Transcriptome Analysis Reveals the Potential Roles of Long Non-Coding RNAs in Feed Efficiency of Chicken

Parastoo Karimi
University of Tehran
Mohammad Reza Bakhtiarizadeh (mrbakhtiari@ut.ac.ir)
University of Tehran
Abdolreza Salehi
University of Tehran
Hamid Reza Izadnia
Animal Science Improvement Research Department, Agricultural and Natural Resources Research and Education Center, Safiabad AREEO, Dezful

Research Article

Keywords: Feed efficiency, molecular mechanisms, lncRNAs, RT-qPCR, metabolism

Posted Date: September 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-882290/v1

Version of Record: A version of this preprint was published at Scientific Reports on February 15th, 2022. See the published version at https://doi.org/10.1038/s41598-022-06528-6.

Abstract

Feed efficiency is an important economic trait and reduces the production costs per unit of animal product. Up to now, few studies have conducted transcriptome profiling of liver tissue in feed efficiency-divergent chickens (Ross vs native breeds). Also, molecular mechanisms contributing to differences in feed efficiency are not fully understood, especially in terms of long non-coding RNAs (lncRNAs). Hence, transcriptome profiles of liver tissue in commercial and native chicken breeds were analyzed. RNA-Seq data along with bioinformatics approaches were applied and a series of lncRNAs and target genes were identified. Furthermore, protein-protein interaction network construction, co-expression analysis, co-localization analysis of QTLs and functional enrichment analysis were used to functionally annotate the identified lncRNAs.In total, 2,290 lncRNAs were found (including 1,110 annotated, 593 known and 587 novel), of which 53 (including 39 known and 14 novel), were identified as differentially expressed genes between two breeds. The expression profile of lncRNAs was validated by RT-qPCR. The identified novel lncRNAs showed a number of characteristics similar to those of known lncRNAs. Target prediction analysis showed that these lncRNAs have the potential to act in cis or trans mode. Functional enrichment analysis of the predicted target genes revealed that they might affect the differences in feed efficiency of chicken by modulating genes associated with lipid metabolism, carbohydrate metabolism, growth, energy homeostasis and glucose metabolism. Some gene members of significant modules in the constructed co-expression networks were reported as important genes related to feed efficiency. Co-localization analysis of QTLs related to feed efficiency and the identified lncRNAs suggested several candidates to be involved in residual feed intake. The findings of this study provided valuable resources to further clarify the genetic basis of regulation of feed efficiency in chicken from the perspective of lncRNAs.

Introduction

Along with the increase in the world's population and the growing need for food, in particular, protein substances, meeting human nutritional needs is primarily an issue for human societies. Chicken meat is a high-quality protein source for humans due to the high levels of protein, low cholesterol and low calories. The poultry industry plays a vitally important role in converting grains into meat and eggs, which are essential sources of food for humans (Alsaffar and Khalil, 2015). In response to the global demand for chicken meet, genetic selection enabled us to improve many traits related to the functions of chickens and led to faster growth rates and higher feed efficiency in different commercial races such as Ross (Liu et al., 2011; Rekaya et al., 2013). On the other hand, given the importance of native chicken breeds in rural economies in most of the developing and underdeveloped countries in having greater resistance to disease and the ability to tolerate harsh environmental conditions, their genetic improvement is reasonable in order to increase their efficiency. In this regard, Esfahan chicken is one of the native and dual-purpose breeds (to produce meat and eggs) in Iran (Yousefı Zonuz et al., 2013). Previous studies have demonstrated that feed efficiency is affected by many factors including food intake, genetic diversity in energy requirements for maintenance, the digestion, metabolism of nutrients, diversity in body composition and body temperature regulation) Luiting et al., 1991). Changes in the amount of energy required to digest food, the size of the digestive organs, and the amount of energy consumed within the tissues can all contribute to feeding intake variation (Herd and Arthur, 2009). Herein, birds with low-level RFI (Ross) means to consume less feed than predicted and are identified as efficient birds (Meale et al., 2017). Furthermore, compared to the low efficient chickens, high efficient one may synthesize ATP more efficiently and control reactive oxygen species (Ross) production more strictly by enhancing the mitochondrial function in skeletal muscle (Yang et al., 2020). According to the importance of Esfahan chickens compared to commercial breed in having more resistance to diseases, efforts to improve chickens with efficient performance and a better feed efficiency would be beneficial with emphasis on genetic improvement. In this context, understanding the genetic mechanisms underlying the differences between native and commercial chicken breeds can be a promising strategy for a quick genetic improvement of the Esfahan breed.

