Mar. Sci. Tech. Bull. (2020) 9(1): 15–22
e–ISSN: 2147–9666
info@masteb.com

dergipark.org.tr/en/pub/masteb
www.masteb.com
DOI: 10.33714/masteb.660220

Marine Science and Technology Bulletin

RESEARCH ARTICLE

Characterization, identification and phylogeny of the creatine kinase (*ckma*) gene in medaka (*Oryzias latipes*)

Mehtap Bayır[1*] • Gökhan Arslan[2] • Pınar Oğuzhan Yıldız[2]

[1] *Atatürk University, Agricultural Faculty, Department of Agricultural Biotechnology, Erzurum, Turkey*

[2] *Atatürk University, Faculty of Fisheries, Department of Seafood Processing Technology, Erzurum, Turkey*

ARTICLE INFO

Article History:

Received: 17.12.2019
Received in revised form: 03.02.2020
Accepted: 03.02.2020
Available online: 03.02.2020

Keywords:
Medaka
Genomic organization
Model organism
Bioinformatics

ABSTRACT

Creatine kinase (*ckma*) has been characterized and described in the medaka (*Oryzias latipes*), an aquatic model organism and the gene structure has been designed using the exons, introns, produced amino acids of the gene, TATA box, poly A tail and 5' UTR and 3' UTR regions of the *ckma* gene. In another step, firstly, the chromosome region of the *ckma* gene was determined in medaka and then the other genes which placed in the same region were determined. Then the locations of these genes were determined in zebrafish and human which are the orthologs of medaka. Finally, the conserved gene synteny was designed manually, using these data. However, genetic identity and similarity ratio between medaka and its orthologs were calculated. In this study, characterization and identification, phylogenetic relationship, conserved gene synteny of *ckma* gene in medaka (*O. latipes*) which is an important model organism were analyzed by using bioinformatics tools (NCBI database, Ensembl genomic database, Expasy, Reverse Complementary and some programs such as MEGA6 program, BLOSUM62 matrix program and BioEdit software). All these data will be used in future studies on molecular stress response in fish and they were presented to the scientific world with this study.

Please cite this paper as follows:

Bayır, M., Arslan, G., Oğuzhan Yıldız, P. (2020). Characterization, Identification and Phylogeny of the Creatine Kinase (*ckma*) Gene in Medaka (*Oryzias latipes*). *Marine Science and Technology Bulletin*, 9(1): 15-22.

Introduction

Medaka (*Oryzias latipes*) is a small freshwater fish lives in East Asia. It is an omnivore fish which feeds on vegetable animal foods such as phytoplankton and zooplankton (Hori, 2011). The male medaka can be easily distinguished from the female by its external morphology. Embryos are transparent. Medaka is the first vertebrate in which Mendel inheritance is also exhibited (Ishikawa, 2000; Jacquet et al., 2004; Shima and Mitani, 2004). Although the physiology, embryology and genetics of medaka (*Oryzias latipes*) have been extensively studied for the last 100 years, the studies carried out in this organism have focused on the use of genetic model systems for early development, pigmentation, sex determination and human diseases and the biological history of this fish in the recent years (Naruse et al., 2011). Medeka embryos are used

* Corresponding author
E-mail address: mehtap.bayir@atauni.edu.tr (M. Bayır)

especially in transplantation, microinjection, transgenesis and gene expression studies. Medaka has contributed to important steps in the studies on oncology, ecotoxicology, endocrinology and determination of conserved gene structure (Shima and Shimada, 1991, 2001).

Quantification of fish muscle protein levels indicates that creatine kinase is one of the most highly expressed proteins in fish muscle. This has both cytosolic and mitochondrial forms of regulation of energy production (mitochondria) and use (cytosol) through actions related to adenosine triphosphate (ATP) (McLean et al., 2007).

There is a chemical cycle in the muscle of alive fish. These chemical events provide energy to the muscle during the swimming of the fish and provide the substances necessary for growth and regeneration of dead tissues. Enzymes are substances that create and control chemical reactions in living muscle. Chemical energy is converted to mechanical energy for ATP production which provides the necessary energy. While ATP consumption regeneration and contraction-relaxation events are continuous in living tissue, the amount of ATP decreases rapidly after blood circulation and oxygen supply is cut off in post mortem tissue and contraction and relaxation events continue to be limited during this decrease. The energy required for muscle contraction in live fish is provided by ATP formed during glycolysis. ATP breaks down into adenosine diphosphate (ADP) and inorganic phosphate (P) by the ATPase enzyme, and the energy is used for muscle contraction. ADP and creatine are catalyzed by the creatine kinase enzyme to regenerate ATP from phosphate (Stryer, 1995).

