Marine Science
and Technology Bulletin

RESEARCH ARTICLE

In silico characterization of *bHLH* transcription factor genes in the genome of rainbow trout *(Oncorhynchus mykiss)*

Yasemin Celik Altunoglu[1*] iD • Gülsüm Dedeeli[1] iD

¹ Kastamonu University, Faculty of Engineering and Architecture, Department of Genetics and Bioengineering, 37150, Kastamonu, Turkey

ABSTRACT

The significance of seafood in nutrition has started to be better understood after the change in the understanding of nutrition in the world. One of the most common species cultivated in the world is rainbow trout (*Oncorhynchus mykiss*) from the origin of North America. Transcription factors are a group of proteins containing different functional components for the accomplishment of various activities. The basic helix (bHLH) domain is a highly preserved amino acid motif that characterizes a family of transcription factors. The *bHLH* gene family in the rainbow trout (*Oncorhynchus mykiss*) genome has been identified in the current study for the first-time using bioinformatics tools. According to the results, 441 *bHLH* genes (*OmybHLH*) were identified in the rainbow trout genome and the physicochemical properties of those proteins were determined. The highest number of the genes was in 7th chromosome of rainbow trout with 29 *OmybHLH* genes. 38 of *OmybHLH* genes had no intronic regions. OmybHLH proteins were divided into 4 main groups in the phylogenetic tree consistent with their motif content. The common biological function of OmybHLH proteins was the regulation of biological processes. The mode of action of OmybHLH proteins was binding activity. The *OmybHLH* gene family in the rainbow trout and the *bHLH* gene family in the Atlantic salmon (*SsabHLH*) had 95 orthologous gene relationships and average separation times of those orthologous genes were found to be 298 million years ago (MYA). Almost all the OmybHLH protein family members have dominated by the α-helix motif which is a stable conformation. Identification of the bHLH proteins and evaluation of their properties in rainbow trout can open new perspectives for aquaculture applications and fish culture to get better yield using genetic data.

* Corresponding author
E-mail address: yasemincelikbio@gmail.com (Y. Celik Altunoglu)

Introduction

Fish meat is a highly nutritional food and is rich in protein and unsaturated fatty acids, as well as containing essential amino acids such as methionine and lysine (İzci et al., 2009). Unsaturated fatty acids, which have benefits such as lowering blood cholesterol levels and preventing cardiovascular diseases, are abundant in fatty fish. Besides, iron, phosphorus and calcium are also abundant (Sayılı et al., 1999; Kocaman and Sayılı, 2014). The significance of seafood in nutrition has started to be better understood after the change in the understanding of nutrition in the world. Fish farming has gained importance because of the need for animal protein for the increasing population (Doğan and Güven, 2005).

In the beginning, carp cultivation was preferred because of its easy cultivation, and then the cultivation of sea bass, sea bream and trout species, which have an economic value over time, gained importance (Kocaman and Sayılı, 2014). One of the most common species cultivated in the world is rainbow trout (*Oncorhynchus mykiss*) from the origin of North America. Properties of rainbow trout, such as resistance to high temperatures and good adaptation to environmental conditions, easy to feed and good growth and having a shorter incubation period at higher temperatures than other trout types, provide easy adaptation to cultural conditions (Aydın, 2009). Trout, which is one of the most important freshwater fish, has become an important option in the market compared to marine fish in terms of both increasing the amount of production and being preferred (Yiğit and Aral, 1999).

Transcription factors are groups of proteins containing different functional components for the accomplishment of various activities such as DNA binding, activation, phosphorylation and protein oligomerization. The basic helix (bHLH) domain is a highly preserved amino acid motif that characterizes a family of transcription factors. The bHLH domain includes two regions: one of them includes about of 10–15 predominantly basic amino acids (the basic region) and other includes about 40 amino acids to construct two α-helices separated by a loop of variable length (the helix-loop-helix region) (Jones, 2004; Pires and Dolan, 2010). bHLH proteins are characterized by protein-protein interaction and highly protected areas for DNA binding (Murre et al., 1989; Atchley et al., 1999). The effects of these transcription factors are seen in many events such as neurogenesis, myogenesis, cell line detection, gender detection, cell proliferation and differentiation in organisms ranging from plants to mammals (Atchley et al., 1999).

Many of bHLH proteins have been detected in organisms ranging from yeast *Saccharomyces cerevisiae* to zebrafish and human (Robinson et al., 2000; Ledent et al., 2002; Wang et al., 2009), however, in our knowledge, there is no study about the *bHLH* gene family in rainbow trout. The *bHLH* gene family in the rainbow trout (*Oncorhynchus mykiss*) genome has been identified in the current study for the first-time using bioinformatics tools. Therefore, it was aimed to determine the properties of the bHLH proteins such as their chromosomal localization, motif regions, homology models, gene structure, the orthologous relationships between the Atlantic salmon (*Salmo salar*) and the predicated roles, molecular functions and cellular localization in the rainbow trout.

Material and Methods

Identification of bHLH Genes and Their Properties in Rainbow Trout Genome

CLC (Genomics Workbench 11) was used for the determination of rainbow trout *bHLH* proteins. BLASTp search was performed for derived protein sequences against NCBI (National Center for Biotechnology Information) database. Then, conserved domains were determined according to the Hidden Markov Model (HMM) and these regions were controlled in Pfam (https://pfam.xfam.org/) database. Repeated sequences were discarded and rainbow trout *bHLH* proteins were identified and numbered based on their chromosomal localization. Physicochemical characterization of identified proteins such as molecular weight, amino acid length, isoelectric point, etc. was determined by the ExPASy ProtParam tool (https://web.expasy.org/protparam/) (Gasteiger et al., 2005). Besides, genomic sequences of *bHLH* genes were searched against the NCBI (https://www.ncbi.nlm.nih.gov/) database to determine their chromosomal localization. The data obtained as a result of this search were visualized using MapChart (https://mapchart.net/) program (Voorrips, 2002). Gene Structure Display Server (Hu et al., 2014) was utilized to determine the exon-intron structure of rainbow trout *bHLH* genes.

Phylogenetic Tree Construction and Determination of Conserved Motifs

ClustalW algorithm was used for protein sequence alignment and the phylogenetic tree was created by the Maximum Likelihood method with bootstrap analysis for 1000 replicates in MEGAX (https://www.megasoftware.net/) program (Kumar et al., 2018).

MEME-Suite database, one of the motif determination tools, was used to define sequence motifs with short and repeating patterns in DNA that are assumed to have a biological

function (Bailey et al., 2015). The maximum motif length was selected as between 4-20 aa.

