

MOLECULAR AND STRUCTURAL CHARACTERIZATION OF MHC CLASS II β GENES REVEALS HIGH DIVERSITY IN THE COLD-ADAPTED ICEFISH *Chionodraco hamatus*

Marco Gerdol^{1°}, Daniela Lucente^{2,5°}, Francesco Buonocore^{3*}, Elia Poerio³, Giuseppe Scapigliati³, Simonetta Mattiucci⁴, Alberto Pallavicini^{1,6}, Roberta Cimmaruta⁵

¹Department of Life Sciences, University of Trieste, Trieste, Italy

² University of Tuscia, Viterbo, Italy

³Department for Innovation in Biological, Agrofood and Forest Systems, University of Tuscia, Viterbo, Italy

⁴Department of Public Health and Infectious Diseases, Section of Parasitology, Sapienza University of Rome, Rome, Italy

⁵Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy

⁶National Institute of Oceanography and Applied Geophysics

[°]These two Authors contributed equally to the paper

*Corresponding author at: Department for Innovation in Biological, Agro-food and Forest systems, University of Tuscia, Largo dell'Università snc, 05100 Viterbo (VT), Italy, E-mail address: fbuono@unitus.it

atg gtt ctgt ctc agc gtct gc ctg ctgc tc a cgt ctg ctgc a gat ggat at
 M A S V L S V C L L V L S V C S A D G Y
 ctgtttacacagtagccgctgtgagttcaactcctctgagctgaaggacatccagtac
 L F Y T S S R C E F N S S E L K D I Q Y
 atcaagtcatattacaacaagttggagttcatcaggttgcacagcagcgtggggag
 I K S F Y Y N K L E F I R F D S S V G E
 tttgttgatacactgagctgggggtgaggaacgc aaaggcgttcaacagggttctcca
 F V G Y T E L G V R N A K R F N R G S P
 gaactggccgcgatgagagcagagaaggagacgttctgc caaaacaacgttaagctgaa
 E L A A M R A E K E T F C Q N N V K L E
 taccagtacgc cctgc cttctcagcgaaggcctacgtccggcttactccacggtctcc
 Y Q Y A L P F S A K P Y V R L H S T V S
 cccagcggatcacaccggccatgttgtctgcagcgttacgagtctacccaaagtc
 P S G S H P A M L V C S V Y E F Y P K V
 atcaaagtgagctggatcaggaacggacaggaagtcacctctgatgtcacttcctctgaa
 I K V S W I R N G Q E V T S D V T S S E
 gagctggctgacagtgactggta caccagg tccactcccac tggagtacacgcccagg
 E L A D S D W Y Y Q V H S H L E Y T P R
 tctggagacaagatctcctgc atggggagc acgtc agccaggagagc ctgg tta c
 S G D K I S C M V E H V S Q G E P L V T
 gactgggacccctccatgccagagtctgagagagaacaaggtagccatcgagctgcagga
 D W D P S M P E S E R N K V A I G A A G
 ctgatcctgggtctgac ttatctctggccgggttcatctactacaaggagaagtcccga
 L I L G L T L S L A G F I Y Y K R K S R
 ggacggatcctgggtccagccactaa
 G R I L V P S H -

Figure S1. The nucleotide and amino acid sequence of icefish MHC class II β Chha DAB. Start and stop codons are highlighted in red and signal peptide is in green.

