Caspian Journal of Environmental Sciences
Online ISSN: 1735-3866
Print ISSN: 1735-3033

Effect of consuming blood-substitute feed on the transcriptome of the medicinal leech, *Hirudo orientalis* Utevsky and Trontelj, 2005 (Clitellata: Hirudinida)

Abbas Torabian[1], Majidreza Khoshkholgh[1*], Sajad Nazari[2]

1. Fisheries Department, Faculty of Natural Resources, University of Guilan, P.O. Box 1144, Sowmeh Sara, Iran
2. Shahid Motahary Cold-Water Fishes Genetic and Breeding Research Centre, Iranian Fisheries Sciences Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Yasouj, Iran

** Corresponding author's E-mail: majidreza@guilan.ac.ir*

ABSTRACT

While leeches in the genus *Hirudo* have long been models for neurobiology, the molecular underpinnings of nervous system structure and function in this group remain largely unknown. We sequenced the transcriptome of the primary salivary glands of *Hirudo orientalis* Utevsky and Trontelj, 2005 using short-read sequencing (Illumina) technology. A de novo-assembled transcriptome revealed 28,485 high-quality contigs with an average size of 598 bp. A total of 10,542 contigs had significant BLASTx or tBLASTx hits ($E \leq 1.0E^{-6}$) to known proteins, whereas a high percentage (59.36%) of contigs exhibited no apparent protein or nucleotide hits. Comparison of the *Hirudo orientalis* salivary gland transcriptome against a published full-body transcriptome assembled from Roche-454 reads revealed several contigs with putative annotations associated with salivary gland functions. KEGG pathway analysis revealed that the majority (15 out of the top 20 predicted KEGG pathways) of the salivary gland contig sequences match proteins involved in metabolism. We identified several genes likely to be involved in detoxification and inhibition of plant defense responses including aldehyde dehydrogenase, metalloprotease, glucose oxidase, glucose dehydrogenase, and regucalcin. We also identified several genes that may play a role in the extra-oral digestion of plant structural tissues including b-glucosidase and pectin lyase, in addition to the extra-oral digestion of sugars, including a-amylase, maltase, sucrase, and a-glucosidase. Our results provide information on the genes related to head formation and insights into the function of proboscis-related genes during organogenesis with the potential roles of genes not yet characterized.

Article type: Research Article.

INTRODUCTION

Approximately 800 species of leeches have been identified, of which 100 are marine, 90 are terrestrial and the rest are freshwater (Tong *et al.* 2020). Medicinal leeches are worms belonging to the phylum Annelida and subclass Hirudinea, which are sensitive to vibrations of water, touch, light, heat, sound, and various chemicals (Sig *et al.* 2017). They are sanguivorous and depending on nutritional conditions, available prey and growth stage feed on vertebrate blood or invertebrate hemolymph (Whitaker *et al.* 2004; Liu *et al.* 2019; Tong *et al.* 2020). Ten species of leeches or their secretions are used for medical and therapeutic purposes (Petrauskiene 2008). *Hirudo orientalis* Utevsky & Trontelj 2005, is one of the most important and well-known medical leeches in the world playing an important role and share in the industry of breeding as well as utilizing as medical leeches. Grosser & Pešić (2006) described the medicinal leech *H. orientalis* found in the northern part of Iran. According to the literature, *H. orientalis* was found in Georgia, Armenia, Azerbaijan, Iran, and Uzbekistan (Lukin 1976). There is little information available on the population genetics and phylogeographical study of *H. orientalis*. Although the genome is fixed in all cells of a living organism (except mutants), the pattern of gene expression changes under different conditions (growth stage, physiological conditions, and external environmental stimuli) (Chandhini and

Rejish Kumar 2019). Transcript analysis is an effective technology for accurately investigating the relationship between genotype and phenotype as well as identifying mechanisms for controlling cellular fate, development, stress, and disease progression (Inbakandan 2020). In addition, in the absence of a comprehensive genome sequence, it is possible to identify the molecular markers associated with the trait by transcriptomics (Chandhini & Rejish Kumar 2019). In the recent past, transcription profiles have been widely used in aquaculture to identify and analyze the expression of genes involved in growth, reproduction, development, immunity, disease, stress, and toxicology (Nazari *et al.* 2021). The use of microarrays in biological studies of aquatic organisms has its technical limitations. However, NGS mapping technologies and genome sequencing algorithms provide the genome reference of aquaculture species more simply and efficiently. In addition, although it is expensive and time-consuming, but still an essential tool for creating sustainable and profitable aquaculture (Yue & Wang 2017; Chandhini & Rejish Kumar 2019). Omics-based tools and technologies are a recent development in the strengthening of marine biological resources used for aquaculture and fisheries development (Inbakandan 2020) and have helped scientists meet many of the challenges in aquaculture (Santos *et al.* 2014; Martinez - Porchas & Vargas - Albores 2017). Many recent studies in aquaculture reflect omics techniques. However, there is a lack of comprehensive research on the use of transcriptomics in aquaculture, and existing studies focus mainly on safety, disease, and nutrition (Martin *et al.* 2016; Martin & Król 2017; Sudhagar *et al.* 2018; Ye *et al.* 2018). Transcripts of leeches that had already been sequenced were compared using different methods to confirm the predicted molecules or mechanisms (Hibsh *et al.* 2015; Liu *et al.* 2018; Liu *et al.* 2018; Liu *et al.* 2019). In addition, by applying the techniques and methods of analysis and description of salivary gland secretion profiles from leeches, the success of showing multiple proteins (more than 100) has been achieved (Baskova *et al.* 2004; Baskova *et al.* 2008). Previous proteomic analysis, however, showed minimal information about proteins or sequences and should be combined with transcriptomic or genomic analysis. Despite the great diversity of leeches and their close and long-standing relationship with humans, we still have little information about the leech genome. Rapid advances in genomic technologies, especially those associated with high-performance sequencing technologies, have transformed our understanding of gene expression, gene regulation, and molecular mechanisms (Liu *et al.* 2018; Lu *et al.* 2018). Medical leeches are highly regarded by pharmaceutical companies because of the important bioactive substances they contain. A comparative study of these existing leech genomes facilitates the discovery of biologically active substances (Tong *et al.* 2020). A look at recent transcription studies in aquaculture shows that in 2013, by analyzing the transcription sequence of crucian carp, *Carassius auratus* enzymes involved in the glycolysis/gluconeogenesis pathway and microsatellite markers used in selective breeding of this fish were identified (Liao *et al.* 2013). Transcriptional analysis of farmed rainbow trout (Köbis *et al.* 2013) identified genes associated with glucose homeostasis and cholesterol biosynthesis associated with rearing density. In 2014, the molecular mechanism of body color and the pigment genes responsible for the skin color of Oujiang common carp, *Cyprinus carpio* var. *color* were discovered (Wang *et al.* 2014). During 2015, Atlantic salmon, *Salmo salar* microarray analysis was performed to identify the nutrigenomics needs of fish farming (Xue *et al.* 2015). Analysis of zebrafish, *Danio rerio* larvae transcripts revealed differently expressed genes associated with higher protein meals (Rurangwa *et al.* 2015). During 2016, liver transcripts of *Hypophthalmichthys nobilis* were studied to understand growth rate, as well as regulate glucose levels and lipid metabolism (Fu *et al.* 2016). Feeding of rainbow trout, *Oncorhynchus mykiss* broodstock with marine and plant materials, and its effect on trout alevin were evaluated using microarray transcription (Lazzarotto *et al.* 2016). The genomic basis of osmotic regulation in freshwater palaemonid prawn, *Macrobrachium australiense* was reported by sequencing (Moshtaghi *et al.* 2016). Physiological mechanisms of adaptation to molting and salinity were identified in the freshwater giant shrimp, *Macrobrachium rosenbergii*, and genes with different expressions that maintained homeostasis against salinity stress and metamorphic osmotic regulation (Chakrapani *et al.* 2016). in 2017, transcript analysis of sea bass (*Dicentrarchus labrax*) identified mechanisms involved in the developmental ability and quality of eggs in fish (Zarski *et al.* 2017). Investigation of transcription factors for intrinsic safety and heat stress in infected delta smelt, *Hypomesus transpacificus* infected with *Ichthyophthirius multifiliis* infection identified molecular markers and genes responsible for ich infection to assess fish health (Frank *et al.* 2017). Comparative analysis of transcriptional responses of *Lepeophtheirus salmonis* lice revealed that the lice response is independent of host sensitivity and that host-parasite evolutionary relationships affect contemporary host-parasite interactions (Braden *et al.* 2017). Through RNA-Seq studies, the molecular mechanism of Cyprinid herpesvirus 3 infection was discovered (Neave *et al.* 2017). A study of whole-body mRNA changes in European sea bass, *Dicentrarchus*