Gene Ontology Analysis

Functional data analysis of bHLH proteins belonging to rainbow trout was performed using Blast2GO software (Conesa et al., 2005). Analyses were carried out in three steps: in the first step, BLASTP search against the NCBI database was performed, in the second step, mapping according to the BLASTp results was accomplished (MAPPING) and in the third step, an information file about the sequences was prepared (ANNOTATION). At the end of those analyses, three categories have been created as predicted molecular functions, cellular locations and determination of biological functions for bHLH proteins in rainbow trout.

Identification of Orthologous Genes between Rainbow Trout and Atlantic Salmon (*Salmo salar*)

The amino acid sequences of bHLH proteins in rainbow trout and Atlantic salmon were aligned by the BLASTp algorithm. As a result of the BLAST query, genes meeting the ≥50 similarity condition and the expectation value of e^{-50} were selected. Orthologous of *bHLH* genes in Rainbow Trout and Atlantic salmon fish were aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) software (Li, 2003). The homologous (Ks) and non-homologous (Ka) exchange rates of the aligned orthologous protein sequences were calculated, and this calculation was performed with the PAL2NAL (http://www.bork.embl.de/pal2nal/) (Suyama et al., 2006) database.

Prediction of the Three-Dimensional Structure of bHLH Proteins in Rainbow Trout

Predicted three-dimensional structures of the bHLH proteins were shown by PHYRE2 (Protein Homology/Analog/YRecognition Engine; http://www.sbg.bio.ic.ac.uk/phyre2) database (Kelley et al., 2015). BlastP search was performed to search for similar sequences and the best result was determined based on the three-dimensional structure in Protein Data Bank (PDB).

Results and Discussion

Identification of bHLH Genes and Their Properties in Rainbow Trout Genome

After multiple searches for determination of *bHLH* genes in rainbow trout genome, 441 *bHLH* genes were identified and they were named from *OmybHLH-01* to *OmybHLH-441* based

on their chromosomal localization. Amino acid (aa) lengths of their protein product were between 79 and 2435 aa. The protein with the shortest aa sequence was found to be the OmybHLH-52 protein, and the protein with the longest aa sequence was OmybHLH-262. Considering the isoelectronic points (pI) of those proteins, these points varied between 4.49 and 11.02; it was found that 257 OmybHLH protein were acidic, 184 OmybHLH protein had basic properties. The percentage of the acidic OmybHLH proteins was 58.28%. The molecular weights of those proteins ranged from 9051.53 kDa to 259871.6 kDa. Besides, 97.96% of the OmybHLH proteins were unstable (Supplementary Material 1).

Considering of the previous studies, 183 in the rice (*Oryza sativa*) genome (Buck and Atchley, 2003; Wei and Chen, 2018) 35 in the worm (*Caenorhabditis elegans*) genome, 56 in Drosophila (*Drosophila melanogaster*) genome (Ledent and Vervoort, 2001), 52 in the silkworm (*Bombyx mori*) genome (Wang et al., 2007), 139 in the zebrafish (*Brachydanio rerio*) genome (Wang et al, 2009), 115 in the dog (*Canis lupus familiaris*) genome (Wang et al., 2015), 86 in the zebra finch (*Taeniopygia guttata*) genome (Liu and Zhao, 2011), 107 in the giant panda (*Ailuropoda melanoleuca*) genome (Dang et al., 2011), 57 in the ponerine ant (*Harpegnathos saltator*) genome (Liu et al., 2012), 48 in the jewel wasp (*Nasonia vitripennis*) genome (Liu et al., 2014), 109 in the pig (*Sus scrofa*) genome (Liu, 2015), 116 in the cattle (*Bos taurus*) genome (Li and Liu, 2017) were determined as bHLH proteins. When the results of this study conducted by different researchers are compared, it is seen that the number of *bHLH* genes found in organisms is close to each other, and the number of *bHLH* genes in the rainbow trout genome is higher than other organisms. This can be explained by the fact that bHLH proteins in rainbow trout can be involved in many key developmental processes and require many bHLH proteins to regulate these processes. Moreover, wide and detailed searches and high resolute genomes may be the reason for these high numbers.