atg ggtatgaagttctcgccccactgctgttctgatccttttttcaagagctgat
 M G M K F S F S L L F L I L F F S R A D
 gctcttttggcatgcttgcactgcccactgtttactccctgatgacttttat
 A L F G H A L F H C Q F T S P D D F V Y
 ttggacaactttcttcaataaaagtgtacaactccaataacaacagcacttaggaaag
 L G Q L F F N K V L Q L Q Y N S T L G K
 tataccggctacacagagaaaacgaaagatattgcagaaggcctcaacaaaaatccaaaa
 Y T G Y T E K T K D I A E G L N K N P K
 tttataaaagaagaaaaaaaaatgaaactgaaatgcaagaaccacatcgcaatgttttt
 F I K E E K K N E L K C K N H I A M F F
 gatgtcttttaaaaccagacacctccatggagccctgtcagggtaggtcagtaaa
 D V F L K P D L H L E P S V R V R S V Q
 gcagcgagcagtcgacacccaggcatgctcgactgcacttcttcccaa
 A A S S R H P G M L V C S V H Y F F P K
 ccaatccgagtgacttggtgaggaacggaaaggaggtgacatctgatgtgacgtccact
 P I R V T W L R N G K E V T S D V T S T
 gagaaactgtccaatgggattggcattatcagatccactcctacggagttcacac
 E K L S N G D W H Y Q I H S Y L E F T P
 gtacctggagagaaaatcacctgcattggagcagccccacctcatgaagccaaagctt
 V P G E K I T C M V E H A H L M K P K L
 tgccgtggatccgagaactgatcgagagtcaagagaacaacaagattgtcgccaca
 C E W D P R T D R E S E N N K I A V G T
 gcggggctgctgggtctgggtttttgtctggctgatattacttcaagaagaaa
 A G L L G L V F F V A G L I Y F K K K
 acttatggacgagagttggtgccaaacttctatt**taa**
 T Y G R E L V P T N S I -

Figure S2.The nucleotide and amino acid sequence of icefish MHC class II β Chha DBB. Start and stop codons are highlighted in red, signal peptide is in green, and the potential N-glycosylation site is underlined.

ATGGGTATGAAGTTCTGTTTCACTGCTGTTCTGATCCTTTTTCAAGAGCTG	Exon 1 58 bp (signal peptide)
G TGAGTTACATAAAATAGCTGACATTTACTTTCTAGTGCTATAACATACTACATTACATCTG TGGC ATTTATTGGAGTGTGCAATGCTTATTCA	Intron 1 134 bp
CATTTAGTAATACTGGTGGCTTGTTACAG	
ATGCTCTTTGGTCATGCTTGTCCACTGCCAGTTACTTCCCTGATGACTTGT TATTGGACA ACTTTCCTCAATAAAGTGC TACA ACTCCA	
ATACAA CAGCACTTAGGGAA GTATA CCGG CTACACAGAGAAA C GAA AGATATTGCAGAAGGCCTCAACAAA ATCCAAA ATTATA AAAAGAAGAAAAA	
AAAATGA ACTGAA ATGCA AGAACCA CACAT CGCA ATGTTTGATGTCTTTAAAACAG	Exon 2 261 bp (β -1 domain)
G TGATTGTGTATTGAAACTCTGTTAATAATTAAATTACATATTGCATAATTATTCGTAATTATGGGTTAAGTAACAGACTTTAATCATATTTC	
TGTGAAAAGTGTGCATAGAAAATACAAATACCCACTGTAATATTGTTAACCTAAATATCCAGAGAATTACTTGTGAAAACATTGAAGAGT	
TGTTGTTCATGCCAATTGTTCGTCTATAG	Intron 2 232 bp
ACCTCCATTGGAGCCCTGTCAAGGTGAGGT CAGTACAAGCAGCGAGCAGTCGACACCCAGGCATGCTCGTGCAGTGTGCACTACTTCTTCCAA	
ACCAATCCGAGTGACTTGGCTGAGGAACGGAAAGGAGGTGACATCTGATGTGACGTCCACTGAGAAACTGTC CAATGGGATTGGCATTATCAGATCCAC	
TCCTACCTGGAGTTCACACCTGTACCTGGAGAGAAAATCACCTGCATGGGGAGCACGCCACCTCATGAAGCCCAAGCTTGCAGTGG	
	Exon 3 291 bp (β -2 domain)
G TGGGGTAGGAAAAGTGACCGCTGTTTGAGTTAAATGTATCCAGCCACATTGTATTGATTAGAGTTATAGAAAATGTGTGCTAAATGTGTTTCAG	
	Intron 3 99 bp
ATCCGAGAACTGATCGAGAGTCAGAGAACACAAAGATTGCTGTCGGACAGCGGGCTGCTGGTCTGGTGTGCTGGCTGATTTACTT	
CAAGAAGAAAATTTAG	Exon 4 117 bp (connecting peptide + transmembrane region)
G TGAGAGACACATACACTTACCAAATTCTACCATCTTATCTCAAGACTATCAAGACACAAACATACAGTCTGTCAACAAACTGTCGTTTTGGTC	
ACCCCGTGTAGGCTACGTTTTACTGCTTGCCCTGTGTTCCCTCCTGTTGATTACCTTCTGTTGCTTGCCTCCCTTGTGCTGTC	
TTGGGATTAGTCTGCATGTATTAAAGGTCTGGTTTGCAATTGTTGTCAGATCCTGTTCAATGGTCAAGTTAAAGTTGGTGTCTGTTGTTGATTT	
AGAGATTTCAACTTATTTTGTCTACCGCTGCTCAGCCTGATTACCTGAAATTAAAGTTGGTGTCTGTTCAACACCTGTGCGTCCGTCTGG	
GTTCACCTGCTGCTTCAATCGTGAGAGCTGCTTTGATGAAACTGTTAAAGATTGTATCACTGTTGCTGAAGTCATAAACTCTGTAA	
TTTGTCAAAATTTCTTAGATAACAAAAGAACAGACGACCACTGATTCTTCTAAAGATCAGTTAAAGTAGCTTGTGACTTGTGTTCTATTCTAG	
	Intron 4 688 bp
GACGAGAGTTGGTGCCAACTAATTCTATTAA	Exon 5 32 bp (cytoplasmic tail)

