Journal of Agricultural Production (2020) 1(1): 12-21 https://agriprojournal.com/
e-ISSN: 2757-6620 agriprojournal@gmail.com

JOURNAL OF AGRICULTURAL PRODUCTION

RESEARCH ARTICLE

Phylogeny, Characterisation and Identification of Creatine Kinase Genes (*ckma* and *ckmb*) in Zebrafish (*Danio rerio*)

Gökhan Arslan[1] • Erdal Özdemir[2] • Mehtap Bayır[2*]

[1]*Atatürk University, Faculty of Fisheries, Department of Fisheries and Seafood Processing Technology, Erzurum/Turkey*
[2]*Atatürk University, Faculty of Agriculture, Department of Agricultural Biotechnology, Erzurum/Turkey*

ARTICLE INFO

Article History:
Received: 17.12.2020
Accepted: 28.12.2020
Available Online: 31.12.2020

Keywords:
Danio rerio
Genomic organisation
Model organism
Bioinformatics

ABSTRACT

Creatine kinase genes (*ckma* and *ckmb*) in zebrafish (*Danio rerio*), an aquatic model organism, have been characterized and identified. However, the gene structure is designed using exons, introns, amino acids produced by the exons. TATA boxex, poly A tails and 5' UTR and 3' UTR regions of zebrafish *ckma* and *ckma* genes are showed at the gene structure. In addition, chromosomal regions of *ckma* and *ckmb* genes were determined. The other genes which are placed in the same region with *ckma* and *ckmb* genes were found in medaka and human which are the orthologs of zebrafish, and conserved gene syntheny was designed manually according to these regions. In addition, phylogenetic relationship was determined between zebrafish and it's some orthologs using *ckma* and *ckmb* gene sequences. Genetic affinity between zebrafish and its orthologs was calculated as similarity-identity % rate and given as a table. For all these studies, bioinformatics databases (NCBI database, Ensembl genomic database, Expasy, Reverse Complementary) and programs (MEGA6 program, BLOSUM62 matrix program and BioEdit software) were used. In this study, characterization and identification of *ckma* and *ckmb* genes in zebrafish (*D. rerio*) was completed using bioinformatics tools and some data to be used in the future studies on molecular stress response were presented.

Please cite this paper as follows:

Arslan, G., Özdemir, E. and Bayır, M. (2020). Phylogeny, Characterisation and Identification of Creatine Kinase Genes (*ckma* and *ckmb*) in Zebrafish (*Danio rerio*). *Journal of Agricultural Production*, 1(1): 12-21.

Introduction

Zebrafish (*Danio rerio*) is a tropical freshwater fish that has its natural habitat in rivers of Northern Pakistan, Bhutan and Nepal, as well as rivers of South Asia and Northeast India (Carpio and Estrada, 2006). Zebrafish belonging to the Cyprinidae family in the ray-finned fish (Actinopterygii) class is a teleost fish (Carpio and Estrada, 2006). It has many advantages, such as being a rapidly developing creature such as being completed to a large extent afterwards (Gilmour et al., 2002). It has been one of the most researched model organisms due to its existence (Carpio and Estrada, 2006). In addition, zebrafish embryos are an aquatic model organism that is powerful enough for experimental-mental manipulations such as microinjection and cell transplantation experiments, and therefore highly preferred in genetic studies (Gilmour et al., 2002). Due to the transparency of the embryo, it has been possible to directly observe its internal development and the embryos can be genetically and embryologically manipulated through microinjection, making zebrafish an excellent complementary research model for human disease and development (Ma, 2004; Lieschke and Currie, 2007).

Analysis of fish muscle protein levels indicates that creatine kinase is one of the most highly expressed proteins in fish muscle (McLean et al., 2007). It has both cytosolic and mitochondrial forms involved in the regulation of energy production (mitochondria) and utilization (cytosol) through actions related to ATP. There is a chemical cycle in the alive fish muscle. These chemical events provide energy to the muscle while the fish swim, providing the substances necessary for growth and regeneration of dead tissue. The substances that create and control chemical reactions in living muscle are

* Corresponding author
E-mail address: mehtap.bayir@atauni.edu.tr

enzymes, and the energy source required for this application is ATP, which converts chemical energy into mechanical energy. While ATP consumption and re-formation and contraction-relaxation events in living tissue are continuous, the amount of ATP decreases rapidly as a result of the interruption of blood circulation and oxygen supply in the post-mortem tissue, and contraction and relaxation events continue to be limited during this decrease. The energy required for the contraction of the muscle in living fish is provided by the ATP formed during glycolysis. ATP is broken down into adenosine diphosphate (ADP) and inorganic phosphate (P) by the ATPase enzyme, and the energy released at this time is used for the contraction of the muscle. ADP and creatine are catalyzed by the keratin kinase enzyme to regenerate ATP from phosphate (Stryer, 1995).

Although it is known that there is genetic similarity between species in organisms, the thesis that studies on an organism can be used as a data source for other species (Collins et al., 1998) increased the importance of model organisms in scientific studies. Therefore, in this study, bioinformatics analysis of *ckma* and *ckmb* genes in zebrafish will provide pioneering data for molecular studies in other organisms. Creatine kinase one of the enzymes that maintains cellular energy homeostasis and high ATP/ADP and ATP/AMP ratios in vertebrates in order to meet the high energy demand for physiological responses in living organism (Wallimann et al., 1992). In order to provide higher energy, it is necessary to catalyze phosphocreatine and ADP of creatine kinase, creatine and ATP in ADP more effectively (Wu et al., 2011). Also, zebrafish cannot survive in water temperatures below 12 ^{o}C or cannot be fed when the temperature is below 16 ^{o}C (Chou et al., 2008). Therefore, this study is of great importance in providing basic information for studies on both zebrafish and other teleost fish.

