical MHC-I lineages Z, S, L, P, and H, some studies have suggested that the Z lineage may bind to peptides with modification to its N-terminal [17,21].

In Asian seabass, predicted MHC-I sequences have been obtained from a single Asian seabass using next-generation sequencing [22]. However, the lineages of these sequences have not been validated, and the polymorphisms in MHC-I molecules have not been explored. Therefore, in the present study, we sequenced and characterized the peptide binding regions of a single MHC-I gene in Asian seabass obtained from Singapore and Australia.

2. Results

2.1. Amino Acid Sequences of Asian Seabass MHC-I α1 and α2 Domains

The classical MHC-I heavy chain gene is organized into eight exons with distinct functional domains [23,24]. To sequence the peptide-binding regions formed by the α1 and α2 domains coded by exon 2 and exon 3, respectively, RNA was isolated and cDNA was generated from the muscle tissues of Asian seabass. Exon 2 and exon 3 were amplified using PCR with primers designed on exon 1 and exon 4. The nucleotide sequences obtained from 12 Asian seabass (9 Asian seabass from Singapore and 3 Asian seabass from Australia) (Figure S1) were translated using the EMBOSS Transeq tool into protein sequences before aligning using Cluster Omega [25]. The amplified α1 domain (exon 2) comprises 87 amino acids and the α2 domain (exon 3) comprises 93 amino acids (Figure 1).

Comparing Lca17733 and the sequences from the 12 Asian seabass, our results indicated that the amino acid sequence variation among the α1 domain with pairwise comparisons ranged from 12.64 to 20.69%, with $\pm$s.e.m. of 9.84 $\pm$ 0.522%, and amino acid sequence variation among the α2 domain with pair-wise comparisons ranged from 6.45 to 26.88%, with $\pm$s.e.m. 19.17 $\pm$ 1.04%. This suggests that the sequences of the MHC-I α1 domain were more conserved as compared with those of the MHC-I α2 domain (Figure 2a). Furthermore, analysis using the variability metric (V), where $V > 1$ indicates highly polymorphic sites, identified 4 sites in the α1 domain and 16 sites in the α2 domain (Figure 2b). Genetic differentiation found between Asian seabass obtained from Australia and Singapore was not observed (Figure 2a).