Genetic similarities among species present in all organisms mean that studies on one organism can be used as a data source for other species (Collins et al., 1998). Therefore, in this study, the bioinformatics of *ckma* gene in aquatic model organism, medaka (*O. latipes*) will be completed and the leading data will be provided for molecular studies in other fish.

Material and Methods

Bioinformatics of ckma gene in medaka (O. latipes)

In this study, firstly The National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) was used to investigate whether the creatine kinase (*ckma*) gene functional in medaka (*O. latipes*) and then its cDNA sequence was obtained from ENSEMBL. However, ensembl database was used to characterize the *ckma* gene in medaka (*O. latipes*).

We determined that this gene encode a 381 amino acid protein and has a single isoform (https://www.ensembl.org/Oryzias_latipes/Info/Index) and its ENSEMBL ID and UNIPROT ID have been found as ENSORLT00000033423.1 and A0A3B3I369, respectively.

In the next step, location and chromosome of these genes in zebrafish (*Danio rerio*) and human (*Homo sapiens*) were determined (Table 1) and manually conserved gene synteny was designed (Figure 1) in order to prove the conservation of these genes in these two orthologs of medaka.

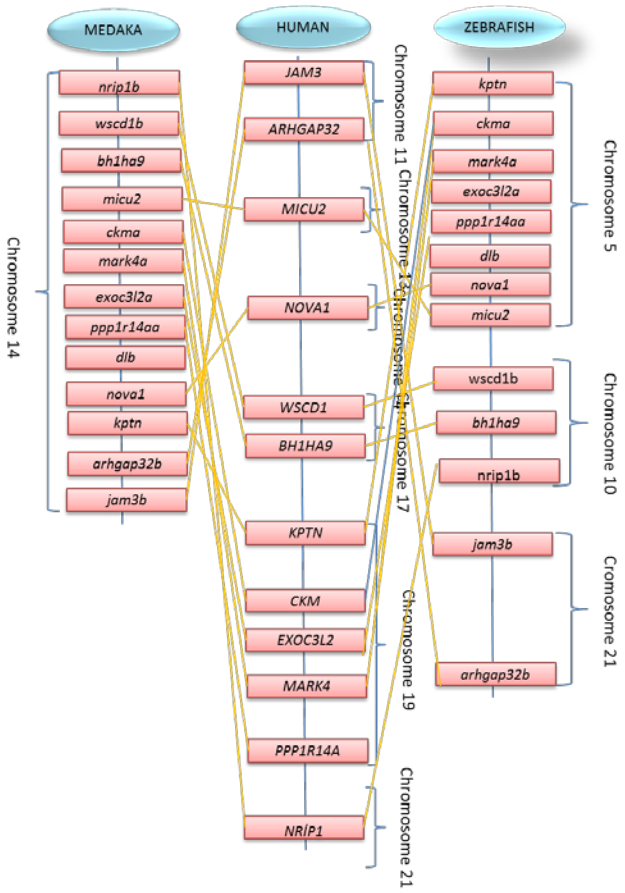

Figure 1. Conserved gene synteny of *ckma* in medaka

For the designing of phylogenetic tree among medaka (*Oryzias latipes*), Monterrey platyfish *(Xiphophorus couchianus),* platyfish *(Xiphophorus maculatus),* Amazon molly *(Poecilia formosa),* stickleback *(Gasterosteus aculeatus),* Midas cichlid *(Amphilophus citrinellus),* tilapia (*Oreochromis niloticus),* lyretail cichlid (*Neolamprologus brichardi*), Makobe island cichlid *(Pundamilia nyererei),* fugu *(Takifugu rubripes),* zebrafish *(Danio rerio), human (Homo sapiens),* mouse *(Mus musculus) ckma/CKM* gene sequences aligned by BioEdit (http://www.mbio.ncsu.edu/bioedit/page2.html) using CLUSTALW (Thompson et al., 1994) and then MEGA6 (Tamura et. al., 2013) program was used according to the maximum likelihood method (Kell et al., 2018) (Figure 2). Medaka (*Oryzias latipes*) glutathione reductase (*gsr)* (A0A3P9I169) was chosen as an external group.