labrax larvae helped to understand the mechanisms of probiotic interaction with the host and its beneficial effects (Schaeck *et al.* 2017). Microarray analysis of reproductive maturity and transcriptional profiles of testes and vas deferens in male black tiger shrimp, *Penaeus monodon* showed that polychaete feeding may increase spermatogenesis by increasing spermatogonia proliferation in testes and by regulating mature spermatozoa in vas deferens. (Leelatanawit *et al.* 2017). To our knowledge, there are still few studies on global gene expression profiles of leech species. Although transcriptome or transcripts of several leeches were sequenced and analyzed by different methods as follows: proteomic analysis of *Hirudo medicinalis* salivary gland secretion showed season-dependent changes in salivary proteins (Baskova *et al.* 2007). The resulting sequences for *Hirudo* transcripts provided the first major database of genes expressed in this important model system. (Macagno *et al.* 2010). In the study on the saliva transcripts of North American medical leeches, *Macrobdella decora*, transcripts of bioactive salivary polypeptides were identified and a better understanding of the evolutionary origin of blood-eating leeches was provided (Min *et al.* 2010). A freshwater leech from the family Glossiphoniidae and a species of annelid called *Helobdella robusta* was sequenced to study evolution, and its early developments were extensively studied (Simakov *et al.* 2013). *Haemadipsa interrupta* leech pyrosequencing from the Haemadipsidae family was compared with known anticoagulants to discuss the current understanding of the distribution and evolution of anticoagulants in leeches (Kvist *et al.* 2014). A comprehensive catalogue of annotated genes for *Hermodice carunculata* was developed to expand reproductive knowledge and immune response genes in the annelid (Mehr *et al.* 2015). In the study of the evolutionary history of the common ancestor of existing leeches, namely rhynchobdellid and arhynchobdellid, with a comparative transcriptional approach and study on the diversity of anticoagulants in the species of these orders, the concept of the single origin of blood nutrition in ancestral leeches was confirmed (Siddall *et al.* 2016). Examination of the transcriptomes of *Poecilobdella javanica*, *Whitmania pigra*, and *Haemadipsa cavatuses* showed that the genetic mechanisms of leeches' environmental adaptations are well-adapted to different environments (Liu *et al.* 2018). In the transcriptional profile of the salivary glands of *Hirudo nipponia* (Lu *et al.* 2018), over 21 genes were identified that may play an important role in the treatment of various diseases. In reviewing the current status and transcription applications in aquaculture, the application of transcriptomics in the production of genetic resources, gene expression, analysis of molecular pathways and the development of molecular markers in various fields of aquaculture are explained (Chandhini & Rejish Kumar 2019). In the study on *Limnobdella mexicana*, a positive relationship was observed between anticoagulant diversity and host diversity in blood-eating leeches (Iwama *et al.* 2019). Proteomic and transcriptomic studies on the leech salivary glands (Liu *et al.* 2019) showed that two-thirds of biologically active genes play a key role in the process of leech bites. RNA-seq was performed on salivary cells of three species: *H. medicinalis*, *H. orientalis* and *Hirudo verbana*, and new salivary proteins, as well as new homologues of anticoagulant coding genes were identified (Babenko *et al.* 2020), hence it turned out that the expression of genes encoding known anticoagulants is not limited to salivary cells. The genome of the leech *Whitania pigra* was prepared and compared with two leeches, *H. medicinalis*, and *Helobdella robusta*, in a genomic comparison study for the exploration of bioactive substances, hence a new hirudin gene g17108 (hirudin_2) was discovered. The genome of *H. medicinalis* from the important family Hirudinidae was examined (Babenko *et al.* 2020; Kvist *et al.* 2020). In order to evaluate the ability of annelid to regenerate, experimental and descriptive data were reviewed at different structural levels to understand the basal cell and molecular diversity of regeneration in annelid and also the role of stem / dedifferentiated cells, as well as molecular morphallaxis in annelid regeneration (Kostyuchenko & Kozin 2021). The problem with using new species as model specimens is the lack of knowledge about how they reproduce and, most importantly, their nutrition in the laboratory. Species that are so few in nature and cannot be continuously harvested from the environment should be bred in the laboratory for long-term studies. Extensive use of medical leeches in various scientific and applied fields has reduced its populations in natural reservoirs and has put it in danger of extinction in many countries. Preparing fresh and healthy blood is difficult or impossible in many areas, and the maintenance and consumption of fresh blood also has its problems and limitations. Therefore, for the protection of leeches in ecosystems and also due to the observance of medical standards, only leeches bred in the laboratory and controlled conditions are used for medical purposes (Utevskaya & Atramentova 2002). So, it should be possible to feed them with blood-substitute feed. Another problem in leech feeding is the lack of uniform feeding behavior among leeches, which necessitates frequent feeding for different proportions of the herd to meet their nutritional needs at different times (Dickinson & Lent 1984; Ceylan & Erbatur 2012). Blood substitute feed can be available at the right time and place and in the required amount,

eliminating the problems of consuming fresh blood. Based on the available information, there has been no report on the production and consumption of blood-substitute feed in scientific articles, and for the first time in this study, by examining the transcript of leeches fed with blood-substitute feed, the effectiveness of this feed instead of fresh blood should be investigated. The results of this study in the field of designing and producing blood-substitute feed for feeding leeches can be provided to leech breeding or research centers that use leeches as model animals.

MATERIALS AND METHODS

Time, place, and schedule of study

The present study was carried out in the Aquatic Breeding Centre of Gil Berkeh Company, Rasht, Iran during 2020-2021 in two consecutive phases. In the first phase, adult leeches were collected from their natural habitat and reproduced. In the second phase, which started with 200 offspring produced in the first phase, under two treatments, calf blood feeding and blood substitute feeding were reared for ten months and the effect of blood substitute feed on leech transcript was investigated.

Preparation and reproduction of adult leeches

Adult leeches (100 pieces) were obtained from local leech fishermen in the natural habitat of the above-mentioned species in Guilan Province and underwent an adaptation period for one month, then reproduced according to the common methods used by researchers (Wilkin & Scofield 1991; Davies & McLoughlin 1996; Utevskaya 1998; Sağlam 2011). Leeches were induced to reproduce by feeding (Ceylan *et al.* 2021). After one month in the mating tank, mated leeches were identified from the clitellum (Wilkin 1989) and kept individually in PET containers with moist sphagnum moss. The produced cocoons were transferred to the cocoon storage tank with the same moss and stored at 25 °C for 1 month with sufficient moisture on the moss (by water spray) to prevent the death of embryos inside the cocoon (Ceylan *et al.* 2021).

Leech breeding

The 200 offspring produced in the first phase of the study were reared for 10 months under two feeding groups to investigate the effect of feed on leech transcriptome. In the first group, fresh blood of calves prepared from Rasht slaughterhouse, and in the second group, blood substitute feed (Table 1) was used for feeding.

Table 1. Blood substitute feed Ingredients.

Nutrient	Amount (g)	Nutrient	Amount (g)
Water	863	Citric acid	0.02
Casein	115	Salt	8.75
vitamins	1	Bicarbonate	0.54
Trace elements and minerals	1	sodium chloride	0.16
Glycerol	2.2	Potassium chloride	0.05
lecithin	2.2	Dextrose	0.71
Amino acids	4.2	Other materials	1.17

Species identification

To identify the samples, first, using 1.5 g of phenoxyethanol (99.9%, Pishgaman Shimi Company, Tehran, Iran) in each liter of water for 5 to 15 minutes, the samples were anesthetized, then biometry and identification were performed alive. Species identification was made according to color patterns and body characteristics based on identification keys (Lukin 1976; Elliott & Mann 1979; Sawyer 1986b).

Tissue samples and RNA extraction

To take the sample required for genetic tests, the leeches were not fed for two months so that the feed eaten in the last meal was completely fixed in the body of the leech and it was ready to bite or feed again. Ten leeches were randomly selected from each treatment and transferred to the laboratory of Shahid Motahari Cold Water Fishes Genetics and Breeding Research Center, Yasouj, Iran to carry out the necessary steps for sequencing. The leeches were anesthetized in a 5% phenoxyethanol solution for 10 min. After emptying the contents of the digestive system and opening the body, the leech's salivary glands were removed and prepared for further examinations. Afterward, RNA was extracted and the rest of the sample was stored at -20 °C in five times the volume of RNAlater solution

(Thermo Fisher Scientific, Massachusetts, United States). RNA extraction was performed according to Trizol reagent protocol 15596026 (Thermo Fisher Scientific, Massachusetts, United States).