The MHC-I α1 and α2 domain sequences were then compared with representative sequences of other MHC-I lineages found in teleost fishes and with the human MHC-I molecules HLA-A2, HLA-B15, and HLA-C1 (Figure 1). The highly conserved eight residues Y7, Y59, Y/R84, T143, K146, W147, Y159, and Y171, thought to be important to anchor the N- and C-terminal peptide ends [26–28], are highlighted in red (Figure 1). The Asian seabass MHC-I in all 12 fishes possessed all of the eight peptide binding termini residues, including residue R84, which is common in non-mammalian vertebrates such as sharks and teleost fish, instead of Y84, found in mammals and some reptiles [27,29] (Figure 1). Similar residues that were (1) probably found before the molecular evolution of MHC classes I and II (indicated in black) [30,31] or (2) common in classical MHC-I (indicated in blue) [19,30] were also found in the Asian seabass MHC-I α1 and α2 domain sequences, albeit with <11 substitutions out of the 70 highlighted conserved residues (Figure 1). Based on the length and pattern expression, the Asian seabass MHC-I showed an overall similarity to classical human MHC-I molecules and may have a similar structure.
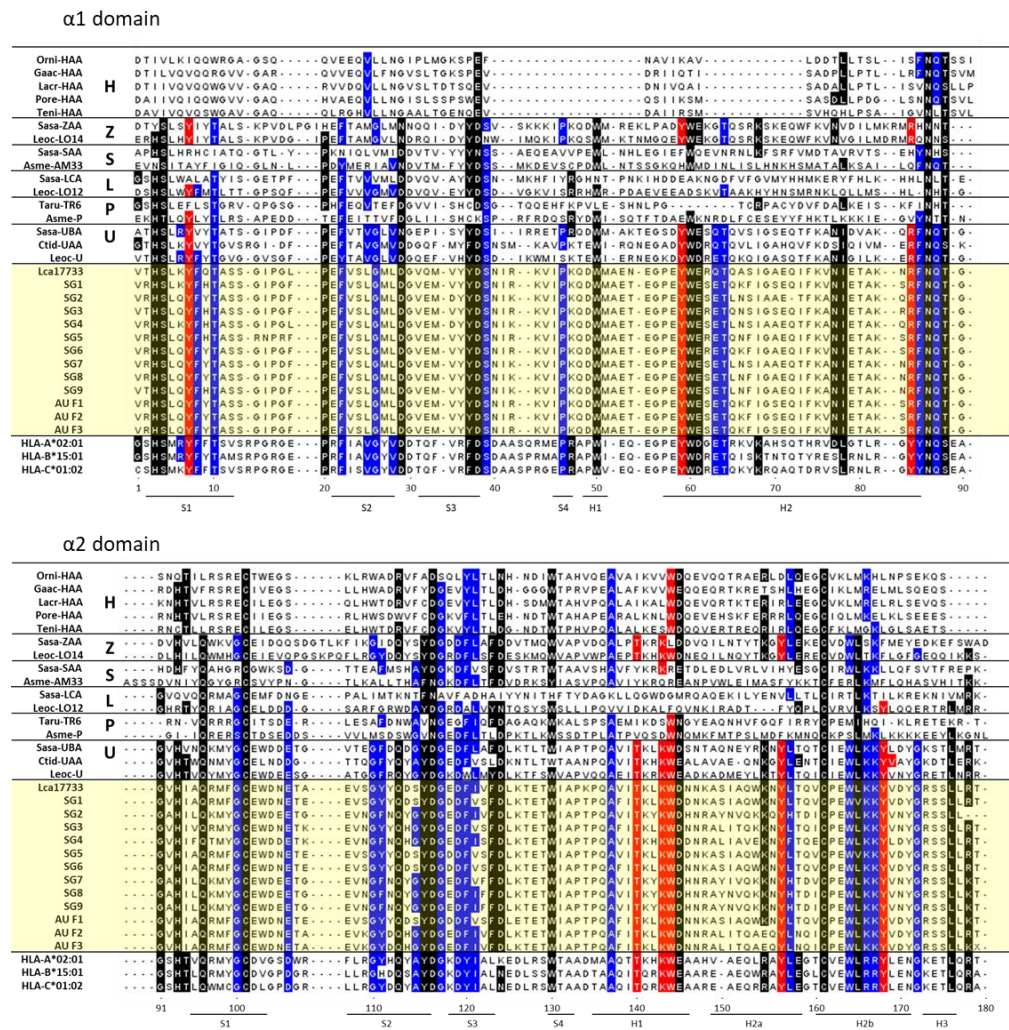
Figure 1. Alignment of Asian seabass MHC-I α1 and α2 amino acid sequences with other representative MHC-I sequences. Multiple sequence alignment of Asian amino acid sequences with other representative fish MHC-I and human HLA sequences. Black indicates similar residues that were probably found before the molecular evolution of MHC classes I and II, blue indicates similar residues that are common in classical MHC-I, and red indicates conserved binding residues in classical MHC-I. The sequences are divided into α1 and α2 regions, which correspond to exon 2 and exon 3, respectively. The residue positions indicated below the alignment are based on the HLA-A2 protein. Structural indications S denotes the β-strand and H denotes the helix based on the pHLA-A2 structure (PDB database accession 3PWN). The sequence references are as follows: Orni (Oreochromis niloticus, Nile tilapia) -HAA (TSA: GBAZ01123113); Gaac (Gasterosteus aculeatus, Stickleback) -HAA (DW655318); Lacr (Larimichthys crocea, Yellow croaker) –HAA (XP_010741942.1); Pore (Poecilia reticulata, Guppy) -HAA (XP_008432358.1 and TSA: GFHH01045885); Teni (Tetraodon nigroviridis, Tetraodon) -HAA (CAG07665.1); Sasa (Salmo salar, Atlantic salmon) -SAA (ACY30362.1), -ZAAa (ACX35596.1), and -UBA (*0301, XP_014032819); Leoc (Lepisosteus oculatus, Spotted gar) -LO14 (TSA: GFIM01040660), -LO12 (JH591577:52,184-56,541), and -U (TSA: GFIM01032149); Ctid (Ctenopharyngodon idella, Grass carp) -UAA (5H5Z-A); Asme (Astyanax mexicanus, Mexican tetra) -AM33 (ENSAMXG00000017444) and -P (TSA: GFIF01000014); Taru (Takifugu rubripes, Fugu) -TR16 (Scaffold_497:44,782-49,340); HLA (Human leukocyte antigen) -A2 (AAA76608.2); -B15:01 (HG794370.1), -C1:02 (HG794388.1), and Lca17733 (Lates calcarifer, Asian seabass) (XM_018684417). SG denotes samples obtained from Singapore and AU denotes samples obtained from Australia.
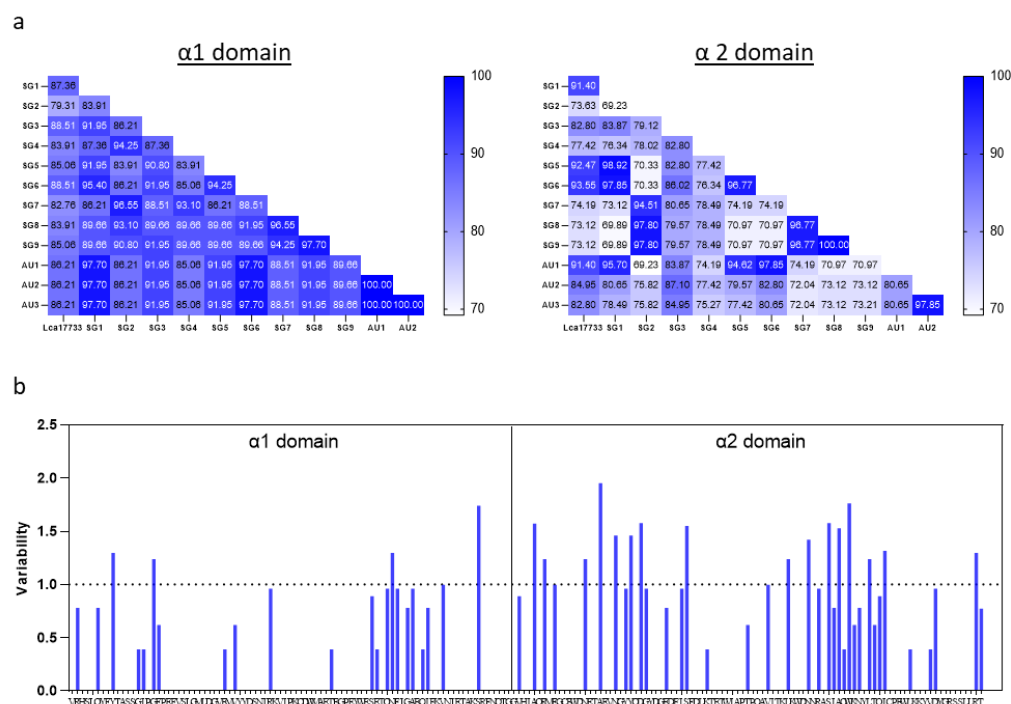
Figure 2. Amino acid identity and sequence variability within the Asian seabass α1 and α2 domains. (**a**) Heatmap of amino acid similarity values for Asian seabass and Lca17733 MHC-I α1 and α2 domains. (**b**) Sequence variability calculated using the variability metric (*V*) analysis based on the Shannon entropy equation. Sites with *V* >1 are considered as highly polymorphic in the sequence alignment of the Asian seabass and Lca17733 MHC-I α1 and α2 domains.