Map Chart software was used to show the genomic distribution of *bHLH* genes on rainbow trout chromosomes. It was realized that the highest number of genes were carried by in the 7th chromosome of rainbow trout with 29 *OmybHLH* genes. This was followed by 8th, 5th, 4th and 3rd chromosomes with 26, 25 and 24 *OmybHLH* genes, respectively. The least number of genes were located on the 29th chromosome with 2 *OmybHLH* genes (Figure 1). Besides, chromosomal localizations of 11 *OmybHLH* genes were mapped in the scaffold level (Figure 2).
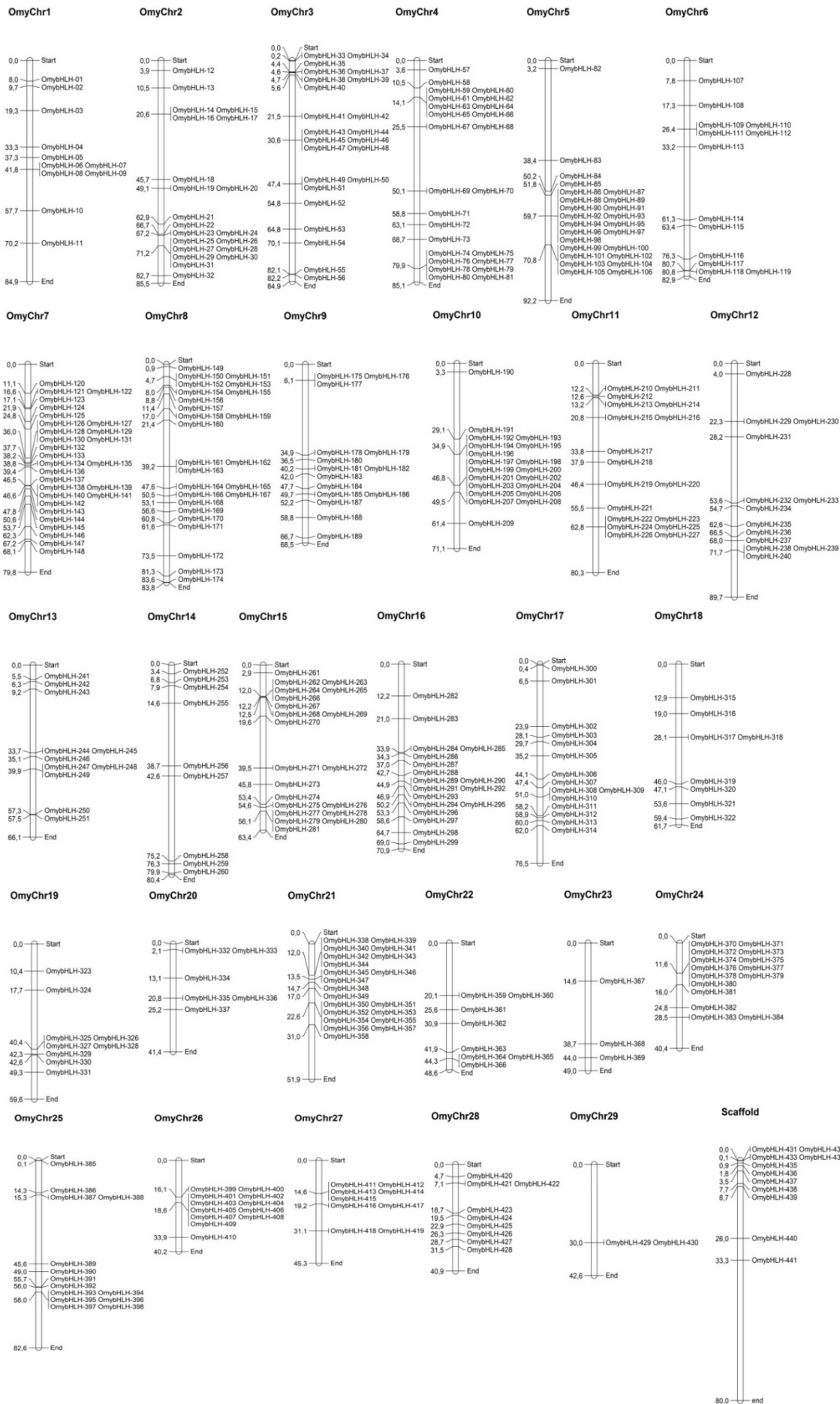
Figure 1. Presentation of OmybHLH genes localizations on rainbow trout chromosomes
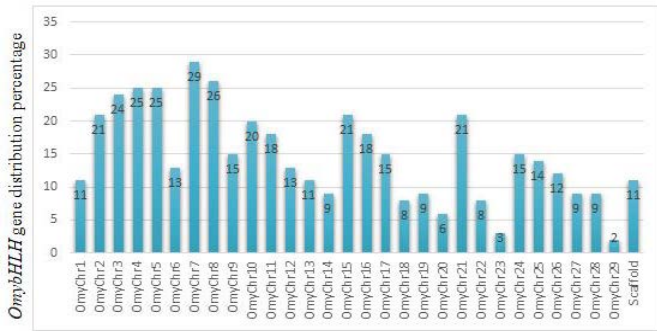
Figure 2. Chromosomal distribution of *OmybHL*H genes on rainbow trout genome

Exon regions are known as sequences encoded in DNA in eukaryotes, while intron regions appear as the parts of DNA that are not read. Exon-intron distribution of *bHLH* genes were determined by Gene Structure Display Server. According to the results, 38 of *bHLH* genes had no intron regions. Other genes had variable number of intron regions. Exon-intron of the genes were presented Supplementary Material 2.

Phylogenetic Tree and Conserved Motifs Analysis

Phylogenetic relationships are tree-like diagrams that show the kinship between species and how long the species split apart in the evolutionary process. The phylogenetic tree was drawn to evaluate evolutionary relations of OmybHLH proteins of rainbow trout. After visualization of the constructed phylogenetic tree by Interactive Tree of life (iTOL) online tool (Letunic and Bork, 2019), OmybHLH proteins are divided into 4 main groups (Class I, Class II, Class III, Class IV). It was determined that there were 63 proteins in Class I, 128 in Class II, 122 in Class III and 128 in Class IV. Class II and Class IV were found to contain an equal number of the OmybHLH proteins (Figure 3). The phylogenetic tree of OmybHLH proteins may provide information about protein sequence similarities, differences and their ancestors.

Protein sequence motifs are often used in genome research to identify and classify proteins, to determine specific binding sites in proteins and to find functional regions in proteins. Conserved motif structures in OmybHLH proteins identified in rainbow trout were determined and detailed results are shown in Supplementary Material 3. According to the MEME Suite database analysis to identify conserved motifs, 12 preserved motifs were defined for the 441 OmybHLH proteins. Those motif analyses verified the phylogenetic tree allocation. Proteins with similar motif compositions were especially in the same cluster in the phylogenetic tree.

Considering the studies on the relationship and evolutionary processes of bHLH protein families in different organisms, it was observed that atonal bHLH protein homologs in organisms such as mice, chickens, and humans are preserved throughout the evolutionary process and the coding regions are quite similar (Ben Arie, 1996). Besides, according to the phylogenetic analysis with 9 species of land plants, bHLH proteins found in these plants have come from a common ancestor (Pires and Dolan, 2010). In addition, it has been concluded that motifs of bHLH proteins found in mushrooms were preserved with aa sequences and were associated with bHLH proteins included in group B in animals (Sailsbery et al., 2012).

Gene Ontology Analysis

Gene ontology is one of the bioinformatics tools used to reveal the features and functionality of gene products in the species. Blast2GO software was utilized to define cellular localization, biological process and molecular function of OmybHLH proteins (Figure 4). According to the analysis results, OmybHLH proteins were found to have biological functions such as regulation of biological processes, roles in metabolic, cellular and developmental processes, multicellular organismal processes and response to a stimulus. Consistent with those results, in the literature, bHLH proteins are involved as positive regulators in biological processes (Norton, 2000), are involved as negative regulators (Benezra et al., 1990; Norton, 2000; Perk et al., 2005), are promoting myogenic cell proliferation and differentiation and are involved in determining flexibility in skeletal muscles and respond to mechanical or neuronal stimuli (Voytik et al., 1993; Molkentin and Olson, 1996; Puri and Sartorelli, 2000; Walters et al., 2000; Perry et al., 2001; Pownall et al., 2002; Buckingham et al., 2003; Ishido et al., 2004; Tapscott, 2005; Legerlotz and Smith, 2008). Also, they are involved in the nervous system development stages and function in neuronal cell development and differentiation in the brain (Campuzano, 1985; Guillemot et al., 1993; Guillemot, 1995; Yasunami et al., 1996; Borges et al., 1997; Miyata et al., 1999; Olson et al., 2001). Moreover, they play an important role in embryonic development and embryonic cell differentiation (Malecki et al., 1999; Norton, 2000).