Gene expression profiling is a powerful approach in functional genomics to uncover the complexity of cells and tissues at the transcriptional level. Next generation sequencing (NGS) based technologies are revolutionizing the field and paved the way to profile the gene expression of many organisms. In this context, RNA sequencing (RNA-Seq) is the current gold-standard method for genome-wide gene expression studies, which takes advantage of NGS. Hence, transcriptome analysis can help us to identify the molecular mechanisms affecting different traits such as feed efficiency. On the other hand, it is well known that the majority of the genome is expressed to non-coding RNAs (ncRNAs), which play critical roles in various biological processes (Korostowski

et al., 2012). One type of important ncRNAs that have been taken into consideration is Long noncoding RNA (lncRNA). LncRNAs are mostly defined as RNAs greater than 200 nucleotides in length, without any coding potential (Liu et al., 2019). Most of the lncRNAs are processed like mRNA to be 5' capped and 3' poly adenylated as well as to be alternatively spliced (Hezroni et al., 2015; Meseure et al., 2016). In spite of lacking protein-coding potential, recent studies have shown the presence of short open reading frames (ORFs) that encode small peptides with vital biological functions. It has been reported that lncRNAs are less evolutionary conserved and tend to be expressed lower than that of protein-coding genes, as most of them are tissue specific. The most important functions of lncRNAs are regulating gene expression (by cis or trans-acting) through epigenetic, transcriptional, post-transcriptional, or post-translational mechanisms (Hezroni et al., 2015; Iyer et al., 2015; Sui et al., 2019).

A growing number of reports have demonstrated regulatory functions of lncRNAs in different chicken tissues using RNA-SEq. For instance, Li et al., identified differentially expressed lncRNAs from chicken skeletal muscle and reported that these genes might play important roles in the muscle development (Li et al., 2017). Moreover, the previous study showed that three lncRNAs may be associated with immune response regulation and could function as novel biomarkers in chickens infected with Avian leukosis virus subgroup J (Qiu et al., 2017). In a recent study, Muret et al., for the first time, reported a new lncRNA (lnc_DHCR24) as a key enzyme of cholesterol biosynthesis in chicken lipid metabolism (Muret et al., 2017). Also, a study on six Beijing-you cocks, divergent in sperm motility led to the identification of 2,597 lncRNAs, including 1,267 intergenic lncRNAs (lincRNAs), 975 anti-sense lncRNAs and 355 intronic lncRNAs (ilncRNAs). Of these, 124 were differentially expressed and two lncRNAs including MSTRG.3652 and MSTRG.4081 and their target genes, showed their potential in sperm motility regulation (Liu et al., 2017). In addition, many studies have suggested that lncRNAs play an important controlling role during lipid metabolism (Muret et al., 2017; Li et al., 2018), carbohydrate metabolism (Ren et al., 2018; Tang et al., 2020), growth traits (Li et al., 2020), liver development (Muret et al., 2017; Ning et al., 2020), ion transfer (Cao et al., 2017; Yang et al., 2019) and digestive enzymes (Li et al., 2020) in chickens, which can affect feed efficiency in this way. In spite of these studies, the regulatory roles of lncRNAs in feed efficiency, especially in native breeds remain to be elucidated and studies which adopt genome-wide approaches are required to gain greater insight into these mechanisms. Additionally, there is a lack of information concerning the differences between native and commercial chicken breeds. In this study, RNA-Seq technology was applied to compare expressed lncRNAs in the liver tissue of two chicken breeds (Ross as commercial and Esfahan as native breed). The main goal was to gain a better understanding of biological and molecular mechanisms that are involved in the feed efficiency between native and commercial chicken breeds and establishing a foundation for future molecular studies. Moreover, using a stringent bioinformatic pipeline, novel lncRNAs were predicted, which can improve annotation of the chicken genome.

Material And Methods

Animals, RNA Extraction and Sequencing

RNA-Seq data used in this study was obtained from our previous study. Detailed information about the samples and the experimental design has been described previously (Izadnia et al., 2018). A population of 120 roosters came from 60 Ross 708 as a commercial breed with low RFI and 60 Isfahan chicken as a local breed with high RFI was selected from the base population of Safiabad Agricultural and Natural Resources Research and Education Center. They were brooded in individual cages and raised under the same environmental and nutritional conditions for 42 days. To better understand the difference in the growth and feeding characteristics feed intake DFI and BWG of each individual were monitored on a daily basis from 24 to 42 d of age. In addition, RFI, mean BW (MBW), and metabolic BW (MBW0.75) were calculated. At the age of 42 days post-hatch, 10 birds were randomly selected from each breed and sacrificed. Then, 1–2 g of their liver tissues was excised from similar parts of each bird. All tissue samples were snap frozen in liquid nitrogen and then transferred to a − 80°C freezer until required for RNA isolation. Six individual RNA samples that were related to six birds in each breed (including Ross 708 chickens as a commercial breed and Esfahani chickens as a native breed) were used to construct pooled samples. Totally, two pooled RNA samples were constructed (by mixing together equal quantities of RNA) in each breed (including three samples per each pooled sample) and were sent to BGI company (Shenzhen, China). Samples were applied for sequencing if the RNA Integrity Number (RIN) was > 7, based on an Agilent Bio Analyzer 2100 system. All the samples were subjected to a paired-end sequencing of 150 bp on the