2.2. Phylogenetic Analyses of Asian Seabass MHC-I α1 and α2 Domains

Previous studies have described six MHC-I lineages in teleosts denoted by U, Z, S, L, P, and H. Using the Asian seabass α1 and α2 domain sequences (Figure 2), the sequences were used to draw a phylogenetic tree with various teleost fishes' and human HLA. With bootstrap values ranging from 24–100%, the MHC-I lineage sequences form separate clusters. Phylogenetic tree analyses showed that the Asian seabass MHC-I sequences shared an overall specific similarity with the U sequences, forming a distinct cluster (Figure 3 and Figure S2).

2.3. Protein Modeling

To date, the Asian seabass MHC-I crystal structure remains undefined. To understand how the Asian seabass binding residues map onto the classical MHC-I molecules, two high-resolution human and teleost fish crystal structures were selected for model building: HLA-A*02:01 structure (PDB code 3MRE) [36] and grass carp MHC-I (Ctid-UAA) structure (PDB code 5H5Z) [37]. The HLA-A2 family is the largest allele family of the HLA-A locus, with HLA-A*02:01 being the most predominant allele present in 50% of the population in the United States [38], while Ctid-UAA is the only teleost fish with available MHC-I structure [37]. The overlapping of the structures of HLA-A2 (pink) and Ctid-UAA (grey) shows that both molecules have similar structures (Figure S3). The conserved residues from the Asian seabass sequences were then highlighted on HLA-A2 and Ctid-UAA structures (Figure 4). The schematic indicates that the binding residues (red) of the Asian seabass MHC-I sequence were the same in both HLA-A2 and Ctid-UAA, with slight differences in the view from the top and the PΩ side of residues that were found before the molecular evolution of the MHC classes I and II (black) or that are common in classical MHC-I (blue) (Figure 4).
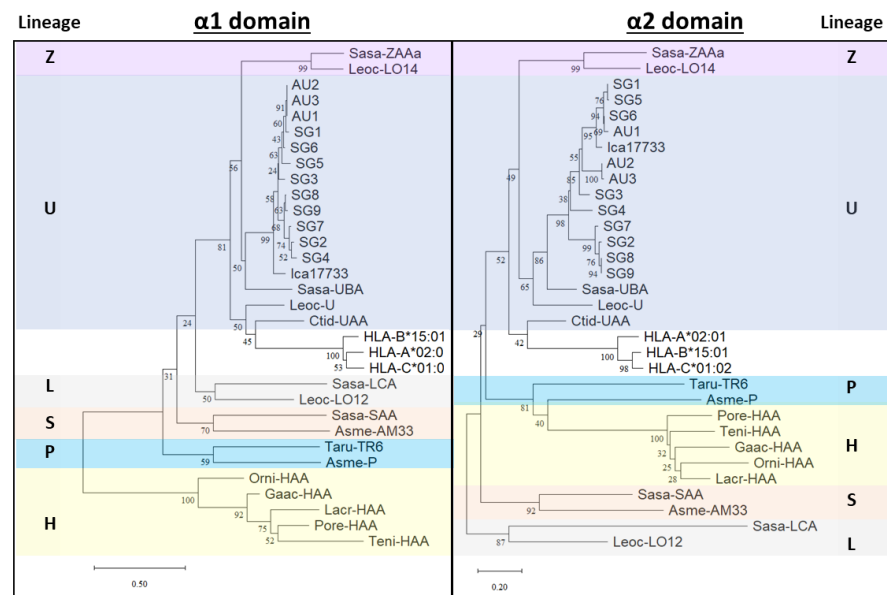
Figure 3. Phylogenetic trees based on α1 and α2 domain amino acid sequences of Asian seabass and other representative teleost fishes' MHC-I molecules. Neighbor-joining tree [32] of α1 (left tree) and α2 (right tree) domain sequences. The trees are drawn to scale, with branch lengths indicating the number of amino acid substitutions per site. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) [33] is shown on each node. The evolutionary distances were computed using the Poisson correction method [34] and are presented in the unit of the number of amino acid substitutions per site. The analysis involved 32 amino acid sequences. There were 99 positions in each of the final datasets. Evolutionary analyses were conducted in MEGA11 [35]. Clusters with the six different MHC class I lineages found in teleosts are shown with colored boxes. The sequence references are listed in Figure 1.
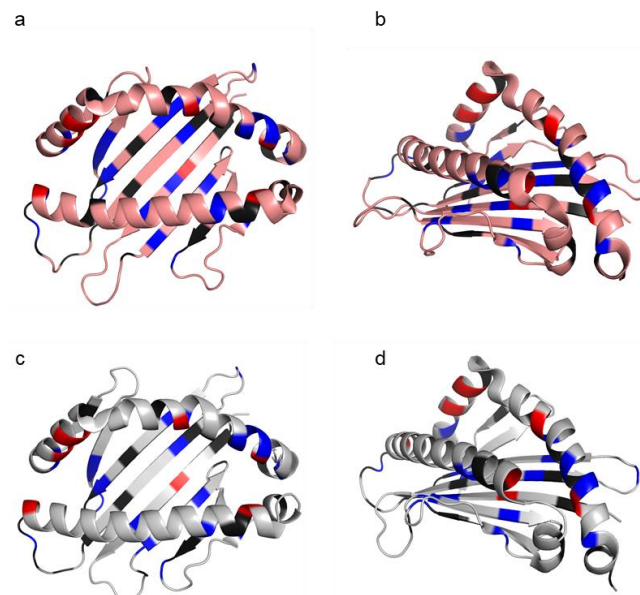
Figure 4. Schematic crystal structure of the human HLA-A*02:01 (PDB code 3MRE) and Ctid-UAA (PDB code 5H5Z) peptide binding groove defined by the α1 and α 2 domains. HLA-A*02:01 structure (**a**,**b**, pink) and Ctid-UAA structure (**c**,**d**, grey) with similar residues from the sequence of representative Asian seabass MHC-I residues mapped viewed from the top (**a**,**c**) and the PΩ side (**b**,**d**). Black indicates similar residues that were probably found before the molecular evolution of MHC classes I and II, blue indicates similar residues that are common in classical MHC-I, and red indicates conserved binding residues in classical MHC-I.