Materials and Methods

To investigate whether the creatine kinase genes (*ckma* and *ckmb*) are functional or they are nonfunctional or pseudo-genes in zebrafish, the cDNA sequences of these genes were obtained from the ENSEMBL database and blasted (http://blast.ncbi.nlm.nih.gov) using the NCBI database described in our previous publication (Bayır et al., 2020), and it was confirmed that both *ckma* and *ckmb* genes are functional genes in zebrafish. Then, ENSEMBL data bank was used to characterize creatine kinase genes. The ensembl number of the zebrafish *ckma* gene was ENSDART00000032481.6 and the UNIPROT number was A2BHA3, while the ensembl number of the *ckmb* gene was ENSDART00000059366.7 and the UNIPROT number was Q7T306. It was also determined that the *ckma* gene encodes a protein of 381 amino acids, while the *ckmb* gene encodes a protein of 380 amino acids.

Conserved gene syntheny was designed manually by detecting conserved common genes and determining their locations to detect genes that are preserved in the same way as their orthologists. For this purpose, conserved genes in zebrafish (*Danio rerio*), human (*Homo sapiens*) and medaka (*Oryzias latipes*) used for the *ckma* gene while conserved genes in zebrafish, human (*Homo sapiens*) and spotted gar (*Lepisosteus oculatus*) was using for *ckmb* gene. First of all, it was determined in which chromosomes and in which regions the *ckma* and *ckmb* genes were found in zebrafish (*D. rerio*) and then the other genes on this chromosome were found and the locations of these genes were recorded. Later, a conserved gene syntheny was formed by detecting the chromosomes and locations of these genes, which were detected in medaka (*Oryzias latipes*), spotted gar *(Lepisosteus oculatus*) and human (*Homo sapiens*), which are orthologists of zebrafish (*Danio rerio*) (Figure 1). Using the CLUSTALW (Thompson et al., 1994) BioEdit program the proteins of these organisms were aligned and then using MEGA6 (Using the program Tamura et al., 2013) a phylogenetic tree (Kell et al., 2018) was created according to the maximum likelihood method (Figure 2). Medaka (*O. latipes*) glutathione reductase (*gsr*) gene was used as outgroup (Figure 2).

The gene structures consisting of the starting point (+1) of transcription, exon-intron organization, amino acids produced by the exons, the 5 'UTR regions of these two genes (with the TATA box in this region) and the 3'UTR region (showing the poly A tail located in this region), of zebrafish (*D. rerio*) *ckma* and *ckmb* genes are shown in Table 1 and Table 2. For the creation of these tables, ENSDART00000032481.6 transcript for the *ckma* gene and the cDNAs of the ENSDART00000059366.7 transcript for the *ckmb* gene were used.

Zebrafish (*Danio rerio*) with medaka (*Oryzias latipes*), platy fish (*Xiphophorus maculatus*), spiny (*Gasterosteus aculeatus*), blowfish (*Fugu rupripes*) human (*Homo sapiens*) *ckma / CKM* proteins and zebrafish (*Danio rerio*) with cave fish (*Astyanax mexicanus*), eel (*Electrophorus electricus*), Mexican tetra fish (*Astyanax mexicanus*), spotted gar (*Lepisosteus oculatus*) human (*Homo sapiens*) ckmb / CKM proteins by aligning the in the Bioedit program, using CLUSTALW. Identity rates were calculated (Thompson et al., 1994) (Table 3,4).

Results and Discussion

The big effect of industrial enterprises wastes, oxygen deficiency is also a major factor in the increase in creatine in fish (Arslan, 2015). The stress responses of vertebrates also include different interactions between physiological pathways that can be characterized in both acute and chronic situations. Creatine kinase (CK); is an important enzyme used in the determination of damage to tissues and organs such as Glutamic pyruvic transaminase (GPT), glutamic oxaloacetic transaminase (GOT), alkaline phosphatase (ALP) and lactate dehydrogenase (LDH) enzymes.

These enzymes except from CK are liver enzymes and they are also used to understand liver-related problems. CK and GOT enzymes tend to increase in fish skin wounds, muscle tissue and brain disorders. In addition, the CK enzyme provides the renewal of ATP in the contraction or transport systems. Therefore, it is of great importance to complete a detailed bioinformatics study of *ckma* and *ckmb* genes, which are the stress markers in fish whose acute or chronic stress response varies with environmental differences (Iwama et al., 1999) in great importance as a model organism zebrafish (*D. rerio*).