Illumina HiSeq 2000 platform (Illumina, San Diego, California, USA). The raw RNA-Seq data were deposited and released in SRA database, with the BioProject accession number of PRJNA707148 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA707148).

RNA-Seq Data Analysis

A computational approach using different filter criteria was applied to identify the lncRNAs (Fig. 1). Preliminary quality control analysis of the raw reads was checked with FastQC (v 0.11.2). Raw reads were processed for initial trimming and filtering of the low-quality reads using the Trimmomatic tool (v 0.35) (Bolger et al., 2014). Then, STAR software (version 2.5.3.e) was applied to align the clean reads to the reference genome (Galgal5) using the default parameters (Dobin et al., 2013). StrnigTie tool (version 1.3.4.d) was used to assemble the aligned reads for each sample into transcripts (Pertea et al., 2015). Then, all the assembled transcripts of four samples were merged using the StrnigTie tool to generate a global and unified set of transcripts across the samples. This approach maximizes the overall quality of the final assembly. Differential expression analysis of the annotated genes was performed using Cuffdiff tool (v.2.2.1). To improve differential expression analysis for less abundant transcripts, the upper quartile normalization option was chosen. Finally, genes with a false discovery rate (FDR) ≤ 0.05 were considered as differentially expressed genes (DEGs) between the two breeds (Trapnell et al., 2010). The reference genome for chicken and the annotation GFT file (release 91) were downloaded from the Ensemble database (Pertea et al., 2015).

LncRNA Identification Pipeline

In order to discriminate the novel and known transcripts, the unified set of transcripts were compared against the Ensemble chicken genome annotation (release 91) by Gffcompare (https://github.com/gpertea/gffcompare). Using this approach annotated lncRNAs were found. Then, only the unknown intergenic (marked with class code “u”) and intronic (marked with class code “i”) transcripts were subjected to further analyses. All unknown transcripts were blasted against NONECODE database (v5.0) to identify the reported chicken lncRNAs in this database. It is worth to note that there are a large number of chicken lncRNAs in this database that are not included in the Ensemble chicken genome annotation. All the significant hits (E-value = 1e-5) were considered as known lncRNAs and were further classified to intergenic lncRNAs (lincRNAs) and intronic lncRNAs (ilncRNAs) based on their class codes. Different kinds of lncRNAs based on their genomic origins relative to nearby protein-coding genes can be grouped as sense or antisense transcript overlapping and as promoter associated, intronic or intergenic (Vance and Ponting, 2014). An ilncRNA possesses a transcribed region overlapped with the intron segments of protein-coding genes, and a lincRNA is located in the intergenic region lacking known protein-coding genes (Ma et al, 2014). Also, lincRNAs have high relative abundance in the genome, for instance, Kuo et al (2017) found that lincRNAs accounted for more than half of chicken lncRNAs. For the remaining transcripts, longer than 200 nt, the following stringent filtering criteria were performed to identify the potential novel lncRNAs. First, the transcripts with one exon that were longer than 10000 nt or those were located in the genomic regions containing simple sequence repeats (based on UCSC RepeatMasker file) were removed. Second, the transcripts with FPKM ≥ 1 expressed in at least two samples were kept. Third, all the remaining transcripts that were blasted (E-value < 1e-5) against one of the following databases were excluded from further analysis: UniprotKB (by BLASTx), miRbase (release 21, by BLASTn) and Rfam (by BLASTn). Also, Hmmscan tool from the HMMER package (version 3.1b2) was applied to evaluate which of the remaining transcripts contains a known protein domain with any of the known protein family domains in the Pfam database (Bateman et al., 2004). The significant hits were then removed. Forth, to predict the coding potential of the remaining transcripts, five software including CPC2 (version 2 beta) (score > 0.5) (Kang et al., 2017), CNCI (version 2) (score > 0) (Sun et al., 2013), CPAT (version 1.2.4) (score > 0.36) (Wang et al., 2013), PLEK (version 1.2) (score > 0) (Li et al., 2014) and FEElnc (version 12/07/2016) (default parameters) (Wucher et al., 2017) were used. Only the transcripts that were simultaneously considered to have coding potential by at least three of the mentioned tools were excluded from further analysis. Next, putative protein-coding transcripts were filtered out by considering open reading frame (ORF) length greater than 300 aa, which were predicted using TransDecoder tool (v2.1.0) (https://transdecoder.github.io/). In the last step, the intergenic transcripts that were located in a distance < 1 kb to a known protein-coding gene, were excluded. Finally, all the transcripts that successfully passed the filtering steps, were considered as putative novel lncRNAs and classified into lincRNAs and ilncRNAs classes (Fig. 1).