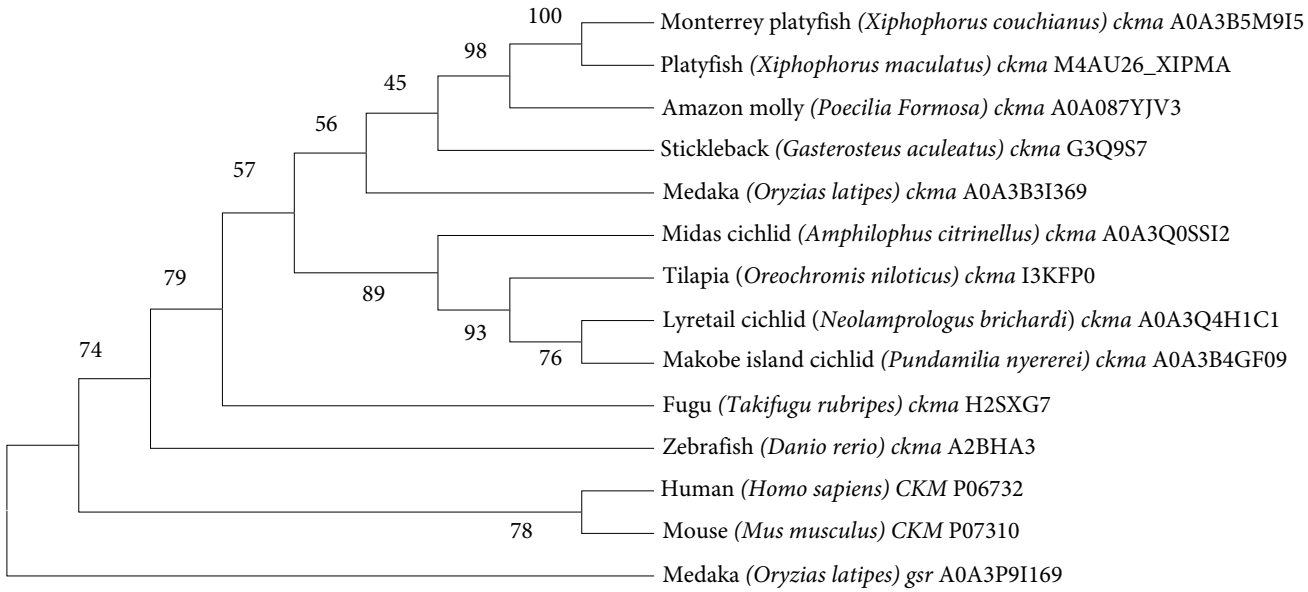

Figure 2. Phylogenetic tree of *ckma* in medaka (*O. latipes*). Phylogenetic relationships between *ckma* sequence from medaka and the other vertebrates. Tree was produced using Maximum Likelihood method (Felsenstein, 1989). Accession numbers (UNIPROT) of the sequences used for phylogenetic tree are shown in phylogenetic tree.

Table 1. The genes which are used in conserved gene synteny and their location in medaka, zebrafish, and human

Gene	Gene symbol	Medaka		Zebrafish		Human	
		Chromosome	*Location*	*Chromosome*	*Location*	*Chromosome*	*Location*
Creatine kinase, muscle a	*ckma*	14	2.16	5	36.83	19	45.30
Junctional adhesion molecule 3b	*jam3b*	14	1.79	21	24.98	11	134.06
Rho GTPase activating protein 32b	*arhgap32b*	14	1.88	21	24.53	11	128.96
Neuro-oncological ventral antigen 1	*nova1*	14	1.98	5	36.61	14	26.44
Kaptin, actin binding protein	*kptn*	14	1.96	5	36.91	19	47.47
DeltaB	*dlb*	14					
Exocyst complex component 3-like 2a	*exoc3l2a*	14	2.41	5	3.67	19	45.21
Protein phosphatase 1 regulatory inhibitor subunit 14A	*ppp1r14aa*	14	2.10	5	36.73	19	38.51
Microtubule affinity regulating kinase 4a	*mark4a*	14	2.14	5	36.76	19	45.07
Mitochondrial calcium uptake 2		14	2.26	5	36.59	13	21.49
Basic helix-loop-helix family member a9		14	2.35	10	37.92	17	1.27
WSC domain containing 1b	*wscd1b*	14	2.40	10	37.98	17	6.05
Nuclear receptor interacting protein 1	*nrip1b*	14	2.65	10	8.25	21	14.96

For the design of gene structure, ENSORLT00000033423.1 cDNA transcript of medaka (*O. latipes*) *ckma* gene was used. exon-intron organization of the medaka *(O. latipes*) *ckma* gene and the amino acids produced by the exons, the 5' UTR and 3' UTR regions of the *ckma* gene, the TATA box, the poly A tail, and the starting point of transcription (+1) were showed in the

gene structure (Table 2). Zebrafish (*Danio rerio*), Nile tilapia (*Oreochromis niloticus*), fugu (*Fugu rupripes*), human (*Homo sapiens*) and mouse (*Mus musculus*) ckma/CKM proteins were used in Bioedit program, CLUSTALW (Thompson et al., 1994) for analyzing the similarity-identity ratios (Table 3).