Table 1. Nucleotide sequence of Zebrafish (*Danio rerio*) *ckma* gene

```
5'atggatatgggaaggaagggggcaccacccacagctgccacctcatctaggatgcct
ggggcctaaattgaagccttcttacactaaacagggcataagagaccagcgccagccaa
tcataattcagtgagctctaaaatgggccagccaatggctgcaggggctagaggtaTATA
                      +1
tatccaaatcaaactcttcttgCTTGGGTGACCCCTATTTCGGCTTGGTGAACAGGATCT
GATCCCAAGGACTGTTACCACTTTTGTTGTCTTTTGTGCAGgtaaa'N1538'atcagTG
TTAGAAACGCAATCATGCCTTTCGGAAACACCCACAACAACTTCAAGCTGAACTACTCAG
              -M--P--F--G--N--T--H--N--N--F--K--L--N--Y--S--
TTGATGAGGAGTATCCAGACCTTAGCAAGCACAACAACCACATGGCCAAGGTGCTGACTA
V--D--E--E--Y--P--D--L--S--K--H--N--N--H--M--A--K--V--L--T--
AGGAAATGTATGGCAAGCTTAGGGACAAGCAGACCTCCACTGGATTCACTGTGGATGATG
K--E--M--Y--G--K--L--R--D--K--Q--T--S--T--G--F--T--V--D--D--
TCATCCAGACCGGTGTTGACAATCCAGgtgag'N95'tccagGCCACCCCTTCATCATGA
V--I--Q--T--G--V--D--N--P--                 G--H--P--F--I--M--
CCGTCGGCTGTGTTGCTGGTGATGAGGAGTCCTACGAAGTGTTCAAGGATCTGTTCGACC
T--V--G--C--V--A--G--D--E--E--S--Y--E--V--F--K--D--L--F--D--
CCGTCATTTCCGACCGTCACGGTGGATACAAGGCAACTGACAAGCACAAGACCGACCTCA
P--V--I--S--D--R--H--G--G--Y--K--A--T--D--K--H--K--T--D--L--
ACTTTGAGAACCTGAAGgtaca'N783'tgtagGGTGGTGATGACCTGGACCCCAACTAC
N--F--E--N--L--K-                  -G--G--D--D--L--D--P--N--Y-
GTCCTGAGCAGCCGTGTGCGTACCGGACGCAGCATCAAGGGATACGCCCTGCCCCCCCAC
-V--L--S--S--R--V--R--T--G--R--S--I--K--G--Y--A--L--P--P--H-
AACAGCCGTGGAGAGCGCAGAGCTGTGGAGAAGCTGTCTGTTGAAGgtctg'N971'tcc
-N--S--R--G--E--R--R--A--V--E--K--L--S--V--E--
agCTCTGAGCAGCTTGGATGGAGAGTTCAAGGGCAAGTACTACCCCCTGAAGTCCATGAC
  A--L--S--S--L--D--G--E--F--K--G--K--Y--Y--P--L--K--S--M--T
TGATGCCGAGCAGGAGCAGCTGATCGCTGACCACTTCCTCTTTGACAAACCCGTCTCCCC
--D--A--E--Q--E--Q--L--I--A--D--H--F--L--F--D--K--P--V--S--P
CCTGCTGCTGGCTGCTGGTATGGCCCGTGACTGGCCCGATGCCAGAGGCATTTGgtgag'
--L--L--L--A--A--G--M--A--R--D--W--P--D--A--R--G--I—W
N555'tatagGCACAATGAGAACAAGACCTTCCTGGTCTGGGTGAACGAGGAGGATCACC
          --H--N--E--N--K--T--F--L--V--W--V--N--E--E--D--H--
TGCGTGTCATTTCCATGCAGAAGGGTGGCAACATGAAGGAAGTGTTCAAGCGCTTCTGCG
L--R--V--I--S--M--Q--K--G--G--N--M--K--E--V--F--K--R--F--C--
TTGGTCTTCAGAGGgtatg'N79'gatagATTGAGGAAATTTTCAAGAAGCACAACCATG
V--G--L--Q--R-               -I--E--E--I--F--K--K--H--N--H--
GGTTCATGTGGAACGAGCATCTTGGTTTCGTCCTGACCTGCCCCTCCAACCTGGGCACAG
G--F--M--W--N--E--H--L--G--F--V--L--T--C--P--S--N--L--G--T--
GCCTGCGCGGTGGAGTCCACGTCAAGCTGCCCAAGCTCAGCACACATGCCAAGTTTGAGG
G--L--R--G--G--V--H--V--K--L--P--K--L--S--T--H--A--K--F--E--
AGATCCTGACCAGACTGCGCCTGCAGAAGCGTGGCACAGgtata'N93'ctcagGTGGTG
E--I--L--T--R--L--R--L--Q--K--R--G--T--                G--G--
TGGACACTGCCTCCGTTGGTGGAGTGTTTGACATTTCCAACGCTGACCGTATCGGCTCTT
V--D--T--A--S--V--G--G--V--F--D--I--S--N--A--D--R--I--G--S--
CAGAGGTTGAGCAGGTGCAGTGTGTGGTTGATGGTGTCAAGCTGATGGTGGAGATGGAGA
S--E--V--E--Q--V--Q--C--V--V--D--G--V--K--L--M--V--E--M--E--
AGAAGCTGGAGAAGGGCGAGTCCATCGACAGCATGATCCCTGCCCAGAAGTAAagcggga
K--K--L--E--K--G--E--S--I--D--S--M--I--P--A--Q--K--*-
gctcttccatttttttctcgtctttgtctgtttttttacagtccaacagcaatgcagagg
aaaactgctgctcaaaaagacagtctcacctttgcacctgtcttctttcctttttttcc
cttcttctctaatttccatgtcatttcgccatctttttttccactttgtttcctattaag
tcggtaacatcttgggatcagatacccggcgcaggagtgagtgcctgttgctgaggcttc
acctcaatttcagccttggttgtaaaaagtgaatcaatcaaagttgtatttcAATAAAAA
taccccataaaaca 3'
```

* The exons of Zebrafish (*D. rerio*) *ckma* gene are shown in capital letters, starting point of transcription with +1, 5 'upstream sequence, 3' downstream sequence are shown in lowercase letters. The first five nucleotides and the last 5 nucleotides of the introns are shown in lowercase letters and in red, and the length of the intron other than these nucleotides is given. The TATA box and poly adenylation signal (AATAAAAA) are shown in capital letters and colored yellow. Stop codon (TAA) is indicated by asterisk.