Functional Analysis

To investigate various aspects of the putative lncRNAs, five different analyses were performed based on our previous study, which is described as follow (Bakhtiarizadeh and Salami, 2019):

1. **Analysis of Conservation**: Previous studies have shown that lncRNAs are not well conserved at the sequence level (Ulitsky, 2016; Foissac et al., 2018). To assess this, all the putative novel lncRNAs were blasted against lncRNA sequences of human, rat and cow using BLASTn. All the lncRNA sequences were obtained from NONCODE database (v5.0). The hits meeting the E-value threshold < 10 − 5 were considered significant.
2. **Genomic Synteny Analysis**: It is well known that establishing homology relationships among lncRNAs is challenging due to their low sequence conservation. Previous studies suggested that genomic position (synteny) of lncRNAs is more conserved than their cross-species sequence conservation. Therefore, lncRNAs conservation is geared toward the maintenance of synteny (Muret et al., 2019). In other words, suppose neighboring upstream and downstream protein coding genes of a lincRNA in chicken are X and Y. If there be a lincRNA in other organism that is located between the same X and Y protein coding genes, it can be considered as a conserved syntenic lincRNA. Here, to identify syntenic lncRNAs, order of surrounding genes flanking each putative novel lincRNA was compared for conserved synteny in the human and bovine genomes. All the genomic positions of lincRNAs were extracted from the NONCODE database (v5.0) for human and bovine.
3. **Target Prediction and Functional Enrichment Analysis**: Typically, lncRNAs regulate their target genes through cis or trans modes. The cis-acting refers to the regulation of genes in a host (in case of ilncRNAs) or nearest neighboring protein coding genes (in case of lincRNAs). According to the location of the lncRNAs and mRNAs on the chicken genome, 10/100 kb upstream or downstream of the differentially expressed lncRNAs were searched to identify the nearest neighboring protein coding genes, which can be considered as potential target genes of these lincRNAs. On the other hand, trans-acting refers to the influence of lncRNAs on other genes at expression levels, which can be assessed by co-expression analysis (correlating expression levels between lncRNAs and mRNAs) (Hezroni et al., 2015; Sui et al., 2019). To predict the potential target genes of the lncRNAs in trans mode, Pearson correlation coefficient was calculated based on the expression levels of each pair of lncRNA and mRNA ($r > 0.95$ or $r < -0.95$). Functional enrichment analysis including Gene Ontology (GO) and KEGG pathway analysis related to the predicted target genes of lncRNAs was conducted using the EnrichR database (Kuleshov et al., 2016). A term with an adjusted P-value < 0.1 (FDR) was considered to be significantly enriched.
4. **QTL Mapping Analysis of LncRNAs**: The co-localization analysis was performed to determine if the putative novel lncRNAs are located in the QTLs related to feed efficiency. For this purpose, all QTLs related to feeding efficiency were retrieved from AnimalQTLdb (Hu et al., 2016). Then positions of the lncRNAs were compared to the QTL's locations.
5. **Target Prediction of the Candidate LncRNAs and PPI Network**: The potential protein-protein interaction (PPI) network was constructed for the predicted cis and trans target genes of the lincRNAs and ilncRNAs (known and novel) using STRING database (Search Tool for the Retrieval of Interacting Genes/Proteins, https://string-db.org/) (Mering et al., 2003). Then, Cytoscape software (version 3.7) (Shannon et al., 2003) was applied to visualize and combine the significant interactions obtained by STRING along with predicted cis and trans target genes of lncRNAs (separately for each of lincRNAs and ilncRNAs). The key sub-networks (module genes) were detected using a Cytoscape plugin, ClusterONE tool (version 1.0), based on minimum size = 5 and a minimum number of the genes in a module = 5 (Nepusz et al., 2012). In order to further explore the characteristics of the significant modules, functional enrichment analysis was performed on their gene members.