Results and Discussion

Bioinformatics of ckma gene in medaka (O. latipes)

Oxygen deficiency is a major factor in creatine increasing in fish, besides the impact of industrial enterprises' waste (Arslan, 2015). Stress responses of vertebrates include different interactions between physiological pathways that can be characterized in both acute and chronic conditions. Creatine kinase (CK) is an important enzyme used in the detection of damage to tissues and organs such as glutamic-pyruvic acid transaminase (GPT), glutamic-oxaloacetic acid transaminase (GOT), alkaline phosphatase (ALP) and lactic dehydrogenase (LDH) enzymes. These enzymes, except from CK, are liver enzymes and those are also used to understand liver problems.

CK and GOT enzymes tend to increase in wounds on fish skin and in case of damage to muscle tissue and brain. In addition, the CK enzyme allows the regeneration of ATP in contraction or delivery systems. Therefore, the completion of the detailed bioinformatics study of the creatine kinase (*ckma*) gene, which is one of the stress markers, in the medaka (*O. latipes*) (Iwama et al., 1999) is important. Therefore, it is of great importance to complete detailed bioinformatics study of the creatine kinase (*ckma*) gene which is one of the stress markers in medaka (aquatic model organism) has great importance, because acute or chronic stress responses of fish change with environmental differences.

Because fish are aquatic organisms, changes in both qualitative and quantitative properties of water can lead to changes in the functional structures of these organisms, resulting in unfolding of protein folds from time to time, and these proteins can combine with other proteins in the cell to form clusters. Consequently, proteins may lose their functions due to conformation deformation (Basu et al., 2000). However, in this research, firstly, *ckma* gene was determined to be a functional gene in medaka (*O. latipes*) by using of bioinformatics tools, and then the other bioinformatics studies were carried out such as gene structure determination, phylogenetic tree design, conserved gene synteny and calculation of the identity-similarity rates between medaka (*O. latipes*) and its orthologs. When a molecular study is planned, firstly bioinformatics studies should be completed before experimental studies to understand how the expression of genes changes with various stress factors. Therefore, this study will provide important bioinformatics data both for fish physiology studies and for the other studies on vertebrates because medaka (*O. latipes*) is an aquatic model organism.

In this study, ENSEMBL, UNIPROT, NCBI databases and BioEdit software, BLOSUM62 matrix program and MEGA6 program were used to reach some knowledge such as the cDNA, exons and introns of the *ckma* gene, the amino acids produced by this gene, the 5' UTR and 3' UTR regions, the chromosome and location where the gene is positioned, and the protein sequences necessary to determine the phylogenetic relationship to other vertebrates. The cDNA sequence of the medaka (*O. latipes*) *ckma* gene was obtained from the Ensembl database (Ensemble number ENSORLT00000033423.1) and it was found that this gene has a single isoform, which encoded a protein of 381 amino acids. Medaka *ckma* gene has 7 exons and 6 introns located between these exons. The amino acids produced by the exons and the 5 'and 3' ends of the gene, TATA box and Poly A tail are given in detail in Table 2.

The sequence identity-similarity ratio was calculated to investigate the orthology between the medaka (*O. latipes*) and zebrafish *ckma* gene. For this purpose, medaka (*O. latipes*), zebrafish (*Danio rerio*), fugu (*Fugu rupripes*), Nile tilapia (*Oreochromis niloticus*) protein sequence produced by *ckma* gene and mouse (*Mus musculus*) and human (*Homo sapiens*) protein sequences produced by CKM gene were aligned using the BioEdit program in the BLOSUM62 matrix algorithm, and the similarity-identity ratios of these organisms were calculated (Gromiha, 2010) and the results were given in Table 3. According to the table, the identity and similarity percentage of medaka (*O. latipes*) *ckma* gene was 98-94% with Nile tilapia, 97-93% with zebrafish, 96-91% with fugu, 93-87% with human, and 92-87% with mouse (Table 3).

In order to define the conserved genes in both medaka and zebrafish and human, the location of *ckma* gene was determined on the 14th chromosome in medaka. Then the other genes and their locations were determined in this chromosome using the Ensemble genome database (Table 1). Conserved gene synteny was determined by detecting the chromosomes and regions of these detected genes (*ckma, jam3b, arhgap32b, nova1, kptn, dlb, exoc3l2a, ppp1r14aa, mark4a, wscd1b, nrip1b*) found in human and zebrafish (Figure 1). These genes on chromosome 14 in medaka (*O. latipes*) are also conserved in humans (chromosomes 11, 13, 14, 19 and 20) and zebrafish (chromosomes 5, 10 and 21). It is known that teleost fish have evolutionary conserved regions in the same gene family, and the designed conserved gene synteny clearly demonstrates it. In addition, it is thought that the *ckma* gene of