Table 2. Nucleotide sequence of Zebrafish (*Danio rerio*) *ckmb* gene

```
5'gtgcctacagtttatttcataactgagcacagaggaataggttaccgtggggtgggga
cagctgtctccccttctcgggtgttgtgatgtagcttctttccttttagaacctgaaacc
cccatcaagaagtgcagctatccaatgaagttcaagcatgcaggagcgggacaaccaatg
                                                        +1
ggcagcaaacattgaagTATAtaaaccaaggtctgacgtcttgcaaaGTTGGGCAACCCC
TACATTGGCTTGGTGAACAGGATCTGATCCCAAGGACTGTTACCTTTTCTCCCATACAGg
taaa'N1073'ggtagATTGCCGGAAAAGTGCAATCATGACTAAGAACTGCAACAACGAT
                                          -M--T--K--N--C--N--N--D-
TACAAGATGAAGTTTGCTGTGGATGAGGAGTTTCCTGACCTCTCCCAGCACAACAACCAT
-Y--K--M--K--F--A--V--D--E--E--F--P--D--L--S--Q--H--N--N--H-
ATGTCCAAGGTCCTCACCAAGGACATCTACAACAAGCTCAGGAGCAAGTCAACCCCCAGT
-M--S--K--V--L--T--K--D--I--Y--N--K--L--R--S--K--S--T--P--S-
GGATTCACCCTTGATGACTGCATCCAGACTGGTGTGGACAACCCTGgtaag'N318'aca
-G--F--T--L--D--D--C--I--Q--T--G--V--D--N--P--
agGTCACCCCTTCATCATGACTGTCGGCTGTGTGGCTGGGGATGAGGAGTCCTATGAAGC
 G--H--P--F--I--M--T--V--G--C--V--A--G--D--E--E--S--Y--E--A
CTTCAAGGAGTTGTTCGACCCCGTCATTTCTGACCGTCATGGTGGCTATAAGCCCACCGA
--F--K--E--L--F--D--P--V--I--S--D--R--H--G--G--Y--K--P--T--D
CAAGCATCTTACTGATCTGAACTGGGAGAACCTGAAGgtatg'N57'cacagGGTGGTGA
--K--H--L--T--D--L--N--W--E--N--L--K-                -G--G--D
TGATCTTGACCCCAACTACGTTCTGAGCAGCCGCGTACGTACCGGCCGCAGCATCAAGGG
--D--L--D--P--N--Y--V--L--S--S--R--V--R--T--G--R--S--I--K--G
ATTCACCCTGCCTCCTCACAACAGCCGTGGTGAGCGCAGAGCTGTGGAGAAGCTGTCCAT
--F--T--L--P--P--H--N--S--R--G--E--R--R--A--V--E--K--L--S--I
TGAGGgtatg'N69'cacagCTCTGAACAGCCTGGATGGTGAGTTCAAGGGCAAGTACTA
--E--                A--L--N--S--L--D--G--E--F--K--G--K--Y--Y
CCCACTGAAGGACATGACTGACAAGGAGCAGGAGCAGCTCATTGCTGACCACTTCCTGTT
--P--L--K--D--M--T--D--K--E--Q--E--Q--L--I--A--D--H--F--L--F
TGACAAGCCTGTGTCCCCCCTGCTGTTGGCTGCTGGCATGGCCCGTGACTGGCCTGACGG
--D--K--P--V--S--P--L--L--L--A--A--G--M--A--R--D--W--P--D--G
TAGAGGTATCTGGCACAACGACAACAAGACCTTCCTTGTGTGGGTGAACGAGGAGGATCA
--R--G--I--W--H--N--D--N--K--T--F--L--V--W--V--N--E--E--D--H
CCTCCGTGTCATCTCTATGCAGAAGGGTGGCAACATGAAGGAGGTCTTCAAGAGGTTCTG
--L--R--V--I--S--M--Q--K--G--G--N--M--K--E--V--F--K--R--F--C
CGTTGGCCTGCAGAAGgtaaa'N77'tgcagATTGAGGATGTTTTCAAGAAGCACAACCA
--V--G--L--Q--K-                -I--E--D--V--F--K--K--H--N--H
CGGTTTCATGTGGAACGAGCATCTTGGTTTCATCTTGACCTGCCCCTCTAACTTGGGTAC
--G--F--M--W--N--E--H--L--G--F--I--L--T--C--P--S--N--L--G--T
CGGTCTGCGCGGTGGTGTCCACGTCAAGCTGCCCAAACTCAGCACACATGCCAAATTTGA
--G--L--R--G--G--V--H--V--K--L--P--K--L--S--T--H--A--K--F--E
GGAGATCCTGACCAGACTGCGTCTTCAGAAGCGTGGCACAGgtgag'N83'tccagGTGG
--E--I--L--T--R--L--R--L--Q--K--R--G--T--                G--G
TGTGGACACAGCCTCCGTCGGTGGTGTGTTCGACATCTCAAACGCTGACCGTCTGGGCTC
--V--D--T--A--S--V--G--G--V--F--D--I--S--N--A--D--R--L--G--S
CTCTGAGGTGCAGCAGGTGCAGCTGGTGGTCGATGGTGTGAAACTCATGGTTGAAATGGA
--S--E--V--Q--Q--V--Q--L--V--V--D--G--V--K--L--M--V--E--M--E
AAAGAAACTGGAGAAGGGCGAGTCCATCGATGACATGATCCCTGCCCAGAAGTAAactgc
--K--K--L--E--K--G--E--S--I--D--D--M--I--P--A--Q--K--*-
caaaagtgtttcttttttttatttgttcatgcagtcgtgccatcatgacttttccaaaag
acactccaccctgcttcctttgtcctactttttttttctccttcctttctttcctgtcg
tcttctttttccagcacttttttcttcttatcttgctgttcctcctgtaacatcctgggat
caaacatccacatagctagaaacccctagctgtgttgtattcacctaaacttttttggtt
gtacaaagtttAATAAAcggccccatgtaacaatcaa 3'
```

* The exons of Zebrafish (*D. rerio*) *ckmb* gene are shown in capital letters, starting point of transcription with +1, 5 'upstream sequence, 3' downstream sequence are shown in lowercase letters. The first five nucleotides and the last 5 nucleotides of the introns are shown in lowercase letters and in red, and the length of the intron other than these nucleotides is given. The TATA box and poly adenylation signal (AATAAAAA) are shown in capital letters and colored yellow. Stop codon (TAA) is indicated by asterisk.

Since fish are aquatic organism, changes in the qualitative and quantitative properties of water can cause changes in the functional structure of the proteins of fish, therefore, from time to time, the protein folds can be opened and these proteins can combine with other proteins in the cell to form clumps.

As a result of this situation, proteins can lose their functions due to conformational deformation (Basu et al., 2000). Therefore, in this study, it was determined that *ckma* and *ckmb* genes are functional genes in zebrafish (*D. rerio*) with bioinformatics tools before the other bioinformatic studies such as determining the gene structure, creating a phylogenetic tree, constructing a preserved gene synthin and calculating the similarity-identity ratios with orthologists of zebrafish.

Bioinformatics studies should be completed before experimental studies in order to understand how the expression of genes changes with various stress factors in molecular studies. Therefore, this study will provide important bioinformatics data both for fish physiology studies and for studies on other vertebrates since zebrafish (*D. rerio*) is a model organism.