3. Discussion

MHC-I is an important immune molecule that is made up of highly polymorphic genes to enhance the recognition of pathogens aiding the resistance of hosts to infectious diseases. Often, MHC-I displays significant differentiation across geographical environments and populations, such as fishes [39–44]. Therefore, to characterize the MHC-I binding domains in Asian seabass, we randomly sampled 12 individual fishes. Comparing the Asian seabass MHC-I α1 and α2 domains, the sequence is more conserved where we identified 4 highly polymorphic sites with $V > 1$ in the α1 domain versus 16 polymorphic sites with $V > 1$ in the α2 domain. In contrast, sequence analyses of HLA-A revealed 6 highly polymorphic sites with $V > 1$ in the α1 domain and 7 polymorphic sites with $V > 1$ in the α2 domain from the 230 sequences obtained in the IMGT/HLA database [45,46].

Regardless of the genetic heterogeneity, the α1 and α2 domains consisted of eight evolutionary conserved residues that anchored the peptide [11,27]. The set of residues in mammals is YYYYYTKW, while that in non-mammalian vertebrates such as fishes is slightly different, YYYRFTKW [18,26]. Indeed, the sequence of the Asian seabass MHC-I contains all eight conserved binding residues, including R84. Further analysis of residues that were either present before the molecular evolution of MHC classes I and II or residues that were common to MHC-I revealed that the Asian seabass MHC-I possesses most of these residues. Moreover, by mapping these residues onto the HLA-A2 and Ctid-UAA crystal structures, we suggest that the Asian seabass MHC-I gene here is likely to function similarly to the classical MHC-I.

Phylogenetic analyses using neighbor-joining trees and maximum likelihood trees based on α1 and α2 domain amino acids suggest that the Asian seabass MHC-I gene sequences clustered close to the U lineages from different fish species. U lineage genes are predicted to provide the classical function of antigen presentation, which explains why the conserved residues were mostly found in these sequences. Interestingly, the Asian seabass MHC-I sequences consistently group with salmonid alleles and may indicate trans-species polymorphisms, which is common for MHC genes [47]. While these observations must be interpreted with caution as the bootstrapping support was relatively low, similar findings were reported in a recent study on the MHC-I alpha 1 domain in European seabass (*Dicentrarchus labrax*) [48].

In conclusion, we characterized a single MHC-I gene that belongs to the U lineage in Asian seabass. Currently, other than the MHC-I sequences that were obtained from an Asian seabass using next-generation sequencing, there is no known information regarding the polymorphism within its species and the genes have not been validated by another study. The availability of MHC-I α1 and α2 sequences enhances the quality of MHC class I genetic information in Asian seabass, provides new tools to analyze Asian seabass immunity against pathogenic infections, and will be applicable for phylogeny and evolution studies of antigen-specific receptors. Furthermore, the investigation of the MHC-I α1 and α2 sequences will help to uncover other polymorphisms at the interaction surfaces in Asian seabass. Future directions would include the search for experimental testing to generate three-dimensional structures and assess antigen binding interactions between MHC-I and T cell receptor molecules to modulate Asian seabass immune responses.