Table 2. Gene structure of *ckma* in medaka (*Oryzias latipes*)

```
5'taaactgcaaggacttgaagggtaaaaggccagatattctggggctaaaaatacccgg          -299
agagcaggctctccacccctgctcaatttcaactggacatctgagccactggaaactgag          -239
cgacacttgttaccaagaatctgcggacagcaccgtttgaaatttgcagctgcccaaaat          -179
gtcatatgctcaaagaaggaaaaagcatcatttgcagcgtccttgctcctcctttatgaa          -119
tgaggctgcaatgacctgtcttcattgtattATATAgcctaagcttgttgtgtttttcag           -59
+1
TGTTAGAAAGCAATCATGCCTTTCGGAAACACCCACAACAACTTCAAGCTCAACTACTCA            60
                -M--P--F--G--N--T--H--N--N--F--K--L--N--Y--S-
GTTGACGATGAGTTCCCAGACCTGTCCAAGCACAACAACCACATGGCCAAAGTCCTGACT           120
-V--D--D--E--F--P--D--L--S--K--H--N--N--H--M--A--K--V--L--T-
AAAGAGCTGTATGGTAAGATGAGGGACAAGCAGACGCCCACTGGATTCACTCTGGATGAC           180
-K--E--L--Y--G--K--M--R--D--K--Q--T--P--T--G--F--T--L--D--D-
GTGATCCAGACCGGCATCGACAACCCTGgtgagacttcaagcaacatttcttctttttc           240
-V--I--Q--T--G--I--D--N--P--
caacagaatccaagatagtaaaagacaagaaacaagtgttagggtcaattcataaccccc           300
acctttgttatcagGTCACCCCTTCATCATGACTGTTGGCTGTGTCGCTGGTGACGAGGA           360
              G--H--P--F--I--M--T--V--G--C--V--A--G--D--E--E
GTCTTATGAGGTCTTCAAAGACCTGCTTGACCCCGTCATCTCTGACCGTCATGGTGGATA           420
--S--Y--E--V--F--K--D--L--L--D--P--V--I--S--D--R--H--G--G--Y
TAAGCCCACTGACAAGCACAAGACTGACCTCAACTTCGAGAACTTGAAGgtgcaatacag           480
--K--P--T--D--K--H--K--T--D--L--N--F--E--N--L--K-
cttctttagagagcagagttacacctagccctttctaatgttcctcacggcccaatctaa           540
ctgtgtctgtagGGAGGTGATGACCTGGACCCCAACTACGTTTTGTCCAGCCGTGTTCGT           600
             -G--G--D--D--L--D--P--N--Y--V--L--S--S--R--V--R-
ACCGGTCGCAGCATCAAGGGATACGCCCTGCCCCCCCACAACAGCCGTGGCGAGCGCAGA           660
-T--G--R--S--I--K--G--Y--A--L--P--P--H--N--S--R--G--E--R--R-
GCTATTGAGAAGCTGTCCATTGAGGgtaagttttcttgattttggggatttccacaggtc           720
-A--I--E--K--L--S--I--E--
aagagtatctgatacccaggtttctgtggtcagtcataaaccagactgaatccaggcttt           780
ctgctctagcaggtcttctaaatcatcatgcaatgcctaatgcatcgatgtatgaaataa           840
agaagtgttctgtttttggtggatgctgacctaacagtgagcctcttcctgcagCTCTG           900
                                                        A--L-
TCCAGCCTTGATGGTGAGTTCAAAGGAAAGTACTATCCCCTGAAGTCAATGACTGATGCT           960
-S--S--L--D--G--E--F--K--G--K--Y--Y--P--L--K--S--M--T--D--A-
GAGCAGGAGCAGCTGATCAGTGATCACTTCCTGTTTGACAAACCTGTGTCCCCCCTGTTG          1020
-E--Q--E--Q--L--I--S--D--H--F--L--F--D--K--P--V--S--P--L--L-
ACCTGCGCCGGTATGGCCCGTGACTGGCCTGACGGCAGAGGCATTTGgtaagtgcagtta          1080
-T--C--A--G--M--A--R--D--W--P--D--G--R--G--I--W
ggaatggtcatcctctgtaaatacaccaaacactcagcttgtatagattcatcaggatta          1140
atcactgacctgcgtagtgctgtccatggtcagtgtccataaatcaagcaagtctcatct          1200
tgtctgagcagtcagagtacaactggaaaacatccacaaatgagtcctcaaggatttcct          1260
ggcagggaaatcatgatggcagtagatacattgggctctgagcttaaattctcattggtc          1320
tgcaagatattgcacattgtccaaatctgtgcccgttggcatctctacatccagGCACAA          1380
                                                        --H--N
CGACAACAAGACCTTCCTGGTGTGGGTGAATGAGGAGGATCACCTGCGTGTCATCTCCAT          1440
--D--N--K--T--F--L--V--W--V--N--E--E--D--H--L--R--V--I--S--M
GCAGAAGGGTGGCAACATGAGGGAGGTCTTCAGGCGTTTCTGCGTGGGCTTGCAGAAGgt          1500
--Q--K--G--G--N--M--R--E--V--F--R--R--F--C--V--G--L--Q--K-
gcatgaagaccgcagatcaaatctgctcagcctgtttaaccaagtcaaacctaaagcagc          1560
tgtgatcctgacccttcttttatgactctcagATTGAGGAGATCTTCAAGAAGCACAACC          1620
                                -I--E--E--I--F--K--K--H--N--
ACGGCTTCATGTGGAATGAGCATCTCGGCTACATTCTGACCTGCCCCTCCAACCTGGGAA          1680
H--G--F--M--W--N--E--H--L--G--Y--I--L--T--C--P--S--N--L--G--
CTGGTCTGCGTGGGGGTGTCCACGTCAAGCTGCCCAAGCTGAGCACACACCCCAAGTTTG          1740
T--G--L--R--G--G--V--H--V--K--L--P--K--L--S--T--H--P--K--F--
AGGAGATCCTCACCAGGTTGCGCCTGCAGAAGCGTGGCACAGgtatggatgtgctccatc          1800
E--E--I--L--T--R--L--R--L--Q--K--R--G--T--
tgtgggacctctacagaggctctgtggacgctcgtatgaggtgttatgtcatgccacatc          1860
ctttctctccagGTGGTGTGGACACTGCATCTGTGGGTGGTGTGTTTGACATCTCCAATG          1920
            G--G--V--D--T--A--S--V--G--G--V--F--D--I--S--N--
CCGACCGTCTTGGATCCTCCGAGGTGGCGCAGGTCCAGTTGGTGGTTGATGGCGTCAAGC          1980
A--D--R--L--G--S--S--E--V--A--Q--V--Q--L--V--V--D--G--V--K--
TGATGGTTGAGATGGAGAAGAAGCTCGAGAAGGGAGAAGCCATCGACAGCATGATCCCCG          2040
L--M--V--E--M--E--K--K--L--E--K--G--E--A--I--D--S--M--I--P--
CCCAGAAGTGAggagggacaatctggcattttccttgtgacctttatgtgcagtcgagc          2100
A--Q--K--*-
cagctgacagcgtgcctgcagagaaaacagccgctcacctagagactcttgactctgcta          2160
actcctttccttccttccagctttgttttttctttttctccttccttgtcgtttttcacg          2220
ttcccctgcgttggtcagtaacatccaggggcagcctcactgagcggggcttgcctagc          2280
ggacatggcatcaccaacttttttgttataagaagtaacaactgttgaataggttcatact          2340
gttcAATAAAAcagcgtcccctgaacacgtctgggtcatccctgtctttcttgttttg 3'          2400
```