In this study, first of all, ENSEMBL, UNIPROT and NCBI databases and computerized algorithms such as BioEdit software, BLOSUM62 matrix program and MEGA6 program were used to reach and evaluate some data such as cDNAs, exons and introns of *ckma* and *ckmb* genes, amino acids produced by these genes, 5 'UTR and 3'UTR regions, chromosomes and locations where genes are located, protein sequences required for determining their phylogenetic affinity with other vertebrates. It was determined that zebrafish *ckma* gene has 8 exons and 7 introns while *ckmb* gene has 7 exons and 6 introns before the gene structures of these two genes were designed (Table 1, 2).

Sequence similarity was calculated to investigate the orthology of zebrafish *ckma* and *ckmb* genes with some other vertebrates. For this purpose, zebrafish (*D. rerio*) with protein sequences produced by *ckma* and *ckmb* genes, medaka (*Oryzias latipes*), platy fish (*Xiphophorus maculatus*), stickleback (*Gasterosteus aculeatus*), puffer fish (*Fugu rubripes*) cave fish (*Astyanax mexicanus*), eel (*Electrophorus electricus*), Mexican tetra fish (*Astyanax mexicanus*), spotted gar (*Lepisosteus oculatus*) and human *(Homo sapiens*) *ckma* and *ckmb*/*CKM* genes were sequenced using the Bioedit program in the BLOSUM62 matrix algorithm and the similarity of these organisms. Identity rates were calculated (Gromiha, 2010). Analysis results include zebrafish (*D. rerio*) with *ckma* gene, medaka (*O. latipes*) 93-97%, platy fish (*X. maculatus*) and stickleback (*G. aculeatus*) 90-96%, puffer fish (*F. rubripes*) and cave fish (*A. mexicanus*) 88-93%, human (*H. sapiens*) 88-93% and zebrafish (*D. rerio*) *ckmb* 88-94% (Table 3), zebrafish (*D. rerio*) with the *ckmb* gene medaka (*O. latipes*) 94-97%, eel (*E. electricus*) 93-97%, zebrafish (*D. rerio*) *ckma* gene 88-95%, human (*H. sapiens*) 84-92%, Mexican tetra fish (*A. mexicanus*) 82-85%, spotted gar (*L. oculatus*) showed 82-91% similarity-identity ratio (Table 4).

In order to determine the conserved genes of zebrafish (*D. rerio*) with medaka (*O. latipes*) and human (*H. sapiens*), it was first determined from the Ensembl genome database that the *ckma* and *ckmb* genes of this organism are on the 5th and 15th chromosomes and other genes located in these chromosomes were determined and their locations determined.

Table 3. Similarity-identity ratio of the zebrafish *ckma* gene with the other teleosts and human

```
Zf Ckma      1   MPFGNTHNNFKLNYSVDEEYPDLSKHNNHMAKVLTKEMYGKLRDKQTSTGFTVDDVIQTG
Me Ckma      1   ................D.F................L...M.....P....L.......
Pf Ckma      1   ..............K.E..F..............N.DI.A.......PS.Y.L.......
St Ckma      1   ..............K.ED.F................L...I..R..PS.Y.L.......
Fu Ckma      1   .AK-.C..DY.MKFA....F....Q............I.....G.S.PS...........
Hu Ckm       1   ........K.....KPE..................L.L.K.....E.PS...........
Zf Ckmb      1   .TK-.CN.DY.MKFA....F....Q.....S.....DI.N...S.S.PS...L..C....

Zf Ckma     61   VDNPGHPFIMTVGCVAGDEESYEVFKDLFDPVISDRHGGYKATDKHKTDLNFENLKGGDD
Me Ckma     61   I.........................L............P..................
Pf Ckma     61   ..........................E.L..I.........P..................
St Ckma     61   ..........................E.L............P.......M..........
Fu Ckma     60   ........................A....L............P..................
Hu Ckm      61   ..........................E....I.........P.........H........
Zf Ckmb     60   ........................A..E..............P....L....W........

Zf Ckma    121   LDPNYVLSSRVRTGRSIKGYALPPHNSRGERRAVEKLSVEALSSLDGEFKGKYYPLKSMT
Me Ckma    121   ................................I....I.....................
Pf Ckma    121   ....................T............I...A....T.................
St Ckma    121   ...................FT............I.........T.................
Fu Ckma    120   ...................FT............I....I...A.............TG..
Hu Ckm     121   ....................T....C................N..T.............
Zf Ckmb    120   ...................FT.................I...N..............D..
```

Table 3 continued

```
Zf Ckma     181 DAEQEQLIADHFLFDKPVSPLLLAAGMARDWPDARGIWHNENKTFLVWVNEEDHLRVISM
Me Ckma     181 ........S.............TC.........G......D..................
Pf Ckma     181 ........S.............TC................DD..................
St Ckma     181 ........N.............TC.........G...M..D..................
Fu Ckma     180 ........................TC.........G......D..S..............
Hu Ckm      181 EK..Q...D...............S..............D..S..............
Zf Ckmb     180 .K..............................G......D..................

Zf Ckma     241 QKGGNMKEVFKRFCVGLQRIEEIFKKHNHGFMWNEHLGFVLTCPSNLGTGLRGGVHVKLP
Me Ckma     241 ......R...R.......K...................YI...................
Pf Ckma     241 ......R...R.......K...................YI...................
St Ckma     241 ......R...........K...................YI...................
Fu Ckma     240 ..........R.......K..A................YI...................
Hu Ckm      241 E.........R.......K.......AG.P....Q...Y...................A
Zf Ckmb     240 ..................K..DV................I...................

Zf Ckma     301 KLSTHAKFEEILTRLRLQKRGTG-GVDTASVGGVFDISNADRIGSSEVEQVQCVVDGVKL
Me Ckma     301 .....P.................-..................L.....A...L.......
Pf Ckma     301 .....P.................-..................L.....D...L.......
St Ckma     301 .....P..D.............S..................L.........L.......
Fu Ckma     300 ....QP.................-..................L.........L.......
Hu Ckm      301 H..K.P.................-.....A..S...V.....L.........L.......
Zf Ckmb     300 ........................-..................L.....Q...L.......

                                                 Similarity(%)     Identity (%)

Zf Ckma     360 MVEME CKMAKKLEKGESIDSMIPAQK      100               100
Me Ckma     360 ............A.........          93                 97
Pf Ckma     360 ............A..G......          90                 96
St Ckma     361 ...L........A...L.....          90                 96
Fu Ckma     359 ...............G......          88                 93
Hu Ckm      360 ...........Q...D......          88                 93
Zf Ckmb     359 ...............D......          88                 94
```