When the cellular locations of OmybHLH proteins were evaluated, it was observed that they dispersed to different parts of the cell and this dispersion covered the cell, intracellular parts and as a cellular anatomical entity. There are studies in the literature about that bHLH proteins were found in embryonic cells (Nambu et al., 1991; Muralidhar vd., 1993), in muscle cells (Perry et al., 2001; Tapscott, 2005), in nerve cells (Guillemot, 2007; Jahan vd., 2010), in blood cells (Drake et al., 1997; Gering et al., 1998), in melanocyte, and mast cells (Hodgkinson et al., 1993; Steingrimsson et al., 2004; Levy et al., 2006).
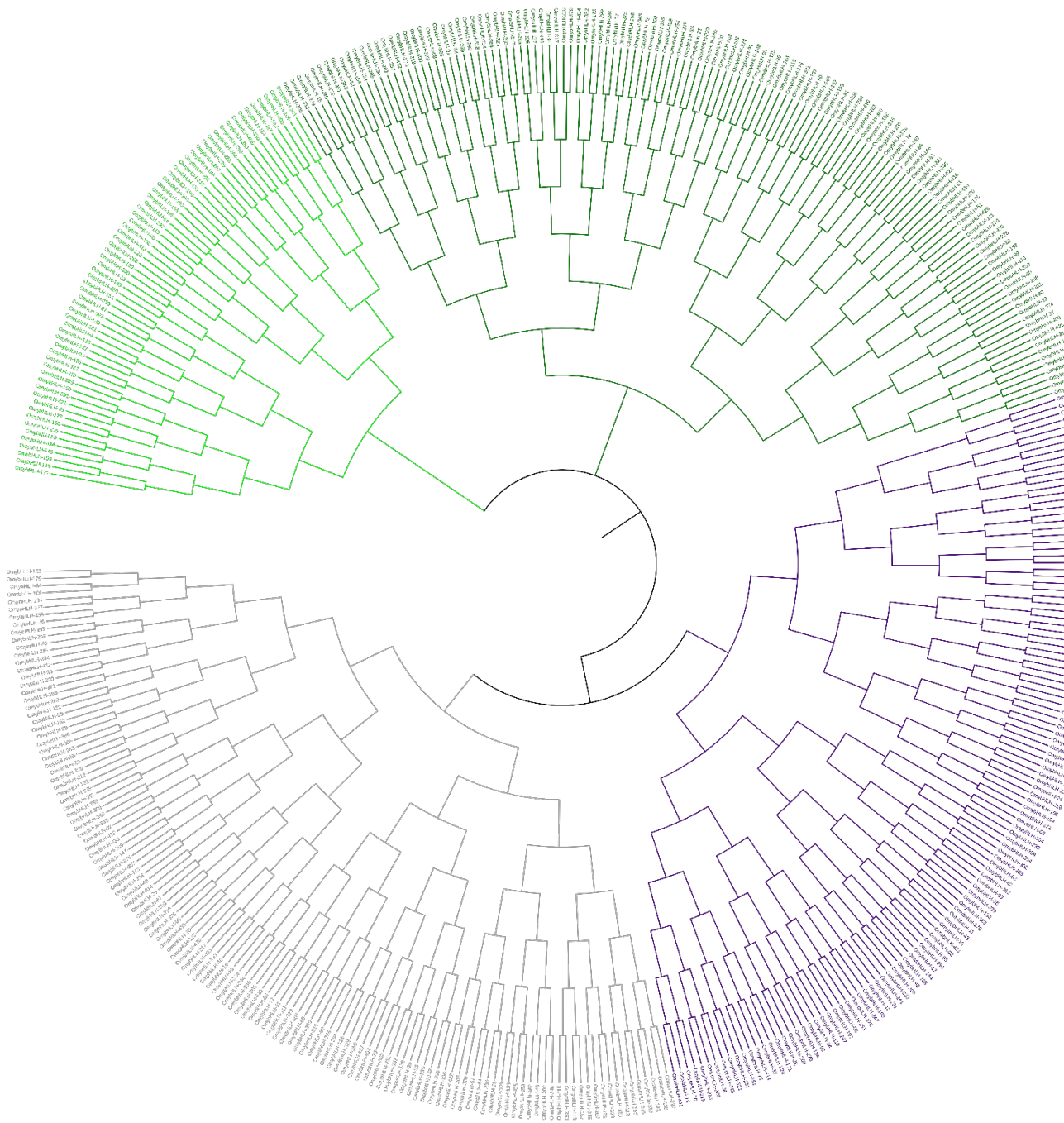
Figure 3. Phylogenetic relations of OmybHLH proteins. The phylogenetic trees were constructed by multiple sequence alignment using the ClustalW program via the MEGA X program (maximum likelihood method)

Mode of action of OmybHLH proteins was found to be the binding activity. In addition, those protein groups had transcription regulator activity which is consistent with their roles. In the literature, it has been revealed that members of the bHLH protein family show their molecular functions as binding in different studies (Ledent and Vervoort, 2001; Ledent et al., 2002; Berkes and Tapscott, 2005; Murre, 2019).

Considering the current results, OmybHLH proteins have significant roles in regulation and development processes in many cell types. High gene numbers of this protein family in rainbow trout may be explained by multiple significant roles of bHLH proteins in different cell types. Gene ontology analysis of OmybHLH proteins can draw a frame about the importance of these multifunctional proteins in the organisms and their predicated roles and localization in the cells.

Orthologous Genes between Rainbow Trout and Atlantic Salmon (Salmo salar)

Although they are found in different organisms, genes that have the same ancestral origins and have structural and

functional similarities but are separated from each other during the species formation are called orthologous genes. By comparing the rainbow trout and Atlantic salmon with the *bHLH* gene sequences, the orthologous gene analysis was performed and the separation times of the species were determined. Considering of the results, The *OmybHLH* gene family in the rainbow trout and the *bHLH* gene family in the Atlantic salmon (*SsabHLH*) had 95 orthologous gene relationships and the substitution ratio [non-synonymous (Ka) versus synonymous (Ks)] was 0.26 (Ka/Ks). Average separation times of those orthologous genes were found to be 298 million years ago (MYA) (Figure 5, Supplementary Material 4). According to the results, it can be estimated that there is a high degree of differentiation between the two fish species.
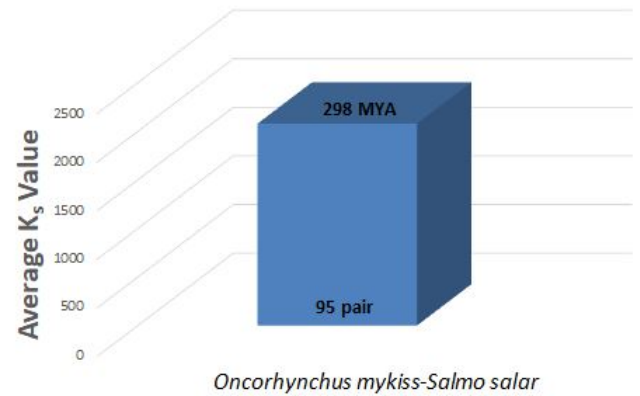