Library construction and sequencing

In this study, using the Turbo DNA-free Kit (Thermo Fisher Scientific, Massachusetts, United States), the possible DNA of each sample was removed according to the protocol. RNA quantity was measured using Eppendorf BioPhotometer D30 (Eppendorf, Schonenbuch, Switzerland). RNA quality was evaluated based on the values of A260/A280 and A260/A230 ratios. Values of 1.9-2 were considered acceptable. In this study, RNA integrity was measured using RNA denaturation gel. The mRNA of the stored samples was purified using Ribominus Eukaryote Kit (Thermo Fisher Scientific, Massachusetts, United States) and MicroPolyA Purist Kit (Thermo Fisher Scientific, Massachusetts, United States) for RNA-Seq analysis. cDNA library was created using SOLiD Whole Transcriptome Analysis Kit (Thermo Fisher Scientific, Massachusetts, United States). Sequencing was done using Illumina high seq 2500 (Applied Biosystems), which is one of the second-generation sequencing methods.

Transcriptome assembly

Analysis of RNA-Seq data of leeches included several steps and was analyzed by several software (Clc Genomic workbench and R packages). Quality control was done using RNA-SeQC software, In addition to control of raw RNA-Seq data, including normalization and correction of various errors and biases caused by technical defects and differences in the frameworks used in generating these data using FastQc software. Assembling the sequence of small RNA fragments of leeches and considering their common and overlapping parts and reconstructing the whole studied transcriptome (Yassour *et al.* 2009) was performed using Cufflinks 1.1 and Cuffmerge 2.1 software packages. Mapping of RNA sequences (total transcriptome) on the reference genome was done using TopHat software. This analysis helped to identify the real and correct binding sites from the alternative and similar ones. Thus, different and similar forms were obtained for the sequence of a particular transcript. Annotation was carried out on each transcript by recording detailed information on each short RNA fragment, including its sequence and genomic position. Alignment analysis of annotated short transcripts sequence with the sequence of haplotypes and/or genomic fragments carrying annotated SNPs was performed using Bowtie software. After formulating each specific strategy and procedure, it was tested on simulated and real data [from valid databases such as the data set in the SRA part of the NCBI database (address: http://www. ncbi.nlm.nih.gov/sra)] and its efficiency, accuracy, and precision were compared with each other and with common genomic methods. Index parameters in RNA-Seq data, i.e., nucleotide sequence and number of readings, were simulated and used for statistical inference. First, the gene expression data were grouped into related categories using the grouping tool and according to the type of treatment, then using the GEO2R tool on the GEO website (ncbi.nlm.nih.gov/geo/geo2r). Then, given the significance level of 0.05 and Benjamini and Hochberg's method to correct the p-statistic, the differentially expressed genes (DEGs) were identified. DEGs were classified into two groups, up-regulated-, and down-regulated- genes, based on the logarithm of the range of changes greater than one (logfc>1) and a significance level of 0.05. Heat map of DEGs was drawn using clustVis (https://biit.cs.ut.ee/clustvis; Metsalu & Vilo 2015). After mapping, the length of the gene and the number of "reads" were corrected. Two of the most important corrections in this field are "RPKM"[1] and "FPKM"[2]. To identify genes with different expressions in each experiment, the samples merged in the previous step were analyzed with statistical tests. After identifying and integrating the DEGs, the DAVID[3] database tool (https:/ /david.ncifcrf.gov) and threshold level of 0.01 were used (Ashburner *et al.* 2000; Huang *et al.* 2009) to check the structural and functional characteristics, as well as identifying the metabolic pathways of the DEGs and checking the ontology and defining the related pathways. The ontology was used in the three paths of biological processes, cellular components, and molecular function for subsequent analyses. KEGG was also used to identify pathways related to DEGs. Classification of hub genes based on 3 indicators of molecular function, biological pathways, and cellular components was done by Panther online software. To identify causal genes, Cytohubba plugin was used in Cytoscape software (Chin *et al.* 2014).

Bioinformatics analysis

Quality control and pre-processing of readings were performed using RNA-SeQC, FastQC, and Trimmomatics software. The positioning of RNA sequences on the reference genome was carried out using TopHat2 software.

[1] - Read Per Kilo base of Transcript per Million Mapped Read
[2] - Fragment Per Kilo base of Transcript per Million Mapped Read
[3] - Database for Annotation, Visualization and Integrated Discovery

Then, using Cufflinks and Cuffmerge software, transcripts were assembled for each sample. Cuffdiff software was used to investigate the difference in gene expression between two feed treatments.

Principles of ethics

The present study was carried out under the conditions and regulations of the code of ethics for working with laboratory animals and following the guidelines for the care and use of laboratory animals in scientific affairs.

RESULTS

Quality and quantity of RNA samples

In this study, the quality of the extracted RNA was evaluated by two qualitative and quantitative methods. A horizontal electrophoresis device and 1.5% agarose gel were used for qualitative evaluation, an example of which is shown in Fig. 1. In addition, for quantitative evaluation, the nanodrop device was used to determine the best concentration of RNA for sequencing (Fig. 2).

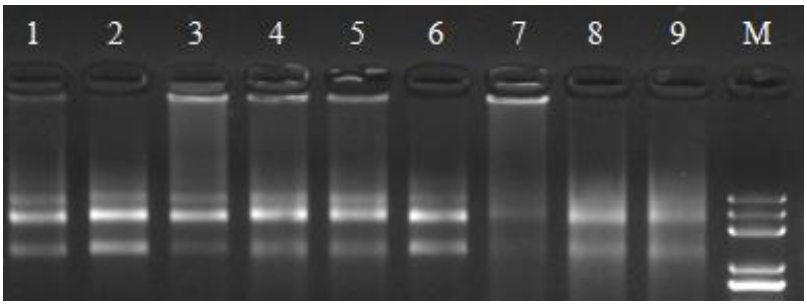
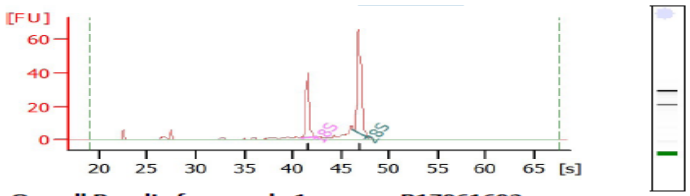


Fig. 1. Quality of RNA extracted in leeches using agarose gel (M = 100 bp gel quality indicator).



Overall Results for sample 1 : R17061602

RNA Area:	177.1
RNA Concentration:	184 ng/µl
rRNA Ratio [28s / 18s]:	1.9
RNA Integrity Number (RIN):	9.4 (B.02.08)
Result Flagging Color:	
Result Flagging Label:	RIN: 9.40

Fragment table for sample 1 : R17061602

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	40.83	42.55	34.2	19.3
28S	46.32	47.94	64.9	36.7

Fig. 2. Quantity of RNA extracted using nanodrop spectrophotometer - sample number 1.

Expression of genes

A typical sequencing process involves fragmenting the genome into millions of size-selected fragments that are ligated to adapters, and each sequenced fragment is considered a read. In this study, the readings obtained are provided according to Table 2. Using the RNA sequencing data obtained from leeches, the genes with different expressions and the most important genes affecting growth were examined. Leech salivary gland tissue samples were classified into two groups of leeches treated with formulated feed and control feeding. Then, based on the *p*-value and the values of the range of changes, the genes were divided into two groups of up-regulated- and down-regulated- genes. According to the results in the *volcano plot*, which consists of two axes, *p*-value and the range of changes, it is possible to identify the DEGs. These genes were separated into up-regulated and down-regulated genes. Another method used to identify DEGs is to check the range of changes and the level of gene expression in the samples. The average difference graph can show this distribution and it is possible to identify the DEGs by considering the statistics of the range of changes and gene expression of the samples. The volcano plot of DEGs of *H. orientalis* transcriptome test is shown in Fig. 3.

Ontology and gene set enrichment

In the category of biological processes, most of the genes with low expression were active in metabolic processes, cellular processes, and biological regulation. In the category of cellular components, most of these genes were in

the intracellular space and part of the anatomical identity of the cell. In addition, in the category of molecular function, binding and enzyme activities were at their highest level. The enrichment of genes with high expression in three categories of biological pathways, molecular function, and cellular components are shown in Fig. 4.

Table 2. Transcriptomic libraries of salivary glands of filtered and mapped reads in *H. orientalis*.