Note: The exons of the *ckma* are shown in capital letters and the nucleotide positions are numbered at the end of the each line. The starting site of transcription is +1,5' upstream sequence, 3' downstream sequence and introns are shown in lower case. The TATA box and the poly adenylation signal (AATAAAA) are shown in capital letters and painted in yellow. Amino acids are shown in capital letters which are placed under exons. Stop codon (TGA) is specified asterisk.

Table 3. Identity and similarity rate between medaka (Me) and Nile tilapia (Nt), zebrafish (Zf), fugu (Fu), human (Hu) and mouse (Mo)

```
Me ckma   1   MPFGNTHNNFKLNYSVDDEFPDLSKHNNHMAKVLTKELYGKMRDKQTPTGFTLDDVIQTG
Nt ckma   1   ............K.EE.............S..V...L......S.Y.........
Zf ckma   1   ...............E.Y.............M...L.....S....V.......
Fu ckma   1   .AK-.C..DY.MKMO..E......Q...........I...L.G.S..S...V.......
Hu CKM    1   ........K.....KPEE.Y............L...K.L...E..S...V.......
Mo CKM    1   ........K.....KPQE.Y............PD..N.L...E..S...........

Me ckma   61  IDNPGHPFIMTVGCVAGDEESYEVFKDLLDPVISDRHGGYKPTDKHKTDLNFENLKGGDD
Nt ckma   61  V.................................................H........
Zf ckma   61  V........................F...........A..................
Fu ckma   60  V....................A.....................................
Hu CKM    61  V......................E.F..I...................H........
Mo CKM    61  V..................T.....F..I.Q................H........

Me ckma   121 LDPNYVLSSRVRTGRSIKGYALPPHNSRGERRAIEKLSIEALSSLDGEFKGKYYPLKSMT
Nt ckma   121 ................FT..........I..R......N............T..
Zf ckma   121 ..............................V....V...................
Fu ckma   120 ................FT....................A............TG..
Hu CKM    121 ..................T....C.......V....V...N..T............
Mo CKM    121 ..................T....C.......V....V...N..T............

Me ckma   181 DAEQEQLISDHFLFDKPVSPLLTCAGMARDWPDGRGIWHNDNKTFLVWVNEEDHLRVISM
Nt ckma   181 ........A.............................E...................
Zf ckma   181 ........A............LA.........A......E...................
Fu ckma   180 ........A.....................................S.............
Hu CKM    181 EK..Q...D............LAS........A.........S.............
Mo CKM    181 EQ..Q...D............LAS........A.........S.............

Me ckma   241 QKGGNMREVFRRFCVGLQKIEEIFKKHNHGFMWNEHLGYILTCPSNLGTGLRGGVHVKLP
Nt ckma   241 ..........................D.................................
Zf ckma   241 ......K...K.......R..................FV...................
Fu ckma   240 ......K.............A.....................................
Hu CKM    241 E.....K..................AG.P....Q....V...................A
Mo CKM    241 E.....K..................AG.P.........V...................A

Me ckma   301 KLSTHPKFEEILTRLRLQKRGTGGVDTASVGGVFDISNADRLGSSEVAQVQLVVDGVKLM
Nt ckma   301 ...................................................E.........
Zf ckma   301 .....A.................................I.....E...C........
Fu ckma   300 ....Q.....................................................E.........
Hu CKM    301 H..K.....................A..S...V..........E.............
Mo CKM    301 N..K.....................A..A.............E.............

                                        Identity (%)   Similarty (%)
Me ckma   361 VEMEKKLEKGEAIDSMIPAQK          100            100
Nt ckma   361 ..........S.........            98             94
Zf ckma   361 ..........S.........            97             93
Fu ckma   360 ..........S..G......            96             91
Hu CKM    361 ..........QS..D......            93             87
Mo CKM    361 ..........QS..D......            92             87
```