Figure 4. Predicated biological roles, molecular function and cellular localizations of OmybHLH proteins by Blast2GO analysis

Figure 5. Presentation of orthologous gene numbers of *OmybHLH* genes and Atlantic salmon *bHLH* genes

Prediction of the Three-Dimensional Structure of bHLH Proteins in Rainbow Trout

Homology is based on the sequence similarity of one or more proteins of which structure is known, and the high sequence similarity of proteins in this context supports the hypothesis that they come from the same ancestor. Predicted three-dimensional structures of the OmybHLH proteins were determined by homology modelling in the Phyre2 database. This modeling was carried out based on >90% confidence interval under intensive mode. It was determined that the similarity level ranged between 4% and 100% and 12 OmybHLH proteins showed similarity over 90% (OmybHLH-45, 58, 145, 359, 370, 371, 372, 373, 380, 391, 424, 433) (Figure 6). According to the predicted three-dimensional structure of the proteins, it was realized that almost all of the protein family members were dominated by the α-helix motif which is a stable conformation formed by the rotation of the alpha carbon atoms of 4 to 50 amino acids in a spiral shape. This structure consists of two spirals, which include the basic region and the variable loop region (Pires and Dolan, 2010). The amino terminal end in the structure allows DNA binding, while the carboxy-terminal end provides dimerization. When modeled three-dimensional structures of OmybHLH proteins are evaluated, it is seen that these two amphipathic α-helices (H1, H2) produced by the dimerization component are present in almost all of these protein structures, which is consistent with the literature (Murre et al., 1989; Atchley et al., 1999).

Conclusion

The study evaluated the bHLH proteins in rainbow trout, which are significant transcription factor family in eukaryotes. Properties of this protein family such as their chromosomal localization, motif regions, gene structure, predicted three-dimensional structures, orthologous relationships with the
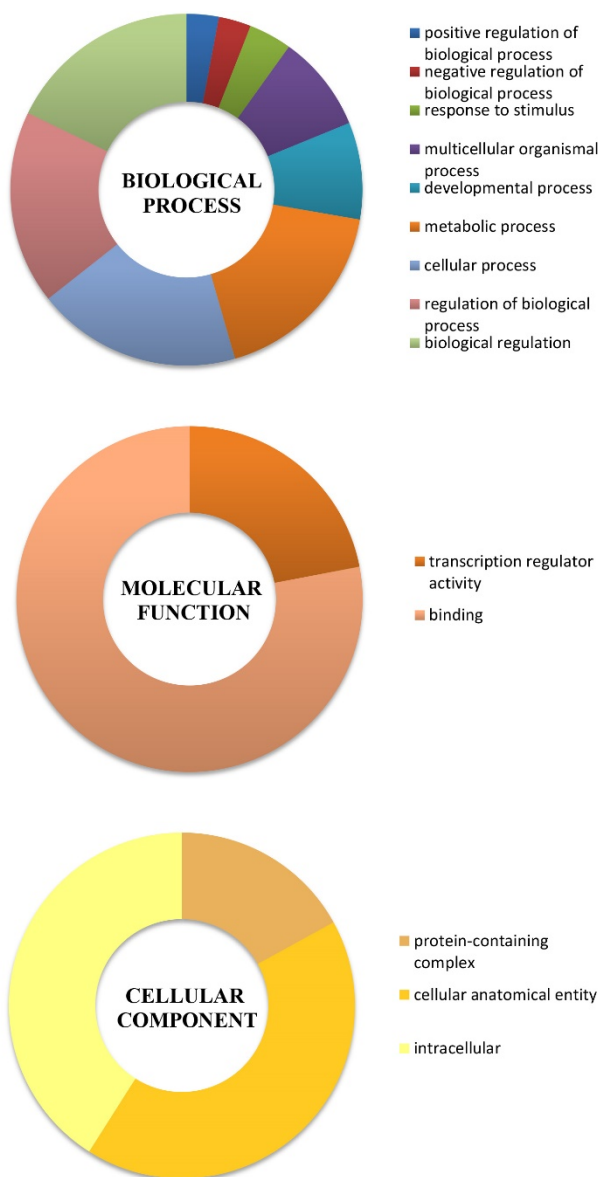
Atlantic salmon (*Salmo salar*) bHLH proteins and the predicted biological roles, molecular functions and cellular localization were determined using *in silico* analysis by bioinformatics tools. Identification of bHLH proteins and evaluation of their properties in rainbow trout can open new perspectives for aquaculture applications and fish culture to get better yield using genetic data. Besides, defined genetic data can be used in healthy and high-yield fish farming and those findings can be applied to other economically important fish species in aquaculture. From these views, the current study includes valuable data for OmybHLH proteins, which has potential usage in aquaculture implications.
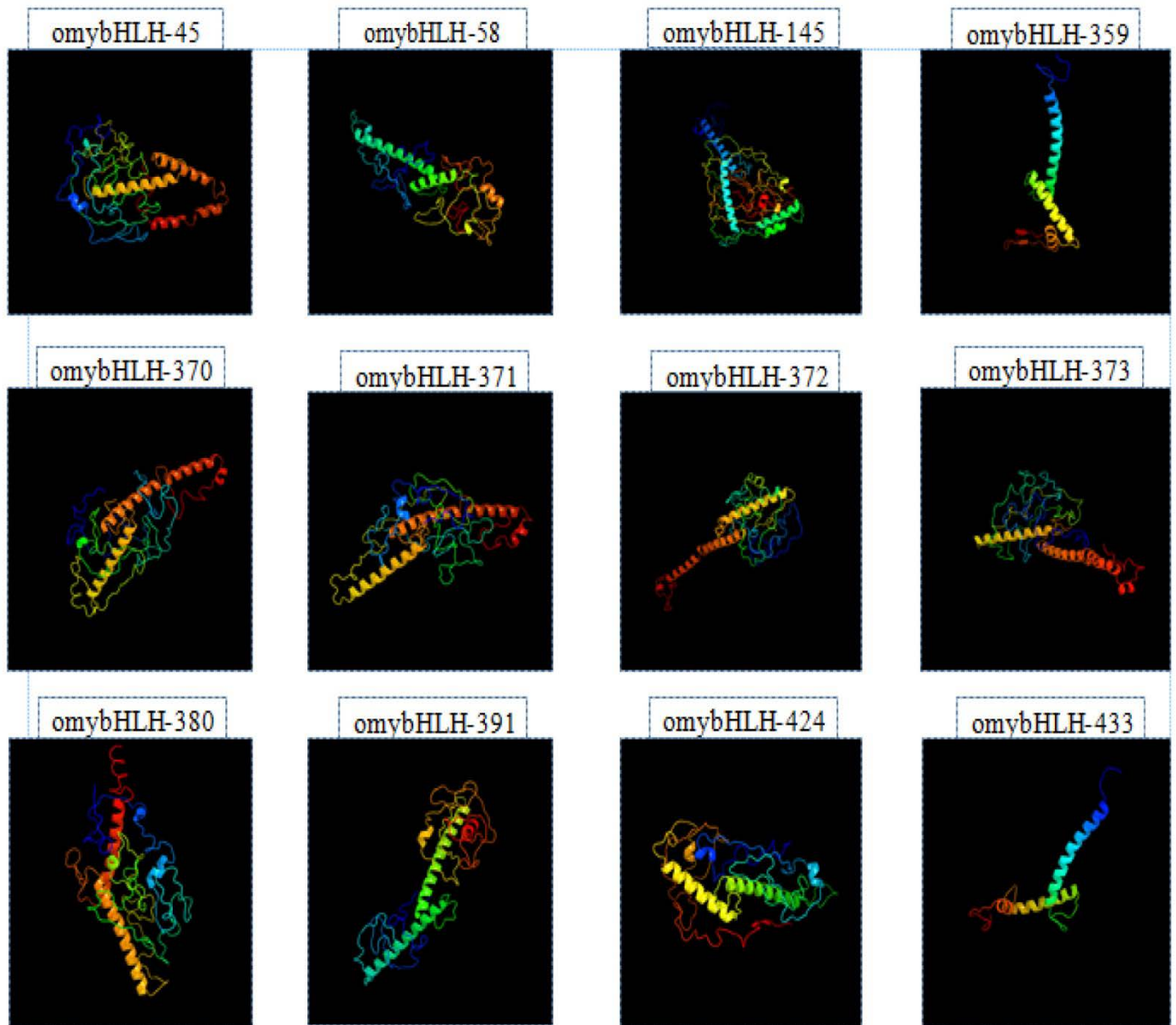