Samples	Raw reads	Clean reads	Total mapped	Multiple mapped	Uniquely mapped
Ho_1	50048582	39805030	24010872 (52.23%)	1384333 (2.98%)	23523338 (39.23%)
Ho_2	55093534	53795542	29988348 (53.73%)	1813980 (3.31%)	28173388 (51.32%)
Ho_3	50737228	50305080	25219332 (50.03%)	1398379 (2.97%)	23720943 (37.04%)
Hc_1	53395893	52933154	29022095 (53.83%)	2080005 (3.93%)	24932090 (50.9%)
Hc_2	50355544	50039020	27310117 (53.77%)	1498835 (3.39%)	25711272 (51.37%)
Hc_3	50144883	39485592	27304315 (53.94%)	2172001 (3.37%)	25133313 (50.59%)

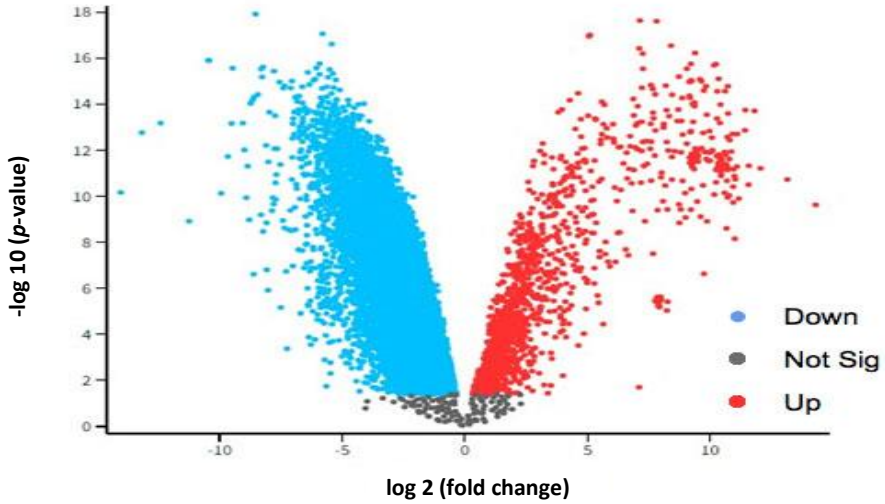


Fig. 3. Volcano plot of the DEGs in *H. orientalis*. Up-regulated genes (Up), down-regulated genes (Down), and genes that did not change in a significant level of expression (Not significant).

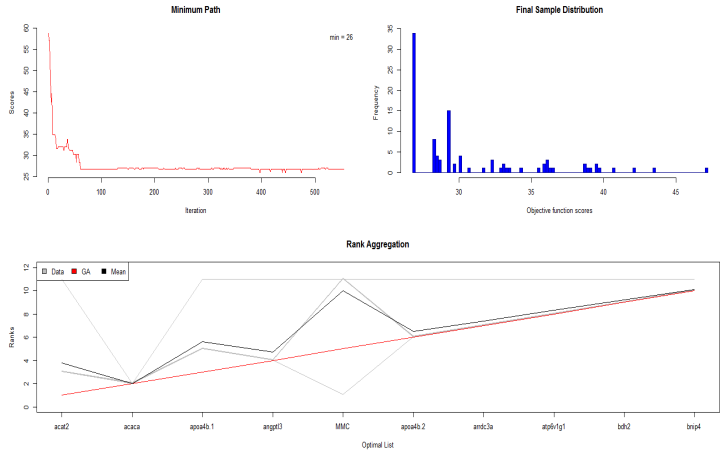


Fig. 4. Identification of causal genes resulting from the network of DEGs in *H. orientalis*.

According to the results of the ranking method (genetic algorithm and cross-entropy Monte-Carlo) on the causal genes from the network of DEGs, genes *acat2*, *acaca*, *apoa4b.1*, *angptl3*, *apoa4b.2*, *arrdc3a*, *atp6v1g1*, *bdh*, *bnip4*, *gcshb* were identified as the final effective causal genes. Among the effective regulatory factors on hub genes, categories such as miRNAs, transcription factors, as well as long and small non-coding RNAs can be

mentioned. By searching and exploring the databases of transcription factors such as animalTFBD, Ggtrd, and Contra V3, no information related to transcription factors related to these hub genes was obtained.

Paths of Gene Ontology

From the collection of 132,884 sequences related to different ontology paths obtained from this study, in the biological process category, gene ontology paths that were significantly enriched were identified, and the number of genes involved in each GO category was determined. The characteristics of the resulting gene ontology paths are given in Table 3.

Table 3. Statistically enriched Gene Ontology terms in the "Biological Process" category.

GO ID	SGGenes	*p*-value	GO Ontology or Category
0040875	47	8.00E-14	Response to stimulus
0007154	32	7.23E-04	Signal transduction
0041177	144	8.00E-14	Localization
0008104	28	3.13E-05	Protein localization
0034513	18	2.77E-04	Cellular protein localization
0041541	25	1.40E-07	Cellular localization
0041234	147	3.04E-07	Establishment of localization
0044184	28	4.48E-04	Establishment of protein localization
0041547	25	8.74E-07	Establishment of localization in cell
0005810	147	2.47E-05	Transport
0014031	28	3.78E-04	Protein transport
0045707	20	3.40E-04	Intracellular transport
0005885	18	4.71E-04	Intracellular protein transport
0005504	4	4.27E-03	Protein targeting
0044084	27	2.33E-11	Cellular component biogenesis
0034522	13	2.07E-04	Cellular macromolecular complex assembly
0043523	8	2.45E-03	Cellular protein complex assembly
0008142			Metabolic process
0004774	47	4.47E-05	Carbohydrate metabolic process
0005055	24	4.47E-05	Alcohol metabolic process
0045154	14	2.07E-04	Alcohol catabolic process
0007045	48	4.01E-11	Catabolic process
0007047	17	1.01E-07	Macromolecule catabolic process
0015042	15	2.53E-05	Carbohydrate catabolic process
0044274	14	1.31E-03	Cellular carbohydrate catabolic process
0044237	388	4.28E-07	Cellular metabolic process
0005137	174	1.87E-07	Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process
0044250	244	2.24E-05	Cellular macromolecule metabolic process
0005247	37	1.23E-03	DNA metabolic process
0005281	14	2.48E-03	DNA repair
0015070	75	3.27E-04	RNA metabolic process
0032774	44	8.12E-05	RNA biosynthetic process
0005375	23	1.08E-03	RNA processing
0034550	27	2.43E-03	ncRNA metabolic process
0005341	43	3.27E-04	Transcription, DNA-dependent
0044248	41	1.24E-08	Cellular catabolic process
0041185	30	3.85E-08	Cofactor metabolic process

0005732 21	8.12E-05	Coenzyme metabolic process
0005730 10	1.08E-05	One-carbon metabolic process
0044252 31	1.42E-05	Cellular carbohydrate metabolic process
0004775 22	3.48E-04	Monosaccharide metabolic process
0017318 21	3.51E-04	Hexose metabolic process
0005075 12	2.07E-04	Glycolysis
0005770 7	2.32E-05	Sulfur metabolic process
0005447 17	1.23E-03	Protein folding
0054007		Biological regulation
0040787 75	8.43E-11	Regulation of biological process
0007887 44	2.27E-04	Regulation of biosynthetic process
0040774 74	1.01E-07	Regulation of cellular process
0031323 48	2.01E-04	Regulation of cellular metabolic process
0031325 44	4.44E-04	Regulation of cellular biosynthetic process
0017217 44	4.20E-04	Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process
0017222 50	7.40E-07	Regulation of metabolic process
0050244 47	2.01E-04	Regulation of macromolecule metabolic process
0010445 44	4.44E-04	Regulation of macromolecule biosynthetic process
0041242 42	2.47E-03	Regulation of RNA metabolic process
0010458 47	4.54E-04	Regulation of gene expression
0005344 42	4.84E-03	Regulation of transcription, DNA-dependent
0042472 12	4.47E-05	Homeostatic process
0017724 12	3.25E-07	Cellular homeostasis

The KEGG pathways that were significantly enriched were identified based on the characteristics and commonalities of the changed expressed genes, and the highest and lowest probability value changes (significance level) in these biological pathways were determined. Of the 33 identified pathways, the salivary secretion pathway exhibited the highest probability, while the fat digestion and absorption pathway was the lowest. Fig. 5 illustrates the most important biological pathways with significant probability value changes in this study. From the data of this study, a total of 4771 genes were identified that had changed expression, of which 4342 genes were decreased in expression, while 425 genes were increased (shown in Fig. 6). The genes that showed different expressions were detected in the KEGG database and the metabolic pathways involved in them were identified. Several pathways were composed of a larger amount of DEGs, and some metabolic pathways also had a lower number of DEGs (overviewed in Fig. 7). Examining 12814 transcripts obtained in the Pfam protein families database (including annotations and multiple sequence alignments) led to the identification of 3517 Pfam domains. A sample of Pfam data is given in Table 4.