Note: The dots and lines refer to repeating amino acids and undetectable amino acids, respectively.

medaka emerged as a result of teleost genome duplication seen in bony fish. As known, teleost fish may have two copies of genes found as a single copy in other vertebrates as a result of whole genome duplication (Amores et al., 1998; Meyer and Schartl, 1999; Postlethwait et al., 2000; Braasch and Postlethwait, 2012; Çapan, 2019). It was observed that tilapia, puffer fish, stickleback, platyfish, Midas cichlid, Makobe island cichlid, fugu, Amazon molly and medaka have just one copy of the creatine kinase gene (*ckma*), while zebrafish has two copies of this gene, *ckma* and *ckmb*, when explored Ensembl database. In this case, it is thought that one copy is lost following teleost whole genome duplication in these species except from zebrafish. Yamamoto (1953), firstly created a gender linkage map for medaka and described differences in the frequency of recombination between genders. It was also reported for the first time that there was an autosomal connection between *i* and

ci loci in fish. Following the development of polymerase chain reaction (PCR) technology, several attempts have been made to create a genetic linkage map in medaka, zebrafish, puffer and other fish species, and finger-print markers were used in the early stages of these experiments, as they did not require prior genome information. In subsequent steps, single locus markers were used to amplify specific regions of the genome in the presence of sequence information, and the map generated using activated single locus markers was used to compare linkage relationships between orthologous genes. All genome amplification specific to the teleosts were then applied (third WGD). Finally, in addition to the tetraodon genome project, the medaka genome sequencing project provided a high quality outline genome sequence for both medaka and tetraodon. All these data confirmed the third WGD, which revealed a potential scenario in which reconstruction of proto-chromosomes prior to duplication and the formation of existing medaka, tetraodon and zebrafish genomes.

Phylogenetic relationship can be seen in the tree (Figure 2) which created using protein sequences of medaka (*O. latipes*), Monterrey platyfish *(X. couchianus),* platyfish *(X. maculatus),* Amazon molly *(P. formosa),* stickleback *(G. aculeatus),* Midas cichlid *(A. citrinellus),* tilapia (*O. niloticus),* lyretail cichlid (*N. brichardi*), Makobe island cichlid *(P. nyererei),* fugu *(T. rubripes),* zebrafish *(D. rerio),* human *(H. sapiens)* and mouse (*M. musculus*) according to maximum likelihood method using MEGA6 (Tamura et. al., 2013) program. It was observed that the medaka showed clustering with other teleost fishes, and that living organisms such as humans, chickens and mice were clustered in a different region (Figure 2).

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V. & Wang, Y. L. (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science,* ***282***(5394): 1711-1714. https://doi.org/10.1126/science.282.5394.1711

Arslan, H. (2015). Pestisit sinerjisinin; gökkuşağı alabalıklarında (*Oncorhynchus mykiss*) yüzme performansı, biyokimyasal hematolojik, histopatolojik ve genotoksik etkilerinin araştırılması. Ph.D. Thesis. Atatürk University, Erzurum, Turkey.