* Amino acid sequence alignment of Zebrafish (Zf) Ckma with medaka (Me) Ckma, eel (Ee) Ckma, Zebrafish Ckmb, Human (Hu) Ckm, Mexican tetra fish (Mt) Ckma, and spotted gar (Sg) Ckma. The dots represent same amino acids with the first line and tires represent amino acids that are not specified. The percent identity and similarity between the zebrafish Ckma and the other teleosts and human Ckma/Ckm proteins is shown at the end of each sequence.

Table 4. Similarity-identity ratio of the zebrafish *ckmb* gene with the other teleosts and human

```
Zf Ckmb   1   MTK-NCNNDYKMKFAVDEEFPDLSQHNNHMSKVLTKDIYNKLRSKSTPSGFTLDDCIQTG
Cf Ckmb   1   ...-..H.......SLE........L.....A............G.........V..V....
Ee Ckmb   1   ...-..H.......SLE..Y..........A.A...E..E..............I....
Zf Ckma   1   .PFG.TH.NF.LNYS....Y....K.....A.....EM.G...D.Q.ST...V..V....
Hu Ckmb   1   .PFG.TH.KF.LNYKPE..Y....K.....A....LEL.K...D.E......V..V....
Mt Ckmb   1   ...-..H.......SLE........L.....A............G.........V..V....
Sg Ckmb   1   .PFG.TH.N..LN.S.........TK.....A.A......A...D.Q......V..V....

Zf Ckmb   60  VDNPGHPFIMTVGCVAGDEESYEAFKELFDPVISDRHGGYKPTDKHLTDLNWENLKGGDD
Cf Ckmb   60  ........................V..D.L.................H.............
Ee Ckmb   60  ........................V..D.L.................N.............
Zf Ckma   61  ........................V..D..............A....K....F........
Hu Ckm    61  ........................V.......I..............K....H........
Mt Ckmb   60  ........................V..D.L.................H.............
Sg Ckmb   61  ......................DV..D......E...N.F......K....FG.......

Zf Ckmb   120 LDPNYVLSSRVRTGRSIKGFTLPPHNSRGERRAVEKLSIEALNSLDGEFKGKYYPLKDMT
Cf Ckmb   120 ................L.........................SV................
Ee Ckmb   120 ......Q............Y.....................MT................
Zf Ckma   121 ...................YA.................V...S.............S..
Hu Ckm    121 ...................Y.....C............V......T...........S..
Mt Ckmb   120 ................L.........................SV................
Sg Ckmb   121 ...................Y.....C.......I..M..D...T.E.............
```

Table 4 continued

```
Zf Ckmb   180 DKEQEQLIADHFLFDKPVSPLLLAAGMARDWPDGRGIWHNDNKTFLVWVNEEDHLRVISM
Cf Ckmb   180 ................................A.........................
Ee Ckmb   180 E...D..................S........A...Y.....................
Zf Ckma   181 .A.............................A......E...................
Hu Ckm    181 E...Q...D..............S........A.........S...............
Mt Ckmb   180 ................................A.........................
Sg Ckmb   181 .E..D...R..............S........A......ND.................

Zf Ckmb   240 QKGGNMKEVFKRFCVGLQKIEDVFKKHNHGFMWNEHLGFILTCPSNLGTGLRGGVHVKLP
Cf Ckmb   240 .L............T......ET.....................................
Ee Ckmb   240 ..............................................................
Zf Ckma   241 ..................R..EI................V.....................
Hu Ckm    241 E.........R..........EI...AG.P....Q...YV...................A
Mt Ckmb   240 .L............T......ET.....................................
Sg Ckmb   241 ..........R...........L....GRS...S....Y.....................

Zf Ckmb   300 KLSTHAKFEEILTRLRLQKRGTGGVDTASVGGVFDISNADRLGSSEVQQVQLVVDGVKLM
Cf Ckmb   300 ..............................................................
Ee Ckmb   300 .....P........................................................
Zf Ckma   301 ........................................I.....E...C........
Hu Ckm    301 H..K.P.....................A..S...V...........E...........
Mt Ckmb   300 ..............................................................
Sg Ckmb   301 Q..K.P.....................AE.............F...E...M........

Zf Ckmb   360 VEMEKKLEKG-------ES----------IDDMIPAQK----------------------
Cf Ckmb   360 ..........-------..----------..N......----------------------
Ee Ckmb   360 I.........-------..----------.........----------------------
Zf Ckma   361 ..........-------..----------..S......----------------------
Hu Ckm    361 ..........-------Q.----------.........----------------------
Mt Ckmb   360 .........CSRTTFPRD.PPCSPSFFLF.LSLF.VAYNVFLLPSFLLSLSHSGHSGLRL
Sg Ckmb   361 I.........--AA---------------...IL--------------------------
```

			Similarity(%)	Identity(%)
Zf Ckmb	380	**----**	100	100
Cf Ckmb	380	**----**	94	97
Ee Ckmb	380	**----**	93	97
Zf Ckma	381	**----**	88	95
Hu Ckm	381	**----**	84	92
Mt Ckmb	420	**SGHV**	82	85
Sg Ckmb	377	**----**	82	91

* Amino acid sequence alignment of zebrafish (Zf) Ckmb with cave fish (Cf) Ckmb, eel (Ee) Ckmb, Zebrafish (Zf ckma, human (Hu) Ckm, Mexican tetra fish (Mt) Ckmb and spotted gar (Sg) Ckmb. The dots represent same amino acids with the first line and tires represent amino acids that are not specified. The percent identity and similarity between the zebrafish Ckma and the other teleosts and human Ckma/Ckm proteins is shown at the end of each sequence.