Figure 6. The predicated 3D structures of OmybHLH proteins

Compliance with Ethical Standards

Authors' Contributions

Author YCA designed the study, YCA and GD wrote the first draft of the manuscript, GD performed and managed the bioinformatics analyses. Both authors read and approved the final manuscript.

Conflict of Interest

The authors declare that there is no conflict of interest.

Ethical Approval

For this type of study, formal consent is not required.

Supplementary Materials

Supplementary data to this article can be found online at https://doi.org/10.33714/masteb.768233.

References

Atchley, W. R., Terhalle, W. & Dress, A. (1999). Positional dependence, cliques and predictive motifs in the bHLH protein domain. *Journal of Molecular Evolution*, *48*(5): 501–516. https://doi.org/10.1007/pl00006494

Aydın, F. (2009). Alabalık biyolojisi ve yetiştirme teknikleri. Ankara Üniversitesi Ziraat Fakültesi Su Ürünleri Bölümü, Ankara, Turkey. 30p.

Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. (2015). The MEME suite. *Nucleic Acids Research*, *43*(1): 39-49. https://doi.org/10.1093/nar/gkv416

Ben Arie, N. (1996). Evolutionary conservation of sequence and expression of the bHLH protein atonal suggests a conserved role in neurogenesis. *Human Molecular Genetics,* *5*(9): 1207–1216. https://doi.org/10.1093/hmg/5.9.1207

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. & Weintraub, H. (1990). The protein id: A negative regulator of helix-loop-helix DNA binding proteins. *Trends in Genetics,* *61*(1): 49-59. https://doi.org/10.1016/0168-9525(90)90161-x

Berkes, C. A. & Tapscott, S. J. (2005). MyoD and the transcriptional control of myogenesis. *Seminars in Cell, Developmental Biology,* *16*(4-5): 585–595. https://doi.org/10.1016/j.semcdb.2005.07.006

Borges, M., Linnoila, R. I., van de Velde, H. J. K., Chen, H., Nelkin, B. D., Mabry, M. & Ball, D. W. (1997). An Achaete-Scute homologue essential for neuroendocrine differentiation in the lung. *Nature, 386*(6627): 852–855. https://doi.org/10.1038/386852a0

Buck, M. J. & Atchley, W. R. (2003). Phylogenetic analysis of plant basic helix-loop helix proteins. *Journal of Molecular Evolution,* *56*(6): 742–750. https://doi.org/10.1007/s00239-002-2449-3

Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S. & Relaix, F. (2003). The Formation of Skeletal Muscle: From Somite to Limb. *Journal of Anatomy,* *202*(1): 59–68. https://doi.org/10.1046/j.1469-7580.2003.00139.x

Campuzano, S. (1985). Molecular genetics of the Achaete-Scute gene complex of D. *melanogaster. Cell, 40*(2): 327–338. https://doi.org/10.1016/0092-8674(85)90147-3

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M. & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics, 21*(18): 3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Dang, C., Wang, Y., Zhang, D., Yao, Q. & Chen, K. (2011). A genome-wide survey on basic helix-loop-helix transcription factors in giant panda. *PLoS One*, *6*(11): e26878. https://doi.org/10.1371/journal.pone.0026878

Doğan, K., & Güven, E. (2005). Ülkemizde (Türkiye) Su ürünleri yetiştiriciliği yapan işletmeler, üretim kapasiteleri, illere göre dağılımları ve ekonomik analizleri. *Su Ürünleri Mühendisleri Derneği Dergisi, 1*(4): 28-33. https://doi.org/1 0.22392/egirdir.414488

Drake, C. J., Brandt, S. J., Trusk, T. C., Little, C. D. (1997). TAL1/SCL Is Expressed in Endothelial Progenitor Cells/Angioblasts and Defines a Dorsal-to-Ventral Gradient of Vasculogenesis. *Developmental Biology,* *192*(1): 17–30. https://doi.org/10.1006/dbio.1997.8751

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*, *571–607.* https://doi.org/10.1385/1-59259-890-0:571