4. Materials and Methods

4.1. Materials

The genomes used in this study are as follows: Sasa-ZAAa *Salmo salar* (Atlantic salmon; accession ACX35596.1) [49], Leoc-LO14 *Lepisosteus oculatus* (spotted gar; accession GFIM01040660) [21,50], Sasa-SAA Salmo *salar* (Atlantic salmon; accession ACY30362.1) [49], Asme-AM33 *Astyanax mexicanus* (Mexican tetra; accession KB882192.1) [21], Sasa-LCA *Salmo salar* (Atlantic salmon, accession XP_013983104) [21,51], *Leoc-LO12 Lepisosteus oculatus* (spotted gar; accession JH591577) [21], Taru-TR6 *Takifugu rubripes* (fugu; accession HE591930) [21,52], Asme-P *Astyanax mexicanus* (Mexican tetra; accession GFIF01000014) [50], Sasa-UBA *Salmo salar* (Atlantic salmon; accession XP_014032819.1) [51], Leoc-U *Lepisosteus*

oculatus (spotted gar; accession GFIM01032149) [50], Ctid (*Ctenopharyngodon idella*, Grass carp) -UAA (5H5Z-A) [37], HLA-A2:01 *Homo sapiens* (human; accession AAA76608.2) [53], HLA-B15:01 *Homo sapiens* (human; accession HG794370.1), HLA-C1:02 *Homo sapiens* (human; accession HG794388.1), and the predicted *Lates calcarifer* class I histocompatibility antigen, F10 alpha chain-like (Asian seabass; accession XM_018684417) [22].

4.2. Data Mining

Lates calcarifer MHC-I gene sequences were identified using known bony fish and human MHC gene sequences [21,49–52] and tblastn against annotated databases available in the Ensembl and NCBI databases.

4.3. Fish Samples

Farmed Asian seabass, SG1-SG9 and AU1-AU3, were collected from a commercial Singapore farm and MainStream Australia, respectively. The fishes sampled were between 40 and 119 days old, with weights ranging from ~2 g to 61 g and lengths ranging from ~2 cm to 15 cm.

4.4. Experimental Analyses of Asian Seabass Mhc-I α1 and α2 Domains Sequence

Total RNA was isolated from Lates calcarifer (Asian seabass) muscle tissues using an Rneasy mini kit (Qiagen, Hilden, Germany) and transcribed into cDNA using oligo (dT) with a SuperScript III First-Strand Synthesis system (Invitrogen, Waltham, MA, USA) as per the manufacturer's instructions. The coding sequence of the Asian seabass MHC-1 α1 and α2 domain genes was amplified by Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Scientific, Waltham, MA, USA) using a forward primer, LCA-MHC-1-F1 (5′-TTCCTGTTCCTGTGCCACG-3′), and a reverse primer, LCA-MHC-1-R1 (5′-AGAGACACTGAGGGAAGGGC-3′), designed at exon 1 and exon 4, respectively, of the RefSeq Genome sequencing and assembly PRJNA345597. The PCR reaction contained 10 μL of 5X Phusion HF buffer, 200 μM of dNTPs, 0.5 μM of each primer, 0.5 μL of Phusion Hot Start II DNA polymerase, and 5 μL of cDNA in a total volume of 50 μL. The PCR conditions were denaturation at 98 °C for 1 min, 35 cycles of denaturation at 98 °C for 15 s, annealing at 60 °C for 30 s, elongation at 72 °C for 30 s, and final elongation at 72 °C for 7 min. The PCR product size was determined using gel electrophoresis (~700 bp; Figure S4), and the PCR products were purified with a Monarch® PCR & DNA Cleanup kit (New England Biolabs, Ipswich, MA, USA). DNA sequencing was performed by Bio Basic Asia Pacific Pte Ltd. (Singapore) using the forward primer LCA-MHC-1-F1 (5′-TTCCTGTTCCTGTGCCACG-3′) and reverse primer LCA-MHC-1-R2 (5′-TCCCAGTACTCAGGTCCTTC-3′). The 12 sequences obtained in this study were deposited at GenBank as accessions OP348933 to OP348944, respectively.

4.5. Sequence Alignments, Variability, and Phylogenetic Analysis

MHC-I amino acid sequences were aligned using Clustal Omega [54] for initial analyses and manually curated subsequently in Jalview [55] for some predicted sequences. Pairwise alignment was conducted using Jalview [55] and plotted as heatmaps with the mean ± s.e. calculated using Graphpad Prism 9. The MHC-I α1 and α2 sequence variability and polymorphisms were determined using variability metric (*V*) analysis [46]. The variability metric is based on the Shannon entropy equation and calculated as $V = -\sum_{t=1}^{M} P_i log_2 P_i$, where P_i is the fraction of residues of amino acid type *i* and *M* is the number of amino acid types [46]. *V* ranges from 0 (most conserved site with only one amino acid type at that position) to 3.7 (most variable site with all 13 amino acids represented differently at that position). Sites with *V* > 1 are considered to be highly polymorphic. For phylogenetic tree analysis, the α1 and α2 domains were inferred using the neighbor-joining method [32], with bootstrap testing according to Felsenstein [33]. The evolutionary distances were computed using the Poisson correction method and were presented in the unit of the number of amino acid substitutions per site [34]. All ambiguous positions were removed for each sequence

pair. There were 99 positions in each of the final datasets. The lineage clustering was further supported by another phylogenetic tree constructed using the maximum likelihood method based on the optimal JTT matrix-based model [56]. Evolutionary analyses were conducted in MEGA11 [35].