Table 4. Sample transcripts of Pfam protein family data in this research.

No.	TPB_SG1 Transcript ID	PfamID	Pfam Domain Description
1	c2_g1_i1	PF00643	B-box zinc finger
2	c10_g1_i1	PF14227	gag-polypeptide of LTR copia-type
3	c20_g1_i1	PF00118	TCP-1/cpn60 chaperonin family
4	c31_g1_i1	PF07728	AAA domain (dynein-related subfamily)
5	c49_g1_i1	PF01172	Shwachman-Bodian-Diamond syndrome (SBDS) protein
6	c58_g1_i1	PF00076	RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)
7	c63_g1_i1	PF02198	Sterile alpha motif (SAM)/Pointed domain
8	c74_g1_i1	PF01757	Acyltransferase family

9	c82_g1_i1	PF00086	Thyroglobulin type-1 repeat
10	c85_g1_i1	PF00078	Reverse transcriptase (RNA-dependent DNA polymerase)
11	c92_g1_i1	PF05033	Pre-SET motif
12	c107_g1_i1	PF00078	Reverse transcriptase (RNA-dependent DNA polymerase)
13	c129_g1_i1	PF02204	Vacuolar sorting protein 9 (VPS9) domain
14	c131_g1_i1	PF03911	Sec61beta family
15	c133_g1_i1	PF06703	Microsomal signal peptidase 25 kDa subunit (SPC25)
16	c135_g1_i1	PF13465	Zinc-finger double domain
17	c137_g1_i1	PF00400	WD domain, G-beta repeat
18	c139_g1_i1	PF05585	Putative peptidase (DUF1758)
19	c159_g1_i1	PF07841	DM4/DM12 family

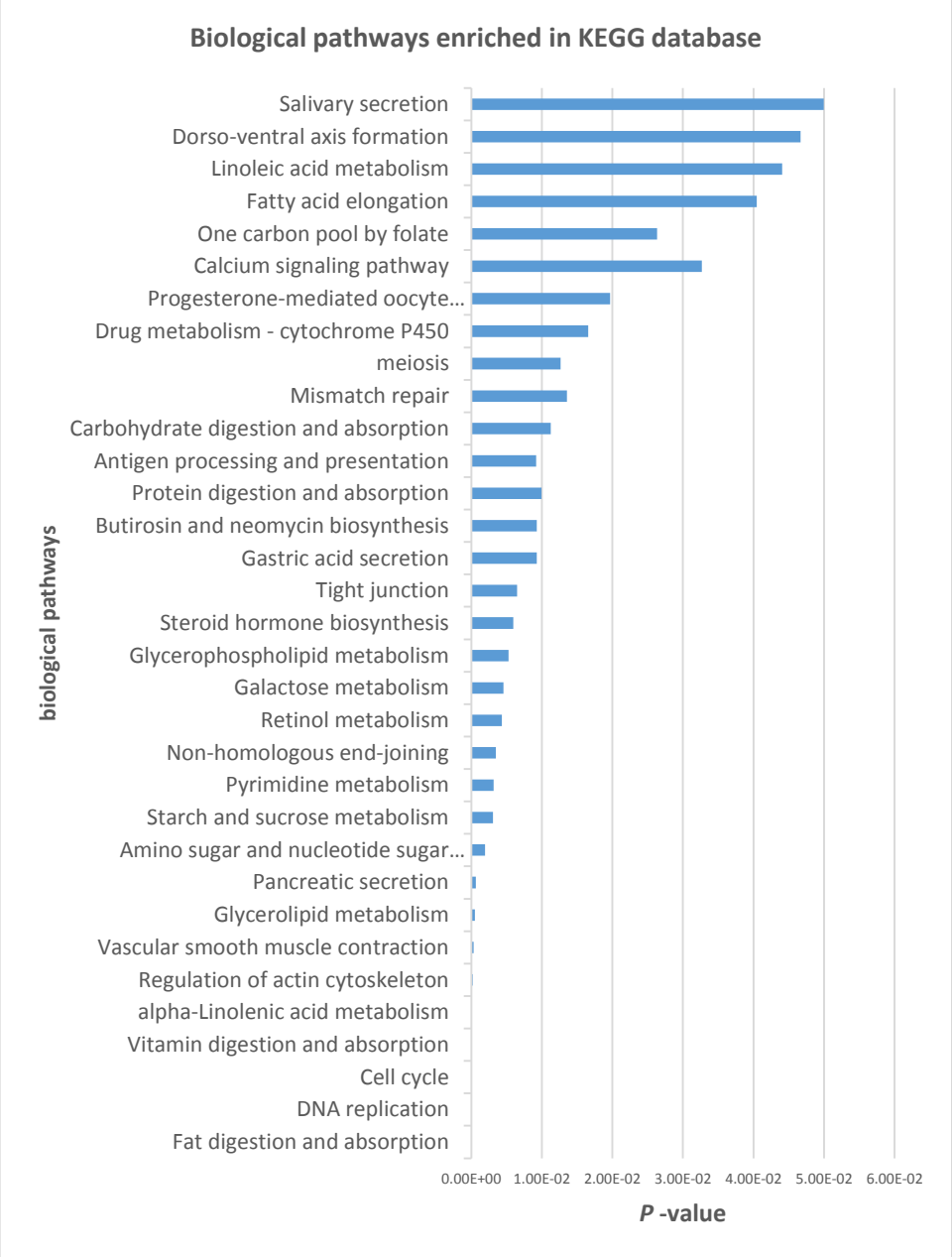


Fig. 5. Significantly enriched KEGG pathways.

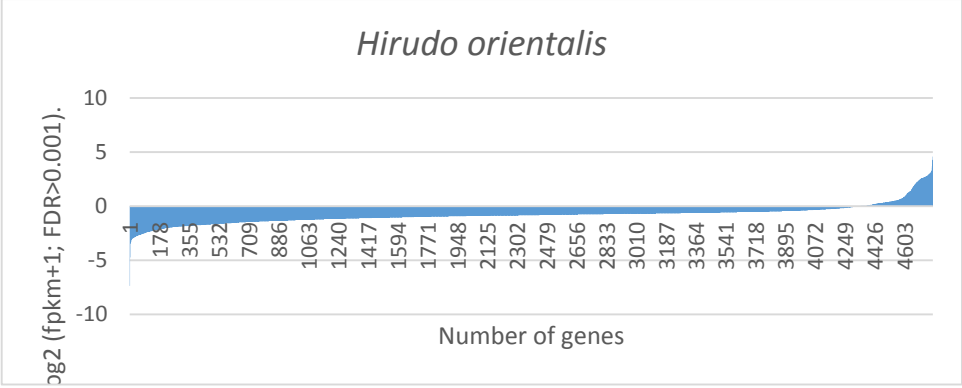


Fig. 6. Differentially expressed genes in *Hirudo orientalis*.