Gering, M., Rodaway, A. R. F., & Göttgens, B., & Patient, R. K. & Green, A. R. (1998). The SCL gene specifies haemangioblast development from early mesoderm. *The EMBO Journal,* *17*(14): 4029–4045. https://doi.org/10.1093/emboj/17.14.4029

Guillemot, F. (1995). Analysis of the role of basic-helix-loop-helix transcription factors in the development of neural lineages in the mouse. *Biology of the Cell, 84*(1-2): 3–6. https://doi.org/10.1016/0248-4900(96)81312-8

Guillemot, F. (2007). Spatial and temporal specification of neural fates by transcription factor codes. *Development, 134*(21): 3771–3780. https://doi.org/10.1242/dev.006379

Guillemot, F., Lo, L. C., Johnson, J. E., Auerbach, A., Anderson, D. J. & Joyner, A. L. (1993). Mammalian Achaete-Scute Momolog 1 is required for the early development of olfactory and autonomic neurons. *Cell, 75*(3): 463–476. https://doi.org/10.1016/0092-8674(93)90381-y

Hodgkinson, C. A., Moore, K. J., Nakayama, A., Steingrimsson, E., Copeland, N. G., Jenkins, N. A. & Arnheiter, H. (1993). Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic- helix- loop- helix- zipper protein. *Cell, 74*(2): 395-404.

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J. & Gao, G. (2014). GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics, 31*(8): 1296–1297. https://doi.org/10.1093/bioinformatics/btu817

Ishido, M., Kami, K. & Masuhara, M. (2004). In Vivo expression patterns of MyoD, p21, and Rb proteins in myonuclei and satellite cells of denervated rat skeletal muscle. *American Journal of Physiology, Cell Physiology*, *287*(2): 484-C493. https://doi.org/10.1152/ajpcell.00080.2004

İzci, L., Günlü, A. & Bilgin, Ş. (2009). Ülkemizde gökkuşağı alabalığı (*Oncorhynchus mykiss* Walbaum, 1792)'nın değerlendirilme şekilleri. *Eğirdir Su Ürünleri Fakültesi Dergisi, 5*(1-2): 73-79.

Jahan, I., Kersigo, J., Pan, N. & Fritzsch, B. (2010). Neurod1 regulates survival and formation of connections in mouse ear and brain. *Cell and Tissue Research, 341*(1): 95–110. https://doi.org/10.1007/s00441-010-0984-6

Jones, S. (2004). An overview of the basic helix-loop-helix proteins. *Genome Biology*, *5*(6): 226. https://doi.org/10.1186/gb-2004-5-6-226

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols, 10*: 845–858. https://doi.org/10.1038/nprot.2015.053

Kocaman, E. & Sayili, M. (2014). Gümüşhane ilinde gökkuşaği alabalik işletmelerinin ekonomik analizi. *Anadolu Tarım Bilimleri Dergisi*, *29*(1): 36-45. https://doi.org/10.7161/anajas.2014.29.1.36

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, *35*(6): 1547-1549. https://doi.org/10.1093/molbev/msy096

Ledent, V., Paquet, O. & Vervoort, M. (2002). Phylogenetic analysis of the human basic helix-loop-helix proteins. *Genome Biology, 3*(6): research0030.1-0030.18. https://doi.org/10.1186/gb-2002-3-6-research0030

Ledent, V. & Vervoort, M. (2001). The basic helix-loop-helix protein family: Comparative genomics and phylogenetic analysis. *Genome Research, 11*(5): 754–770. https://doi.org/10.1101/gr.177001

Legerlotz, K. & Smith, H. K. (2008). Role of MyoD in denervated, disused, and exercised muscle. *Muscle, Nerve, 38*(3): 1087–1100. https://doi.org/10.1002/mus.21087

Letunic, I. & Bork, P. (2019). Interactive tree of life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, *47*(W1): W256-W259. https://doi.org/10.1093/nar/gkz239

Levy, C., Khaled, M. & Fisher, D. E. (2006). MITF: Master regulator of melanocyte development and melanoma oncogene. *Trends in Molecular Medicine, 12*(9): 406–414. https://doi.org/10.1016/j.molmed.2006.07.008

Li, F. & Liu, W. (2017). Genome-wide identification, classification, and functional analysis of the basic helix–loop–helix transcription factors in the cattle, *Bos taurus*. *Mammalian Genome*, *28*(5-6): 176-197. https://doi.org/10.1007/s00335-017-9683-x

Li, K. B. (2003). ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics, 19*(12): 1585–1586. https://doi.org/10.1093/bioinformatics/btg192

Liu, A., Wang, Y., Dang, C., Zhang, D., Song, H., Yao, Q. & Chen, K. (2012). A genome-wide identification and analysis of the basic helix-loop-helix transcription factors in the ponerine ant, *Harpegnathos saltator*. *BMC evolutionary biology*, *12*(1): 165. https://doi.org/10.1186/1471-2148-12-165

Liu, W. (2015). Genome-wide identification, classification and functional analyses of the bHLH transcription factor family in the pig, *Sus scrofa*. *Molecular genetics and genomics*, *290*(4): 1415-1433. https://doi.org/10.1007/s00438-015-1007-9

Liu, W. & Zhao, C. (2011). Molecular phylogenetic analysis of zebra finch basic helix-loop-helix transcription factors. *Biochemical Genetics*, *49*(3-4): 226-241. https://doi.org/10.1007/s10528-010-9401-9

Liu, X. T., Wang, Y., Wang, X. H., Tao, X. F., Yao, Q. & Chen, K. P. (2014). A genome-wide identification and classification of basic helix-loop-helix genes in the jewel wasp, *Nasonia vitripennis* (Hymenoptera: Pteromalidae). *Genome*, *57*(10): 525-536. https://doi.org/10.1139/gen-2014-0171

Malecki, M. T., Jhala, U. S., Antonellis, A., Fields, L., Doria, A., Orban, T. & Krolewski, A. S. (1999). Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. *Nature Genetics, 23*(3): 323–328. https://doi.org/10.1038/15500

Miyata, T., Maeda, T. & Lee, J. E. (1999). NeuroD is required for differentiation of the granule cells in the cerebellum and hippocampus. *Genes Development, 13*(13): 1647–1652. https://doi.org/10.1101/gad.13.13.1647

Molkentin, J. D. & Olson, E. N. (1996). Defining the regulatory networks for muscle development. *Current Opinion in Genetics, Development, 6*(4): 445–453. https://doi.org/10.1016/s0959-437x(96)80066-9