4.6. Three-Dimensional Modeling of Human HLA-A2-01 and Grass Carp MHC-I α1 and α2 Domains

The structures of HLA-A*02:01 (PDB code 3MRE) [36] and Ctid-UAA (PDB code 5H5Z) [37] were studied and contrasted using The PyMol Molecular Graphics System, Version 2.1.1 Schrödinger, LLC, New York, NY, USA.

References

1. Nelson, J.S.; Grande, T.C.; Wilson, M.V. *Fishes of the World*; John Wiley & Sons: Hoboken, NJ, USA, 2016.
2. Mathew, G. Taxonomy, Identification and Biology of Seabass (*Lates calcarifer*). In *Course Manual: National Training on Cage Culture of Seabass*; CMFRI & NFDB: Kochi, India, 2009.
3. Gibson-Kueh, S.; Chee, D.; Chen, J.; Wang, Y.H.; Tay, S.; Leong, L.N.; Ng, M.L.; Jones, J.B.; Nicholls, P.K.; Ferguson, H.W. The pathology of 'scale drop syndrome' in Asian seabass, *Lates calcarifer* Bloch, a first description. *J. Fish Dis.* **2012**, *35*, 19–27. [CrossRef] [PubMed]
4. De Groof, A.; Guelen, L.; Deijs, M.; van der Wal, Y.; Miyata, M.; Ng, K.S.; van Grinsven, L.; Simmelink, B.; Biermann, Y.; Grisez, L.; et al. A Novel Virus Causes Scale Drop Disease in *Lates calcarifer*. *PLoS Pathog.* **2015**, *11*, e1005074. [CrossRef] [PubMed]
5. Dong, H.T.; Jitrakorn, S.; Kayansamruaj, P.; Pirarat, N.; Rodkhum, C.; Rattanarojpong, T.; Senapin, S.; Saksmerprome, V. Infectious spleen and kidney necrosis disease (ISKND) outbreaks in farmed barramundi (*Lates calcarifer*) in Vietnam. *Fish Shellfish Immunol.* **2017**, *68*, 65–73. [CrossRef] [PubMed]
6. Chang, S.F.; Ng, K.S.; Grisez, L.; de Groof, A.; Vogels, W.; van der Hoek, L.; Deijs, M. Novel Fish Pathogenic Virus. International Patent Application No. WO2018/029301A1, 15 February 2018.
7. Chen, J.; Toh, X.; Ong, J.; Wang, Y.; Teo, X.H.; Lee, B.; Wong, P.S.; Khor, D.; Chong, S.M.; Chee, D.; et al. Detection and characterization of a novel marine birnavirus isolated from Asian seabass in Singapore. *Virol. J.* **2019**, *16*, 71. [CrossRef]
8. Girisha, S.; Puneeth, T.; Nithin, M.; Kumar, B.N.; Ajay, S.; Vinay, T.; Suresh, T.; Venugopal, M.; Ramesh, K. Red sea bream iridovirus disease (RSIVD) outbreak in Asian seabass (*Lates calcarifer*) cultured in open estuarine cages along the west coast of India: First report. *Aquaculture* **2020**, *520*, 734712. [CrossRef]

9.   Germain, R.N. MHC-dependent antigen processing and peptide presentation: Providing ligands for T lymphocyte activation. *Cell* **1994**, *76*, 287–299. [CrossRef]

10.   Neefjes, J.; Jongsma, M.L.; Paul, P.; Bakke, O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* **2011**, *11*, 823–836. [CrossRef]

11.   Madden, D.R. The three-dimensional structure of peptide-MHC complexes. *Annu. Rev. Immunol.* **1995**, *13*, 587–622. [CrossRef]

12.   Mungall, A.J.; Palmer, S.A.; Sims, S.K.; Edwards, C.A.; Ashurst, J.L.; Wilming, L.; Jones, M.C.; Horton, R.; Hunt, S.E.; Scott, C.E.; et al. The DNA sequence and analysis of human chromosome 6. *Nature* **2003**, *425*, 805–811. [CrossRef]

13.   Norman, P.J.; Norberg, S.J.; Guethlein, L.A.; Nemat-Gorgani, N.; Royce, T.; Wroblewski, E.E.; Dunn, T.; Mann, T.; Alicata, C.; Hollenbach, J.A.; et al. Sequences of 95 human MHC haplotypes reveal extreme coding variation in genes other than highly polymorphic HLA class I and II. *Genome Res.* **2017**, *27*, 813–823. [CrossRef]

14.   Dijkstra, J.M.; Katagiri, T.; Hosomichi, K.; Yanagiya, K.; Inoko, H.; Ototake, M.; Aoki, T.; Hashimoto, K.; Shiina, T. A third broad lineage of major histocompatibility complex (MHC) class I in teleost fish; MHC class II linkage and processed genes. *Immunogenetics* **2007**, *59*, 305–321. [CrossRef] [PubMed]