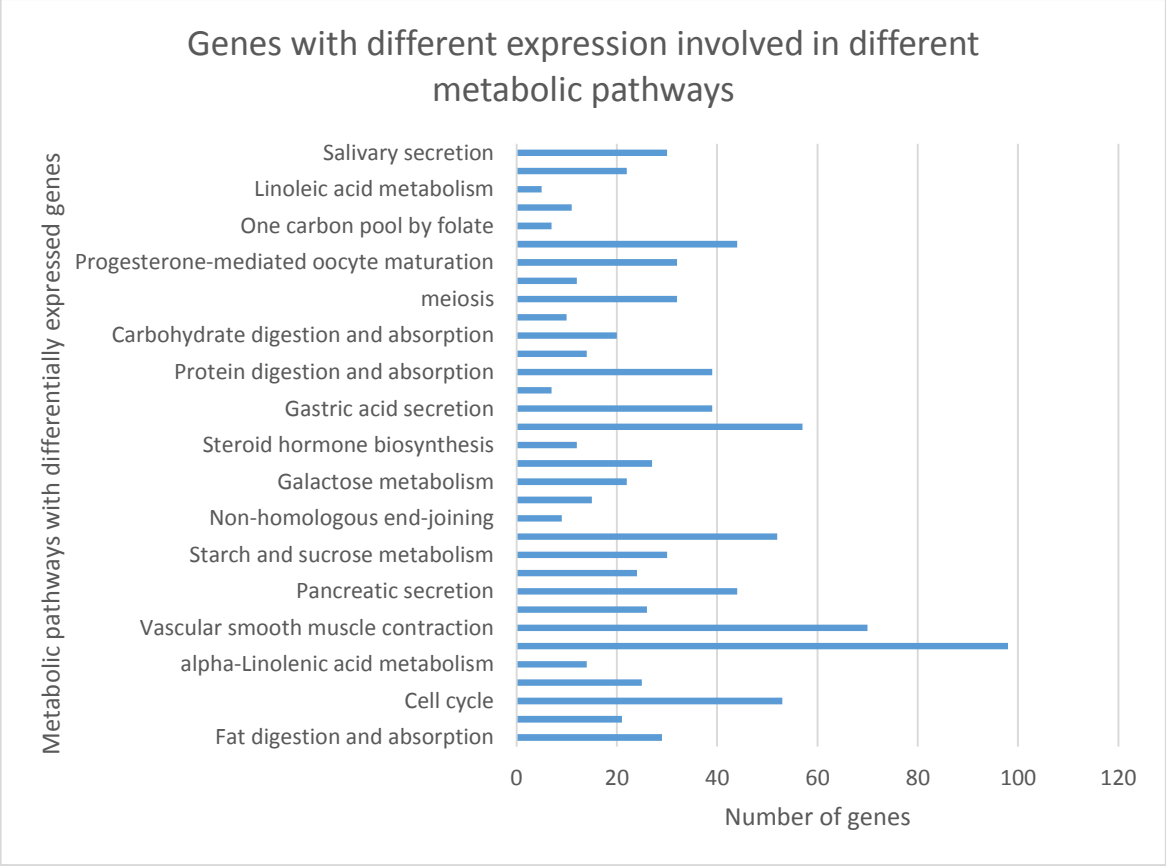


Fig. 7. The number of differentially expressed genes participating in significantly enriched KEGG pathways.

DISCUSSION

In this study, two feed treatments (blood and casein blood substitute feed) on *H. orientalis* changed the expression of 4771 genes in 33 biological pathways, so that the leeches could adapt to the change in diet. The environment affects the activity and expression of genes, and hereditary phenotypic changes (without changes in the DNA sequence) are studied in the field of epigenetics. It has been shown that feed is a very influential environmental factor in the role and function of genes in the living body. Epigenetic changes determine whether genes are turned on, off, or altered expression and affect the production of specific proteins or metabolites in certain cells, or ensure that only necessary proteins are produced. The patterns of epigenetic changes differ among individuals, in different tissues in individuals, and even in different cells (Douglas 2006; Panserat & Kaushik 2010; Ngoh *et al.* 2015; Bizuayehu *et al.* 2016; Hixson *et al.* 2017; Morris *et al.* 2018). It was shown that in the biological activities of leeches, the indices such as growth rate, survival rate, and reproduction performance are affected by nutrition (Sineva 1944; Galun 1966; Dickinson & Lent 1984; Elliott & Tullett 1986; Davies & Mcloughlin 1996; Li *et al.* 2001; Zhang *et al.* 2008; Zulhisyam *et al.* 2011; Ceylan *et al.* 2019).

During the genetic researches of recent years, many metabolic pathways responsible for the life of living organisms have been revealed. However, there is still a lot of work to be done to fully understand the processes and mechanisms involved and coordinating the activity of genes during the growth and development of organisms. Because many regulatory mechanisms related to metabolism, hormonal settings, and the role and function of genes work in harmony with each other (Belmonte *et al.* 2013; Tong *et al.* 2020). The lack of genomic information about medical leeches has caused our knowledge about this organism to be at a very low level. In this study, the transcriptome of eastern medicinal leech was sequenced to identify and functionally investigate transcripts with different expressions due to the consumption of casein blood substitute feed. The identified sequences between the two feed treatments were compared with each other; the transcripts with different expression were functionally interpreted and 132884 sequences were identified. Sequence similarity analyses and functional interpretation of these sequences were performed. GO analysis identified 8813 functional annotations in 4 GO groups. The most abundant GO found was in the biological process category.

Gene expression

The numbers of identified leech genes in *H. robusta*, *W. pigra*, and *H. medicinalis* species are reported to be 23400, 26743, and 14596 respectively (Tong *et al.* 2020). The expression of different genes in different living tissues fluctuates in an unknown and controlled way, and the peak of complexity appears when we know the presence and activity of genes, in addition to the asymmetric spatial distribution in the tissue. Different types have different time expressions in different periods of growth and development (Belmonte *et al.* 2013). The change values of DEGs from Trinity genes [log2(fpkm+1) (FDR>0.001)] showed that 4771 genes changed expression under two nutritional treatments, out of this number, 4342 genes were decreased in expression. Also, 425 genes were increased in expression.

Biological pathways

The analysis of the research data in the KEGG database led to the recognition of 33 biological pathways that were significantly enriched. In the examination of gene ontology data, a total of 132,884 sequences were obtained, of which 31,870 were in the genetic ontology group C (Cellular component), 28,705 were in the genetic ontology group F (molecular function), 1,974 were in the genetic ontology group N (None) and 70335 were placed in genetic ontology group P (biological process). Biological processes, having 53% of the sequences obtained in this study, have assigned the largest contribution and role in living biological activities, undergoing the greatest impact from diet change. It seems that paying attention to the processes that have undergone the most changes will help us in understanding the effects and fruits of diet changes in the organism and corrective actions in optimizing the food formulation according to the needs of the organism (Hixson *et al.* 2017; Morris *et al.* 2018). Naturally, biological processes have played the biggest role in the leech's adaptation to the casein blood substitute diet, since it has shown the greatest change to adapt to new conditions. Afterward, the processes related to cellular components and molecular function, by 24 and 21% of the obtained sequences respectively, were able to gain the next positions in the adaptation of the organism to the new biological conditions (dietary change). From the obtained 132884 sequences, a total of 8813 different biological activities were identified, including 977 activities in genetic ontology group C, 2158 activities in genetic ontology group F, 1 activity in genetic ontology group N, and 5677 activities in group of the genetic ontology of P. Cellular processes are controlled by a set of related and influencing molecules, whose activity and levels are often regulated or expressed together. This evidence shows that the methods used in network reconstruction should consider regulatory relationships and other regulatory factors. Reconstruction of inter-gene relationships related to nutrition with different sources using protein regulatory and interaction networks is in line with this attitude. However, network reconstruction provides a list of hub genes and modules for a specific treatment. In general, a set of genes or their products (proteins) that have similar and common biological features and functions constitute the gene set. Gene set enrichment analysis is a method to identify the group of genes or proteins that are expressed more than usual in a set of genes or proteins and may be related to pathogenic phenotypes (Subramanian *et al.* 2005). In comprehensive experiments to identify sets of genes (for example, genes that are expressed differently in different conditions), often, the goal is to identify the biological function of genes to gain a comprehensive understanding of the mechanisms affecting biological processes. For this purpose, the above goal can be achieved by comparing the desired genes to each of the expressions of biological processes, molecular functions, and cellular components in gene ontology. A statistical

test can be used to check the enrichment of input genes for each expression. According to the results, the number of DEGs was high. The criteria considered to identify DEGs (corrected p-statistic value greater than 0.05 and logarithm of change of amplitude greater than one) can influence the number and type of these genes. Among other factors affecting the amount of identified genes, we can mention the percentage of matching of annotation information and mapping of probes to genes. In the case of RNA-seq test data, about 20% of the probes did not have gene annotation information, so the information related to their gene expression in the gene expression matrix was removed. Nevertheless, the number of deletions will be effective in identifying differently expressed genes. Among the highly expressed genes, the gene family *fabp, elovl, fads, ndufa, apoa, cpt, slc*, and *acaca* are active in the category of metabolism, binding, and transport of lipids. The *acaca* gene (acetyl-CoA carboxylase) is a biotin-containing enzyme that performs the carboxylation of acetyl-CoA to malonyl-CoA, which is the rate-limiting step in the synthesis of fatty acids. The *fabp7* gene is involved in the process of binding to fatty acids. The *elovl* gene is related to the production of the enzyme that lengthens the chain of unsaturated fatty acids and plays a role in the elongation of the side chain of fatty acids. The *fad* gene catalyzes the double bond in the fatty acid chain and plays an important and decisive step in the synthesis of polyunsaturated fatty acids. The *apoa4* gene is an apolipoprotein and is involved in cell transfers. The increase in the expression of *apoa* and *fabp* genes shows that, due to the consumption of casein blood substitute feed, the activities related to the transfer of fatty acids and lipids in extracellular and intracellular have increased. The expression of the important *slc* gene in membrane lipid transfer indicates an elevation in lipid transfer activity. The *cpt* gene is carnitine palmitoyl transferase and the *ndufa* gene is a part of the *nadh* dehydrogenase family, whose expression was increased in this study. The genes identified in this study as genes related to the metabolism and growth of leeches, include those that can be used in future studies from the point of view of their function and life cycle using methods such as gene knockdown (with the help of RNA interference), deactivation of some gene regions (gene knockout) and gene transfer for high expression of a specific gene (transgenic overexpression).