The genes which are on chromosome 5 (*ckma, mark4a, kptn, crx, nfkbib, alkbh6, nova1, micu2, rhogc, nccrp1*) and chromosome 15 (*ckmb klc3, ercc2, zc3h4, sae1, bbc3, rad1, ift20, tmem97*) in zebrafish are found in the medaka, human and spotted gar in different location. However, these locations were identified and recorded before designed the conserved gene syntheny (Figure 1). As can be seen in the figure 1, the genes mentioned are conserved on the 5th, 11th, 13th, 14th, 17th, 19th chromosomes in humans, 14th and 21st chromosomes in medaka, and the 2nd and 22nd chromosomes in spotted gar. As it is known, teleost fish have evolutionarily conserved regions for the gene structure in the same gene family, and the designed conserved gene syntheny clearly demonstrates this. In addition, when the results are examined, it is thought that the zebrafish creatine kinase gene emerged as a result of the teleost genome duplication seen in teleost fish. Teleost fish can have two copies of genes found as a single copy in other living things 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). When the Ensembl database is examined; zebrafish were found to have two copies of the creatine kinase gene, *ckma* and *ckmb*.

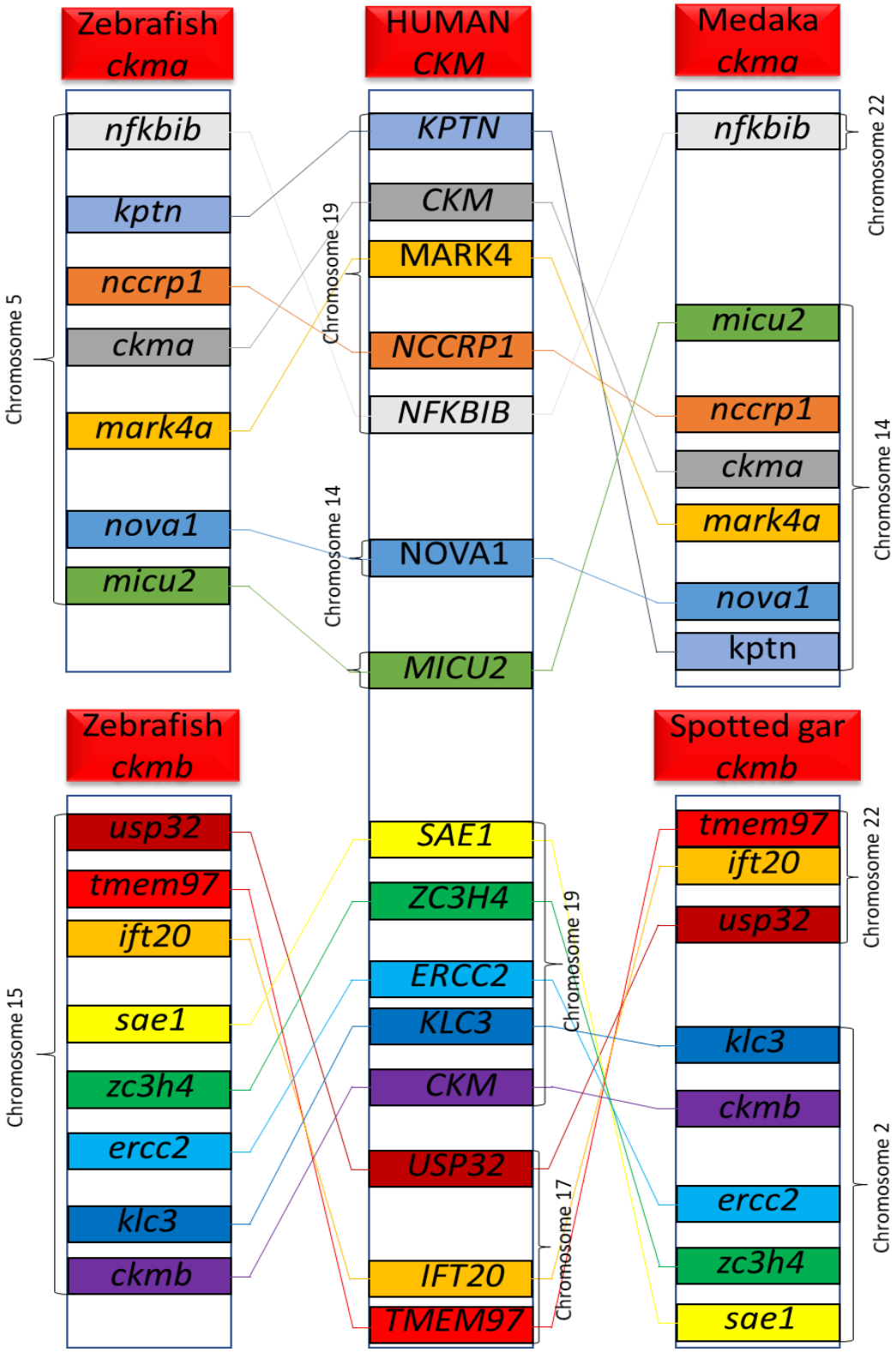

Figure 1. Conserved gene syntheny of Zebrafish (*Danio rerio*) *ckma* and *ckmb* genes