15.   Grimholt, U.; Hordvik, I.; Fosse, V.M.; Olsaker, I.; Endresen, C.; Lie, O. Molecular cloning of major histocompatibility complex class I cDNAs from Atlantic salmon (*Salmo salar*). *Immunogenetics* **1993**, *37*, 469–473. [CrossRef] [PubMed]

16.   Hashimoto, K.; Nakanishi, T.; Kurosawa, Y. Isolation of carp genes encoding major histocompatibility complex antigens. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6863–6867. [CrossRef] [PubMed]

17.   Kruiswijk, C.P.; Hermsen, T.T.; Westphal, A.H.; Savelkoul, H.F.; Stet, R.J. A novel functional class I lineage in zebrafish (*Danio rerio*), carp (*Cyprinus carpio*), and large barbus (*Barbus intermedius*) showing an unusual conservation of the peptide binding domains. *J. Immunol.* **2002**, *169*, 1936–1947. [CrossRef]

18.   Shum, B.P.; Rajalingam, R.; Magor, K.E.; Azumi, K.; Carr, W.H.; Dixon, B.; Stet, R.J.; Adkison, M.A.; Hedrick, R.P.; Parham, P. A divergent non-classical class I gene conserved in salmonids. *Immunogenetics* **1999**, *49*, 479–490. [CrossRef]

19.   Grimholt, U.; Tsukamoto, K.; Hashimoto, K.; Dijkstra, J.M. Discovery of a Novel MHC Class I Lineage in Teleost Fish which Shows Unprecedented Levels of Ectodomain Deterioration while Possessing an Impressive Cytoplasmic Tail Motif. *Cells* **2019**, *8*, 1056. [CrossRef]

20.   Yamaguchi, T.; Dijkstra, J.M. Major Histocompatibility Complex (MHC) Genes and Disease Resistance in Fish. *Cells* **2019**, *8*, 378. [CrossRef]

21.   Grimholt, U.; Tsukamoto, K.; Azuma, T.; Leong, J.; Koop, B.F.; Dijkstra, J.M. A comprehensive analysis of teleost MHC class I sequences. *BMC Evol. Biol.* **2015**, *15*, 32. [CrossRef]

22.   Vij, S.; Kuhl, H.; Kuznetsova, I.S.; Komissarov, A.; Yurchenko, A.A.; Van Heusden, P.; Singh, S.; Thevasagayam, N.M.; Prakki, S.R.; Purushothaman, K.; et al. Chromosomal-Level Assembly of the Asian Seabass Genome Using Long Sequence Reads and Multi-layered Scaffolding. *PLoS Genet.* **2016**, *12*, e1005954. [CrossRef]

23.   Schellens, I.M.; Hoof, I.; Meiring, H.D.; Spijkers, S.N.; Poelen, M.C.; van Gaans-van den Brink, J.A.; van der Poel, K.; Costa, A.I.; van Els, C.A.; van Baarle, D.; et al. Comprehensive Analysis of the Naturally Processed Peptide Repertoire: Differences between HLA-A and B in the Immunopeptidome. *PLoS ONE* **2015**, *10*, e0136417. [CrossRef]

24.   Rammensee, H.G.; Friede, T.; Stevanoviíc, S. MHC ligands and peptide motifs: First listing. *Immunogenetics* **1995**, *41*, 178–228. [CrossRef] [PubMed]

25.   Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [CrossRef] [PubMed]

26.   Hashimoto, K.; Okamura, K.; Yamaguchi, H.; Ototake, M.; Nakanishi, T.; Kurosawa, Y. Conservation and diversification of MHC class I and its related molecules in vertebrates. *Immunol. Rev.* **1999**, *167*, 81–100. [CrossRef] [PubMed]

27.   Kaufman, J.; Salomonsen, J.; Flajnik, M. Evolutionary conservation of MHC class I and class II molecules—Different yet the same. *Semin. Immunol.* **1994**, *6*, 411–424. [CrossRef]

28.   Saper, M.A.; Bjorkman, P.J.; Wiley, D.C. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *J. Mol. Biol.* **1991**, *219*, 277–319. [CrossRef]

29.   Wu, Y.; Zhang, N.; Wei, X.; Lu, S.; Li, S.; Hashimoto, K.; Dijkstra, J.M.; Xia, C. The Structure of a Peptide-Loaded Shark MHC Class I Molecule Reveals Features of the Binding between β_2-Microgulobulin and H Chain Conserved in Evolution. *J. Immunol.* **2021**, *207*, 308–321. [CrossRef]

30.   Dijkstra, J.M.; Yamaguchi, T.; Grimholt, U. Conservation of sequence motifs suggests that the nonclassical MHC class I lineages CD1/PROCR and UT were established before the emergence of tetrapod species. *Immunogenetics* **2018**, *70*, 459–476. [CrossRef]

31.   Kaufman, J. Vertebrates and the evolution of the major histocompatibility complex (MHC) class I and class II molecules. *Verh. Dtsch. Zool. Ges.* **1988**, *81*, 131–144.

32.   Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef]

33.   Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [CrossRef]

34.   Zuckerkandl, E.; Pauling, L. Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins*; Bryson, V., Vogel, H.J., Eds.; Academic Press: New York, NY, USA, 1965.

35. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. [CrossRef] [PubMed]
36. Reiser, J.B.; Legoux, F.; Gras, S.; Trudel, E.; Chouquet, A.; Léger, A.; Le Gorrec, M.; Machillot, P.; Bonneville, M.; Saulquin, X.; et al. Analysis of relationships between peptide/MHC structural features and naive T cell frequency in humans. *J. Immunol.* **2014**, *193*, 5816–5826. [CrossRef] [PubMed]
37. Li, Z.; Zhang, N.; Ma, L.; Zhang, L.; Meng, G.; Xia, C. The Mechanism of β2m Molecule-Induced Changes in the Peptide Presentation Profile in a Bony Fish. *iScience* **2020**, *23*, 101119. [CrossRef]
38. Ellis, J.M.; Henson, V.; Slack, R.; Ng, J.; Hartzman, R.J.; Katovich Hurley, C. Frequencies of HLA-A2 alleles in five U.S. population groups. Predominance Of A*02011 and identification of HLA-A*0231. *Hum. Immunol.* **2000**, *61*, 334–340. [CrossRef]
39. Van Tienderen, P.H.; de Haan, A.A.; van der Linden, C.G.; Vosman, B. Biodiversity assessment using markers for ecologically important traits. *Trends Ecol. Evol.* **2002**, *17*, 577–582. [CrossRef]
40. Sommer, S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front. Zool.* **2005**, *2*, 16. [CrossRef]
41. Miller, K.M.; Withler, R.E. Mhc diversity in Pacific salmon: Population structure and trans-species allelism. *Hereditas* **1997**, *127*, 83–95. [CrossRef]
42. Landry, C.; Bernatchez, L. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Mol. Ecol.* **2001**, *10*, 2525–2539. [CrossRef]
43. Miller, K.M.; Kaukinen, K.H.; Beacham, T.D.; Withler, R.E. Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon. *Genetica* **2001**, *111*, 237–257. [CrossRef]
44. Miller, K.M.; Winton, J.R.; Schulze, A.D.; Purcell, M.K.; Ming, T.J. Major histocompatibility complex loci are associated with susceptibility of Atlantic salmon to infectious hematopoietic necrosis virus. In *Genetics of Subpolar Fish and Invertebrates*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 307–316.
45. Robinson, J.; Malik, A.; Parham, P.; Bodmer, J.G.; Marsh, S.G. IMGT/HLA database—A sequence database for the human major histocompatibility complex. *Tissue Antigens* **2000**, *55*, 280–287. [CrossRef]
46. Reche, P.A.; Reinherz, E.L. Sequence variability analysis of human class I and class II MHC molecules: Functional and structural correlates of amino acid polymorphisms. *J. Mol. Biol.* **2003**, *331*, 623–641. [CrossRef]
47. Klein, J.; Sato, A.; Nagl, S.; O'hUigín, C. Molecular trans-species polymorphism. *Annu. Rev. Ecol. Syst.* **1998**, *29*, 1–21+C1. [CrossRef]
48. Ratcliffe, F.C.; Garcia de Leaniz, C.; Consuegra, S. MHC class I-α population differentiation in a commercial fish, the European sea bass (*Dicentrarchus labrax*). *Anim. Genet.* **2022**, *53*, 340–351. [CrossRef] [PubMed]
49. Lukacs, M.F.; Harstad, H.; Bakke, H.G.; Beetz-Sargent, M.; McKinnel, L.; Lubieniecki, K.P.; Koop, B.F.; Grimholt, U. Comprehensive analysis of MHC class I genes from the U-, S-, and Z-lineages in Atlantic salmon. *BMC Genom.* **2010**, *11*, 154. [CrossRef] [PubMed]
50. Pasquier, J.; Cabau, C.; Nguyen, T.; Jouanno, E.; Severac, D.; Braasch, I.; Journot, L.; Pontarotti, P.; Klopp, C.; Postlethwait, J.H.; et al. Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database. *BMC Genom.* **2016**, *17*, 368. [CrossRef]
51. Davidson, W.S.; Koop, B.F.; Jones, S.J.; Iturra, P.; Vidal, R.; Maass, A.; Jonassen, I.; Lien, S.; Omholt, S.W. Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biol.* **2010**, *11*, 403. [CrossRef]
52. Kai, W.; Kikuchi, K.; Tohari, S.; Chew, A.K.; Tay, A.; Fujiwara, A.; Hosoya, S.; Suetake, H.; Naruse, K.; Brenner, S.; et al. Integration of the genetic map and genome assembly of fugu facilitates insights into distinct features of genome evolution in teleosts and mammals. *Genome Biol. Evol.* **2011**, *3*, 424–442. [CrossRef]
53. Balas, A.; García-Sánchez, F.; Gómez-Reino, F.; Vicario, J.L. HLA class I allele (HLA-A2) expression defect associated with a mutation in its enhancer B inverted CAT box in two families. *Hum. Immunol.* **1994**, *41*, 69–73. [CrossRef]
54. Sievers, F.; Higgins, D.G. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* **2018**, *27*, 135–145. [CrossRef]
55. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [CrossRef]
56. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **1992**, *8*, 275–282. [CrossRef] [PubMed]