Basu, S., Binder, R. J., Suto, R., Anderson, K. M. & Srivastava, P. K. (2000). Necrotic but not apoptotic cell death releases heat shock proteins, which deliver a partial maturation signal to dendritic cells and activate the NF-kappa B pathway. *International Immunology*, ***12***(11): 1539-1546. https://doi.org/10.1093/intimm/12.11.1539

Braasch, I. & Postlethwait, J. H. (2012). Polyploidy in fish and the teleost genome duplication (pp. 341-383). In: Soltis, P. S., Soltis, D. E. (eds.), Polyploidy and Genome Evolution. Springer. 420p. https://doi.org/10.1007/978-3-642-31442-1_17

Collins, F. S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R. & Walters, L. (1998). New goals for the U.S. human genome project: 1998-2003. *Science,* ***282***(5389): 682-689. https://doi.org/10.1126/science.282.5389.682

Çapan, E. C. (2019). Plati balığı (*Xiphophorus maculatus*)'nda katalaz enzim geninin biyoenformatiği ve doku spesifik dağılımı. Master Thesis. Atatürk University, Erzurum, Turkey.

Hori, H. (2001). A glance at the past of medaka fish biology (pp. 1-16). In: Naruse, K., Tanaka, M., Takeda, H. (eds.), Medaka: A model for organogenesis, *human disease, and evolution.* Tokyo: Springer. 387p. https://doi.org/10.1007/978-4-431-92691-7_1

Ishikawa, Y. (2000). Medakafish as a model system for vertebrate developmental genetics. *BioEssays*, ***22***(5): 487-495. https://doi.org/10.1002/(SICI)1521-1878(200005)22:5%3C487::AID-BIES11%3E3.0.CO;2-8

Iwama, G. K., Vijayan, M. M., Forsyth, R. B. & Ackerman, P .A. (1999). Heat shock proteins and physiological stress in fish. *American Zoologist,* ***39***(6): 901-909.

Jacquet, C., Thermes, V., de Luze, A., Puiseux-Dao, S., Bernard, C., Joly, J. S., Bourrat, F. & Edery, M. (2004). Effects of microcystin-LR on development of medaka fish embryos (*Oryzias latipes*). *Toxicon*, ***43***(2): 141-147. https://doi.org/10.1016/j.toxicon.2003.11.010

Kan, B., London, I. M. & Levin, D. H. (1988). Role of reversing factor in the inhibition of protein synthesis initiation by oxidized glutathione. *Journal of Biological Chemistry,* ***263***(30): 15652- 15656.

Kell, A. J. E., Yamins, D. .L. K., Shook, E. N., Norman-Haignere, S. V. & McDermott, J. H. (2018). A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy. *Neuron*, ***98***(3): 630-644.e16. https://doi.org/10.1016/j.neuron.2018.03.044

McLean, L., Young, I. S., Doherty, M. K., Robertson, D. H. L., Cossins, A. R., Gracey, A. Y., Beynon, R. J. & Whitfield, P. D. (2007). Global cooling: Cold acclimation and the expression of soluble proteins in carp skeletal muscle. *Proteomics*, ***7***(15): 2667-2681. https://doi.org/10.1002/pmic.200601004

Meyer, A. & Schartl, M. (1999). Gene and genome duplications in vertebrates: The one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Current Opinion in Cell Biology,* ***11***(6): 699-704. https://doi.org/10.1016/s0955-0674(99)00039-3

Naruse, K., Fukamachi, S., Mitani, H., Kondo, M., Matsuoka, T., Kondo, S., Hanamura, N., Morita, Y., Hasegawa, K., Nishigaki, R., Shimada, A., Wada, H., Kusakabe, T., Suzuki, N., Kinoshita, M., Kanamori, A., Terado, T., Kimura, H., Nonaka, M. & Shima, A. (2000). A detailed linkage map of medaka, *Oryzias latipes*: Comparative genomics and genome evolution. *Genetics*, ***154***(4): 1773–1784.

Postlethwait, J. H., Woods, I. G., Ngo-Hazelett, P., Yan, Y. L., Kelly, P. D., Chu, F., Huang, H., Hill-Force, A., Talbot, W. S. (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Research,* ***10***(12): 1890-1902. https://doi.org/10.1101/gr.164800

Shima, A. & Mitani, H. (2004). Medaka as a research organism: past, present and future. *Mechanisms of Development*, ***121***(7-8): 599–604. https://doi.org/10.1016/j.mod.2004.03.011

Stryer, L. (1995). Biochemistry (4th Ed.). New York: W.H. Freeman and Company. 1064 p.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research,* ***22***(22): 4673-4680. https://doi.org/10.1093/nar/22.22.4673

Yamamoto, T. (1953). Artificial sex-reversal in the genotypic males of the medaka, *Oryzias latipes*. *Journal of Experimental Zoology,* **123***(3):* 517-594. https://doi.org/10.1002/jez.1401230309