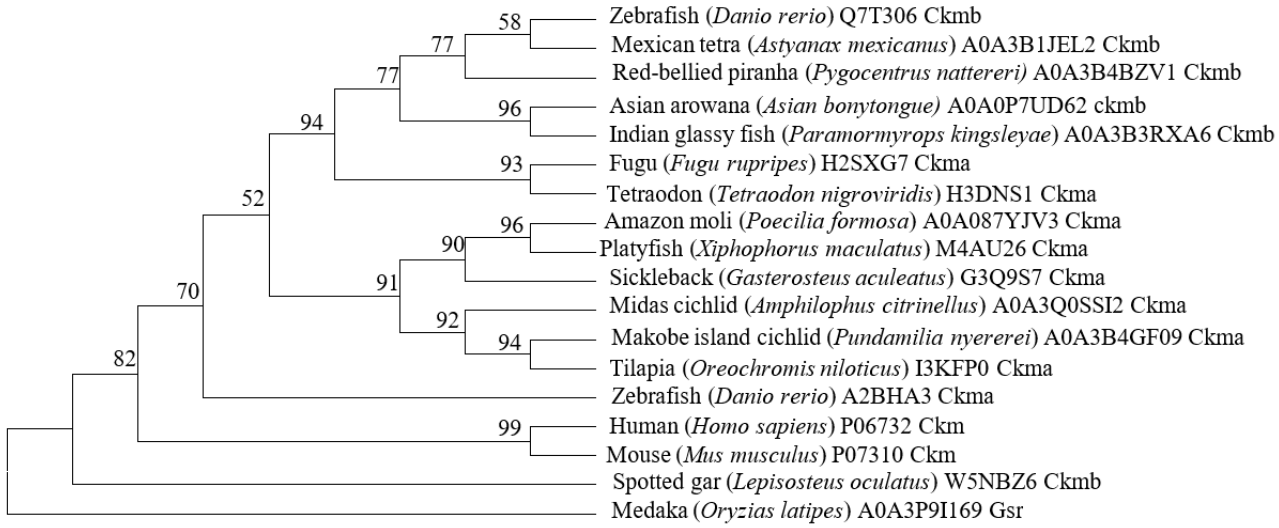

Figure 2. Phylogenetic tree of zebrafish (*Danio rerio*) *ckma* and *ckmb* genes

Phylogenetic tree was designed using protein sequences of Zebrafish Ckma A2BHA3 and Ckmb Q7T306 genes and the other vertebrates such as Amazon moli (*Poecilia formosa*) A0A087YJV3 Ckma, fugu (*Fugu rubripes*) H2SXG7 Ckma, human (*Homo sapiens*) P06732 *CKM*, Makobe island cichlid (*Pundamilia nyererei*) A0A3B4GF09 Ckma, Midas cichlid (*Amphilophus citrinellus*) A0A3Q0SSI2 Ckma, rat (*Mus musculus*) P07310 Ckm, platy fish (*Xiphophorus maculatus*) M4AU26 Ckma, stickleback (*Gasterosteus aculeatus*) G3Q9S7 Ckma, puffer fish (*Tetraodon nigroviridis*) H3DNS1 Ckma, spotted gar (*Lepisosteus oculatus*) W5NBZ6 Ckmb, Mexican tetra (*Astyanax mexicanus*) A0A3B1JEL2 Ckmb, Asian arovan (*Asian bonytongue*) A0A0P7UD62 Ckmb, Indian gladiolus (*Paramormyrops kingsleya*) A0A0P7UD62 Ckmb by obtaining from the UNIPROT genomic database, and the phylogenetic relationship between these genes was determined by Mega program and the maximum likelihood method (Felsenstein 1989). Medaka (*Oryzias latipes*) *gsr* gene (A0A3P9I169) was used as outgroup. In the phylogenetic tree, it was observed that *ckma* protein sequences clustered separately from *ckmb* protein sequences (Figure 2). The reliability of the tree was evaluated by a phylogenetic analysis with 1000 replicates (Felsenstein, 1989).

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: 1711-1714.

Arslan, H. (2015). Pestisit Sinerjisinin; Gökkuşağı Alabalıklarında (*Oncorhynchus mykiss*) Yüzme Performansı, Biyokimyasal Hematolojik, Histopatolojik Ve Genotoksik Etkilerinin Araştırılması, *Doktora Tezi*. Atatürk Üniversitesi, Erzurum, Turkey.

Basu S., Binder, R. J., Suto, R., Anderson, K. M., & Srivastava, P. K. (2000). Necrotic but not apoptotic heat shock proteins, which deliver a partial maturation signal to dentritic cells and activate the NF-kB pathway. Interntional Immunology, 12: 1539-1546.

Bayır, M., Arslan, G., & Yıldız Oğuzhan, 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.

Braasch, I., & Postlethwait, J. H. I. (2012). Polyploidy in fish and the teleost genome duplication. In: Soltis, P. S., Soltis, D. E. (eds.), Polyploidy and Genome Evolution. Springer. 341-383.

Carpio, Y., & Estrada, M. P. (2006). Zebrafish as a Genetic Model organism. Biotecnología Aplicada, 23: 265-270.

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: 682-689.

Felsenstein, J. (1989). PHYLIP-Phylogeny inference package. Cladistics, 5: 164-166.

Gilmour, D. T., Jessen, J. R., & Lin, S. (2002). Manipulating gene expression in the zebrafish. In: Zebrafish: A Practical Approach, Nusslein-Volhard C, Dahm R, (eds.). *Oxford:* Oxford University Press, p: 121-43.

Gromiha, M. M. (2010). Protein Bioinformatics: From Sequence to Function. Academic Press, 339 p.

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

Kell, A. J. E., Yamins, D. L. K., Shook, E. N., & Norman-Haignere, S. V. (2018). A Task-Optimized Neural Network Replicates Human Auditory Behavior, Predicts Brain Responses, and Reveals a Cortical Processing Hierarchy. Neron, 98(3): 630-644.

Lieschke, G. J., & Currie, P. D. (2007). Animal models of human disease: zebrafish swim into view. Nature Reviews Genetics, 8: 353-367.

Ma, C. (2004). Animal models of disease. Modern Drug Discovery, 7(6): 30-36.

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: 2667-2681.

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 Cell Biology, 11: 699-704.

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 Reserch, 10: 1890-1902.

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

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30(12): 2725.

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: 4673-4680.

Wu, C. L., Lin, T. H., Chang, T. L., Sun, H. W., Hui, C. F., & Wu, J. L. (2011). Zebrafish HSC70 promoter to express carp muscle-specific creatine kinase for acclimation under cold condition. Transgenic Research, 20: 1217-1226.