Figure S3. The exon-intron organization of the Chha-DBB sequence. The size of each intron and exon is reported.

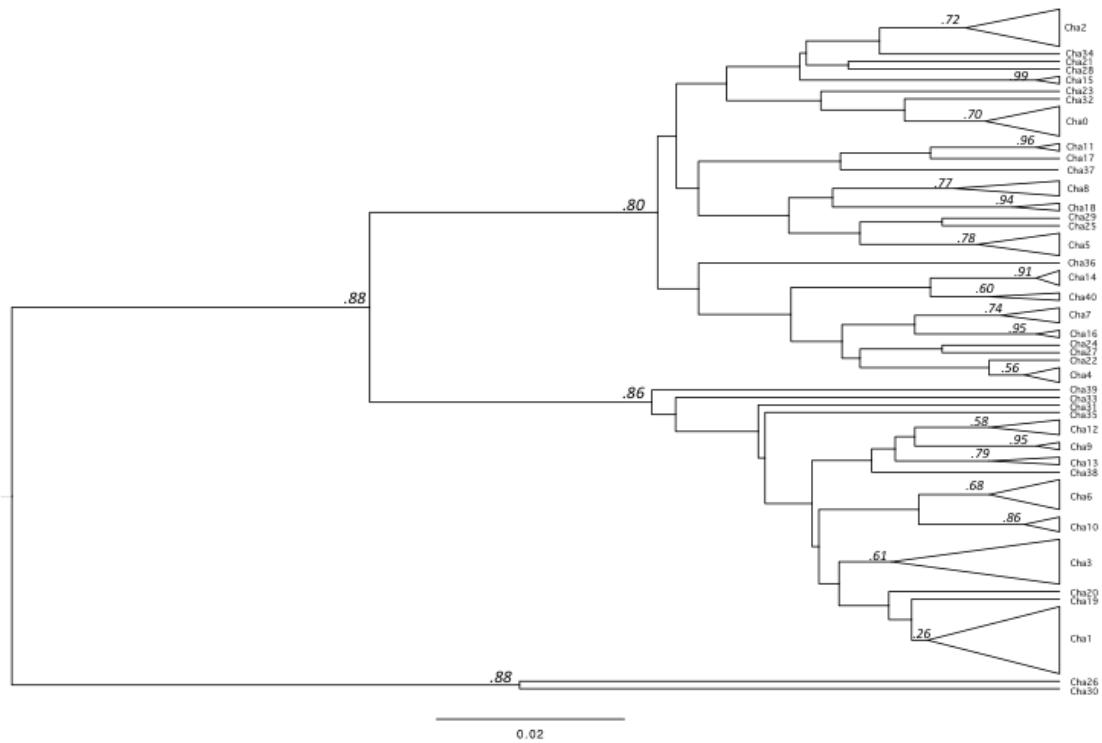


Figure S4. UPGMA dendrogram showing the genetic relationships among the 92 nucleotide sequences of DAB exon 2 recovered. Each triangle represent 2 to 9 genetically close sequences and labels correspond to the 41 peptide clusters obtained by grouping sequences differing by less than 3 aminoacids out of 68 (95%). Bootsrap values over 1000 replicates are reported at each relevant node. 0.02 represents the genetic distance.