Muralidhar, M. G., Callahan, C. A. & Thomas, J. B. (1993). Single-minded regulation of genes in the embryonic midline of the drosophila central nervous system. *Mechanisms of Development, 41*(2-3): 129–138. https://doi.org/10.1016/0925-4773(93)90043-w

Murre, C. (2019). Helix-loop-helix proteins and the advent of cellular diversity: 30 years of discovery. *Genes Development, 33*: 6-25. https://doi.org/10.1101/gad.320663.118

Murre, C., McCaw, P. S. & Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and Myc proteins. *Cell, 56*(5): 777–783. https://doi.org/10.1016/0092-8674(89)90682-x

Nambu, J. R., Lewis, J. O., Wharton, K. A. & Crews, S. T. (1991). The drosophila single- minded gene encodes a helix-loop- helix protein that acts as a master regulator of cns midline development. *Cell, 67*(6): 1157-1167. https://doi.org/10.1016/0092-8674(91)90292-7

Norton, J. D. (2000). Id helix loop helix proteins in cell growth, differentiation and tumorigenesis. *Journal of Cell Science, 113*(22): 3897-3905.

Olson, J. M., Asakura, A., Snider, L., Hawkes, R., Strand, A., Stoeck, J. & Tapscott, S. J. (2001). NeuroD2 is necessary for development and survival of central nervous system neurons. *Developmental Biology, 234*(1): 174–187. https://doi.org/10.1006/dbio.2001.0245

Perk, J., Iavarone, A. & Benezra, R. (2005). Id family of helix-loop-helix proteins in cancer. *Nature Reviews Cancer, 5*(8): 603–614. https://doi.org/10.1038/nrc1673

Perry, R. L., Parker, M. H. & Rudnicki, M. A. (2001). Activated MEK1 binds the nuclear MyoD transcriptional complex to repress transactivation. *Molecular Cell, 8*(2): 291–301. https://doi.org/10.1016/s1097-2765(01)00302-1

Pires, N. & Dolan, L. (2010). Origin and diversification of basic-helix-loop-helix proteins in plants. *Molecular Biology and Evolution, 27*(4): 862-874. https://doi.org/10.1093/molbev/msp288

Pownall, M. E., Gustafsson, M. K. & Emerson, C. P. (2002). Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annual Review of Cell and Developmental Biology, 18*(1): 747–783. https://doi.org/10.1146/annurev.cellbio.18.012502.1

Puri, P. L. & Sartorelli, V. (2000). Regulation of muscle regulatory factors by DNA-binding, interacting proteins, and post-transcriptional modifications. *Journal of Cellular Physiology, 185*(2): 155–173. https://doi.org/10.1002/1097-4652(200011)185:2<155::aid-jcp1>3.0.co;2-z

Robinson, K. A. & Lopes, J. M. (2000). Survey and summary: Saccharomyces cerevisiae basic helix–loop–helix proteins regulate diverse biological processes. *Nucleic Acids Research, 28*(7): 1499-1505. https://doi.org/10.1093/nar/28.7.1499

Sailsbery, J. K., Atchley, W. R. & Dean, R. A. (2012). Phylogenetic analysis and classification of the fungal bHLH domain. *Molecular Biology and Evolution, 29*(5): 1301–1318. https://doi.org/10.1093/molbev/msr288

Sayılı, M., Karataş, M., Yücer, A. & Akça, H. (1999). Tokat ilinde alabalık yetiştiriciliği yapan işletmelerin yapısal ve ekonomik analizi. *Ekin Dergisi, 7*: 66-72.

Steingrimsson, E., Copeland, N. G. & Jenkins, N. A. (2004). Melanocytes and the microphthalmia transcription factor network. *Annual Review of Genetics, 38*(1): 365–411. https://doi.org/10.1146/annurev.genet.38.072902.09

Suyama, M., Torrents, D. & Bork, P. (2006). PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research, 34*(Web Server): W609-W612. https://doi.org/10.1093/nar/gkl315

Tapscott, S. J. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development, 132*(12): 2685–2695. https://doi.org/10.1242/dev.01874

Voorrips, R. E. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity, 93*(1): 77–78. https://doi.org/10.1093/jhered/93.1.77

Voytik, S. L., Przyborski, M., Badylak, S. F. & Konieczny, S. F. (1993). Differential expression of muscle regulatory factor genes in normal and denervated adult rat hindlimb muscles. *Developmental Dynamics, 198*(3): 214-224. https://doi.org/10.1002/aja.1001980307

Walters, E. H., Stickland, N. C. & Loughna, P. T. (2000). The expression of the myogenic regulatory factors in denervated and normal muscles of different phenotypes. *Journal of Muscle Research*, *Cell Motility, 21*(7): 647-653. https://doi.org/10.1023/A:1005683825960

Wang, X. H., Wang, Y., Liu, A. K., Liu, X. T., Zhou, Y., Yao, Q. & Chen, K. P. (2015). Genome-wide identification and analysis of basic helix-loop-helix domains in dog, *Canis lupus familiaris. Molecular Genetics and Genomics*, *290*(2): 633-648. https://doi.org/10.1007/s00438-014-0950-1

Wang, Y., Chen, K., Yao, Q., Wang, W. & Zhu, Z. (2007). The basic helix-loop-helix transcription factor family in *Bombyx mori. Development Genes and Evolution, 217*(10): 715-723. https://doi.org/10.1007/s00427-007-0184-x

Wang, Y., Chen, K., Yao, Q., Zheng, X. & Yang, Z. (2009). Phylogenetic analysis of zebrafish basic helix-loop-helix transcription factors. *Journal of Molecular Evolution, 68*(6): 629-640. https://doi.org/10.1007/s00239-009-9232-7

Wei, K. & Chen, H. (2018). Comparative functional genomics analysis of bHLH gene family in rice, maize and wheat. *BMC Plant Biology, 18*(1): 309. https://doi.org/10.1186/s12870-018-1529-5

Yasunami, M., Suzuki, K., Maruyama, H., Kawakami, H., Nagai, Y., Hagiwara, M. & Ohkubo, H. (1996). Molecular Cloning and Characterization of a cDNA Encoding a Novel Basic Helix- Loop- Helix Protein Structurally Related to Neuro-D/BHF1. *Biochemical and Biophysical Research Communications, 220*(3): 754-758. https://doi.org/10.1006/bbrc.1996.0476

Yiğit, M. & Aral, O. (1999). Gökkuşağı alabalığı'nın (*Oncorhynchus mykiss* W., 1792) tatlısu ve denizsuyundaki büyüme farklılıklarının karşılaştırılması. *Turkish Journal of Veterinary and Animal Sciences, 23*: 53-59.