CONCLUSION

Different factors affect the expression pattern of genes in the leech's body, and feeding, digestion and absorption of feed as the main biological activity during the life of an organism, play an important role in determining the quantity and quality of gene expression. According to living conditions and the expression of its genome, the transcriptome, as the primary product of gene expression, is always adjusted in the best possible state for biological adaptation and plays a role in controlling and guiding the mechanisms of cell fate. The data of this research showed that various genes changed their expression (increase or decrease) due to consumption of blood substitute feed and the resulting transcriptome was also changed.

REFERENCES

Abdelrahman, H, ElHady, M, Alcivar-Warren, A, Allen, A, Al-Tobasei, R, Bao, L & Beck, B 2017, Aquaculture genomics, genetics and breeding in the United States: current status, challenges, and priorities for future research. *BMC genomics*, 18: 1-23.

Abdullah, S, Dar, A, Rashid, A & Tewari, A 2012, Hirudotherapy/leech therapy: applications and indications in surgery. *Archives of Clinical Experimental Surgery*, 1(3): 172-180.

Allen, P & Steeby, J 2012, Aquaculture: challenges and promise, Nature Education, 2(11): 12.

Babenko, VV, Podgorny, OV & Manuvera, VA 2020, Draft genome sequences of Hirudo medicinalis and salivary transcriptome of three closely related medicinal leeches. *BMC genomics*, 21: 1-16.

Baskova, I, Kostrjukova, ES & Vlasova, MA 2008, Proteins and peptides of the salivary gland secretion of medicinal leeches *Hirudo verbana, H. medicinalis,* and *H. orientalis. Biochemistry* (*Moscow*), 73: 315-320.

Baskova, I, Kostrjukova, ES & Vlasova, MA 2004, Protein profiling of the medicinal leech salivary gland secretion by proteomic analytical methods. *Biochemistry* (*Moscow*), 69: 770-775.

Baskova, I, Zavalova, LL, Kostryukova, ES, Titova, GA, Lazarev, VN, & Zgoda, VG 2007, Proteomic analysis methods for characterization of proteins from the salivary gland secretions of the medicinal leech during different seasons, *Biochemistry* (*Moscow*), 72: 219-225.

Braden, L, M, Braden, LM, Monaghan, SJ, Fast, MD 2017, Enhanced transcriptomic responses in the Pacific salmon louse *Lepeophtheirus salmonis oncorhynchi* to the non-native Atlantic Salmon *Salmo salar* suggests increased parasite fitness. *BMC genomics*, 18: 1-14.

Ceylan, M, Küçükkara, R, Erbatur, I, Karataş, E, Tunç, M & Sağlam, N 2021, Growth, survival and reproduction of the Turkish medicinal leech, *Hirudo sulukii*. *Invertebrate Reproduction & Development*, 65: 57-68.

Chakrapani, V, Coulouris, V, Avagyan, N, Papadopoulos, J, Bealer, K & Madden, TL 2016, Comparative transcriptomic profiling of larvae and post-larvae of *Macrobrachium rosenbergii* in response to metamorphosis and salinity exposure. *Genes & Genomics*, 38: 1061-1076.

Chandhini, S & Rejish Kumar, VJ 2019, Transcriptomics in aquaculture: current status and applications. *Reviews in Aquaculture*, 11: 1379-1397.

Davies, R & McLoughlin, N 1996, The effects of feeding regime on the growth and reproduction of the medicinal leech *Hirudo medicinalis*. *Freshwater Biology*, 36: 563-568.

FAO 2018, FAO Yearbook: Fishery and Aquaculture Statistics, FAO.

Frank, DF, Hasenbein, M, Eder, K, Jeffries, KM, Gesit, J, Fangue, NA, Connon, RE 2017, Transcriptomic screening of the innate immune response in delta smelt during an *Ichthyophthirius multifiliis* infection. *Aquaculture*, 473: 80-88.

Fu, B, Wang, XH, Feng, X, Yu, XM, Tong, JG 2016, Comparative transcriptomic analyses of two bighead carp (*Hypophthalmichthys nobilis*) groups with different growth rates. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 20: 111-117.

Grosser, C, Pešić V, 2006, On the diversity of Iranian leeches (Annelida: Hirudinea). Archeives of Biological Sciences 58:21–24

Hibsh, D, Schori, H, Efroni, S, 2015, De novo transcriptome assembly databases for the central nervous system of the medicinal leech. *Scientific Data*, 2: 1-8.

Horgan, RP & Kenny, LC 2011, Omic technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist*, 13: 189-195.

Inbakandan, D 2020, Transcriptomics in Aquaculture. Encyclopedia of Marine Biotechnology 3: 1919-1936.

Iwama RE, Oceguera-Figueroa, A, Giribet, G. & Kvist, S 2019, The salivary transcriptome of *Limnobdella mexicana* (Annelida: Clitellata: Praobdellidae) and orthology determination of major leech anticoagulants. *Parasitology*, 1-34.

Köbis, JM, Rebl, A, Kühn, C & Goldammer, T 2013, Comparison of splenic transcriptome activity of two rainbow trout strains differing in robustness under regional aquaculture conditions. *Molecular Biology Reports*, 40: 1955-1966.

Kostyuchenko, RP & Kozin, VV 2021, Comparative aspects of annelid regeneration: Towards understanding the mechanisms of regeneration. *Genes*, 12(8): 1148.

Kvist, S, Oceguera-Figueroa, A, Tessler, M, Jiménez-Armenta, Jr, RMF, Giribet, G, Siddall, ME 2014, Pyrosequencing the salivary transcriptome of *Haemadipsa interrupta* (Annelida: Clitellata: Haemadipsidae): anticoagulant diversity and insight into the evolution of anticoagulation capabilities in leeches. *Invertebrate Biology*, 133: 74-98.

Kvist, S, essler, M, Jiménez-Armenta, J, Jr, RMF 2020, Draft genome of the European medicinal leech *Hirudo medicinalis* (Annelida, Clitellata, Hirudiniformes) with emphasis on anticoagulants. *Scientific Reports*, 10: 1-11.

Lazzarotto, V, Corraze, G, Leprevost, A, Quillet, E, Dupont-Nivet, M, Medale, F 2016, Does broodstock nutritional history affect the response of progeny to different first-feeding diets? A whole-body transcriptomic study of rainbow trout alevins. *British Journal of Nutrition*, 115: 2079-2092.

Leelatanawit, R, Lee, SY, Keenan, CP & Marsden, GE 2017, Transcriptomic analysis of male black tiger shrimp (*Penaeus monodon*) after polychaete feeding to enhance testicular maturation. *Marine Biotechnology*, 19: 125-135.

Liao, X, Cheng, L, Xu, P, Lu, G, Wachholtz, M, Sun, X, Chen, S 2013, Transcriptome analysis of crucian carp (*Carassius auratus*), an important aquaculture and hypoxia-tolerant species, *PloS ONE*, 8: e62308.

Liu, Z, Shi, P, You, H, Liu, Y, & Chen, S 2019, In-depth profiles of bioactive large molecules in saliva secretions of leeches determined by combining salivary gland proteome and transcriptome data. *Journal of Proteomics*, 200: 153-160.

Liu, Z, Kim, SI, Ha, KS, Joe, CO & Kang, K 2018, De novo assembly and comparative transcriptome characterization of *Poecilobdella javanica* provide insight into blood feeding of medicinal leeches. *Molecular Omics*, 14: 352-361.

Liu, Z, You, H, Liu, Y & Chen, S 2018, Comparative transcriptomic analysis reveals the mechanism of leech environmental adaptation, *Gene*, 664: 70-77.

Lu, Z, Shi, P, You, H, Liu, Y & Chen, S 2018, Transcriptomic analysis of the salivary gland of medicinal leech *Hirudo nipponia*, *PloS ONE*, 13(10): e0205875.

Lukin EI 1976, Piyavki [Leeches]. In: Fauna SSSR. Academy of Sciences of the USSR, vol. 1. Nauka, Moscow

Macagno, ER, Gaasterland, T, Edsall, L, Bafna, V, Soares, MB, Scheetz, T, Casavant, T, Da Silva, C & Wincker, P 2010, Construction of a medicinal leech transcriptome database and its application to the identification of leech homologs of neural and innate immune genes. *BMC Genomics*, 11(1): 407.

Martin, SA, Dehler, CE & Król, E 2016, Transcriptomic responses in the fish intestine. *Developmental & Comparative Immunology*, 64: 103-117.

Martin, SA & Król, E 2017, Nutrigenomics and immune function in fish: new insights from omics technologies. *Developmental & Comparative Immunology*, 75: 86-98.

Martínez-Porchas, M & Vargas-Albores, F 2017, Microbial metagenomics in aquaculture: a potential tool for a deeper insight into the activity. *Reviews in Aquaculture*, 9: 42-56.

Mehr, S, DeSalle, R, Kao HT, Narechania, A, Han, Z & Tchernov, D 2015, Transcriptome sequencing and annotation of the polychaete *Hermodice carunculata* (Annelida, Amphinomidae). *BMC genomics*, 16: 1-13.

Min, GS, Sarkar, IN & Siddall, ME 2010, Salivary transcriptome of the North American medicinal leech, *Macrobdella decora*, *Journal of Parasitology*, 96: 1211-1221.

Moshtaghi, A, Rahi, ML, Mather, PB, Hurwood, DA 2016, A transcriptomic scan for potential candidate genes involved in osmoregulation in an obligate freshwater palaemonid prawn (*Macrobrachium australiense*), *PeerJ*, 4: e2520.

Muttenthaler, M, King, GF, Adams, DJ & Alewood, PF 2021, Trends in peptide drug discovery. *Nature Reviews Drug Discovery*, 20: 309-325.

Nazari, S, Pourkazmi, M 2021, Isolation and characterization of SNP markers of rainbow trout (*Onchorhynchus mykiss* Walbaum, 1792) from transcriptomic sequences. *Molecular Biology Reports*, 48: 989–995.

Neave, MJ, Sunarto, A & McColl, KA 2017, Transcriptomic analysis of common carp anterior kidney during Cyprinid herpesvirus 3 infection: Immunoglobulin repertoire and homologue functional divergence. *Scientific Reports*, 7: 1-13.

Ortiz-Estrada, ÁM, Gollas-Galván, T & Martínez-Córdova, LR 2019, Predictive functional profiles using metagenomic 16S rRNA data: a novel approach to understanding the microbial ecology of aquaculture systems. *Reviews in Aquaculture*, 11: 234-245.

Park, C, Han, YH, Lee, SG, Ry, KB, Oh, J, Kern, EMA, Park, JK & Cho, SJ 2018, The developmental transcriptome atlas of the spoon worm *Urechis unicinctus* (Echiurida: Annelida). *GigaScience*, 7(3): giy007.

Petrauskiene, L 2008, The use of the medicinal leech (*Hirudo* sp.) in ecotoxicological and other scientific research—a short review. *Lauterbornia*, 65: 163-175.

Phillips, AJ & Siddall, ME 2009, Poly-paraphyly of Hirudinidae: many lineages of medicinal leeches. *BMC Evolutionary Biology*, 9: 1-11.

Rodrigues, PM, Silva, TS, Dias, J & Jessen, F 2012, Proteomics in aquaculture: applications and trends. *Journal of Proteomics*, 75(14): 4325-4345.

Rurangwa, E, Sipkema, D, Kals, J, ter Veld, M, Forlenza, M, Bacanu, GM, Smidt, H & Palstra, AP 2015, Impact of a novel protein meal on the gastrointestinal microbiota and the host transcriptome of larval zebrafish *Danio rerio*. *Frontiers in Physiology*, 6: 133.

Sağlam, N 2011, Protection and sustainability, exportation of some species of medicinal leeches (*Hirudo medicinalis* L., 1758 and *Hirudo verbana* Carena, 1820). *Journal of Fisheries Sciences*, 5: 1-15.

Santos, CA, Blanck, DV & de Freitas PD 2014, RNA-seq as a powerful tool for penaeid shrimp genetic progress. *Frontiers in Genetics*, 5: 298.

Schaeck, M, Reyes-López, FE, Vallejos-Vidal, E, Van Cleemput, J, Duchateau, L, Van den Broeck, W, Tort, L & Decostere, A 2017, Cellular and transcriptomic response to treatment with the probiotic candidate *Vibrio lentus* in gnotobiotic sea bass (*Dicentrarchus labrax*) larvae. *Fish & shellfish Immunology*, 63: 147-156.

Siddall, ME, Trontelj, P, Utevsky, SY, Nkamany, M & Macdonald, KS 2016, Comparative transcriptomic analyses of three species of *Placobdella* (Rhynchobdellida: Glossiphoniidae) confirms a single origin of blood feeding in leeches. *Journal of Parasitology*, 102: 143-151.

Sig, AK, Guney, M, Uskudar Guclu, A, Ozmen, E 2017, Medicinal leech therapy—an overall perspective. *Integrative Medicine Research*, 6: 337-343.

Simakov, O, Marlétaz, F & Yue, JX 2013, Insights into bilaterian evolution from three spiralian genomes. *Nature*, 493(7433): 526.

Sudhagar, A, Kumar, G & El-Matbouli, M 2018, Transcriptome analysis based on RNA-Seq in understanding pathogenic mechanisms of diseases and the immune system of fish: a comprehensive review. *International Journal of Molecular Sciences*, 19(1): 245.

Tong, L, Dai, SX, Kong, DJ, Yang, PP, Tong, X, Tong, XR, Bi, XX, Su, Y, Zhao, YQ, Liu, ZC 2020, The genome of medicinal leech (*Whitmania pigra*) and comparative genomic study for exploration of bioactive ingredients, *BMC Genomics*, 23(1):76.

Trontelj, P & Utevsky, SY 2005, Celebrity with a neglected taxonomy: molecular systematics of the medicinal leech (genus *Hirudo*). *Molecular Phylogenetics and Evolution*, 34: 616-624.

Utevskaya, O, 1998, Analysis of reproductive ability, Vestnik zoologii, 32: 119-122.

Wang, C, Gerstein, M & Snyder, M 2014, Analysis of the skin transcriptome in two oujiang color varieties of common carp. *PloS ONE*, 9(3): e90074.

Whitaker, IS, Rao, J, Izadi, D, Butler, PE 2004, Historical Article: *Hirudo medicinalis*: ancient origins of, and trends in the use of medicinal leeches throughout history. *British Journal of Oral and Maxillofacial Surgery*, 42: 133-137.

Wilkin, P 1989, The medicinal leech, *Hirudo medicinalis* (L.) (Hirudinea: Gnathobdellae), at Dungeness, Kent. *Botanical Journal of the Linnean Society*, 101(1): 45-57.

Wilkin, P & Scofield, A 1991, Growth of the medicinal leech, *Himudo medicinalis*, under natural and laboratory conditions. *Freshwater Biology*, 25: 547-553.

Xue, X, Feng, C, Hixson, SM, Johnstone, K, Anderson, D, Parrish, C, Rise, M 2015, Atlantic salmon (*Salmo salar*) liver transcriptome response to diets containing Camelina sativa products. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 14: 1-15.

Yáñez, JM, Newman S & Houston RD 2015, Genomics in aquaculture to better understand species biology and accelerate genetic progress. *Frontiers in Genetics*, 6: 128.

Ye, H, Yoshida, A, Wang, X, Chiba, Y & Jigami Y 2018, Applications of transcriptomics and proteomics in understanding fish immunity. *Fish & Shellfish Immunology*, 77: 319-327.

Young, T & Alfaro, AC 2018, Metabolomic strategies for aquaculture research: a primer, Reviews in Aquaculture, 10(1): 26-56.

Yue, G & Wang, L 2017, Current status of genome sequencing and its applications in aquaculture. *Aquaculture*, 468: 337-347.

Zaidi, S, Jameel, SS, Zaman, F 2011, A systematic overview of the medicinal importance of sanguivorous leeches, *Alternative Medicine Review*, 16: 59-65.

Żarski, D, Nguyen T, Le Cam, A, Montfort, J, Dutto, G, Vidal, MO, Fauvel, C, & Bobe J 2017, Transcriptomic profiling of egg quality in sea bass (*Dicentrarchus labrax*) sheds light on genes involved in ubiquitination and translation. *Marine Biotechnology*, 19: 102-115.

Zhao, Y, Sheng, Z, Huang, J 2012, A systematic analysis of heart transcriptome highlights divergent cardiovascular disease pathways between animal models and humans. *Molecular BioSystems*, 8: 504-510.