Quantitative Real-Time PCR (RT-qPCR) Validation

To validate the RNA-Seq results by RT-qPCR, seven differentially expressed lncRNAs (including four known lincRNAs and three novel lincRNAs) were randomly selected. Two biological replicates were conducted for each breed. Total RNA was extracted from the liver of chickens using TRIzol (Invitrogen, CA, USA) according to the manufacturers' protocols and the concentration and purity of RNAs were measured using NanoDrop 2000. The cDNA was synthesized from purified RNA by using an RT-qPCR kit (Takara, Dalian, China). Primer 3 plus software was applied to design the primers (Untergasser et al., 2007). The details of all the designed primers are provided in Supplementary File S1. RT-qPCR was performed on a Light Cycler 96 instrument (Roche Co. Germany) using HiFi SYBR Green Master Mix (Thermo Fisher Scientific, USA). RT-qPCR was performed with the following thermo cycling conditions: an initial 1 cycle at 95°C for 10 min, 40 cycles at 95°C for 15s, 60°C for 20s and 72°C for 20s, followed by a

72°C elongation for 60s. GAPDH and ACTB were used as house-keeping genes to normalize the expressions of lncRNAs (Kuang et al., 2018). The results were analyzed using the 2 − ΔCt method.

Results

Phenotype measurements

The average BWG of higher RFI and lower RFI groups were 441.06 ± 6.76 g and 1027.88 ± 7.57 g, respectively. There were significant differences ($p < 0.01$) in daily feed intake (DFI) between native and commercial chickens (1329.25 ± 8.05 vs 2007.90 ± 9.25). The lower RFI had a higher average daily weight gain than a higher one. Consequently, the differences in mean RFI values between LRFI and HRFI chickens in each breed were highly significant ($p < 0.01$). The local breed had the RFI values of 13.430 ± 5.393 (g/day) compared with − 11.212 ± 4/435 (g/day) for the commercial breed during 19 days (day 24–42) of the experiment."

RNA-Seq Data Analysis

A total of 35,696,493 and 34,824,326 raw paired-end reads were obtained in native and commercial breeds, respectively. Following a pipeline of adaptor removal, quality filtering and the removal of sequences that were too short, only 565,714 and 534,690 paired reads did not pass the quality filtering and were removed in native and commercial breeds, respectively. In total, 84.70-86.97% of the clean reads from all the samples were successfully aligned to the reference genome. Of the successfully aligned reads, 82.90-84.53% were uniquely aligned (Table 1).

Table 1
.Number of RNA-Seq reads and mapping rates of all samples.

Sample name	Raw reads	Clean reads	Uniquely mapped reads (%)	Total mapped reads (%)
Native 1	14,510,042	14,246,264	11,810,777 (82.90)	12,140,023 (85.22)
Native 2	21,186,451	20,884,515	17,653,576 (84.53)	18,162,985 (86.97)
Commercial 1	19,034,129	18,723,445	15,607,073 (83.35)	16,028,736(85.61)
Commercial 2	15,790,197	15,566,191	12,859,795 (82.61)	13,183,626(84.70)

LncRNA Identification and Characterization

To comprehensively identify lncRNAs in the genome wide, a stringent bioinformatics pipeline was applied. In total, 37,663 transcripts were generated through reconstructing the transcripts of all the samples by StringTie. All assembled transcripts were compared against ENSEMBL chicken GTF file and 20,846 and 1,110 transcripts were annotated as mRNAs and lncRNA, respectively. After comparing the remaining transcripts against chicken lncRNAs from the NONCODE database, 593 transcripts were identified as known lncRNAs (including 525 lincRNA ad 68 ilncRNAs), which are not annotated in the chicken GTF file (ENSEMBL database), yet. Finally, according to our stringent lncRNA prediction pipeline, 587 putative novel lncRNAs including 454 lincRNAs and 133 ilncRNAs were identified (Fig. 1).

To understand the differences in characteristics of novel lncRNAs and verify their accuracy, their exon number, GC content, transcript length and expression level were compared against known and annotated lncRNAs as well as mRNAs.

The results revealed that average GC content of the novel lincRNAs and novel ilncRNAs was similar (approximately 42%), which is lower than known mRNAs (50%) (Fig. 2A). GC content of annotated lncRNAs, known ilncRNAs, known lincRNAs, novel ilncRNAs and novel lincRNAs were 47, 46, 47, 43 and 42%, on average, respectively. The transcript lengths of novel lncRNAs ranged from 203 to 5,889 bp, with an average of 733 bp, which was lower than the values observed for mRNAs (with average

length 3193.793 bp) (Fig. 2B). The average transcript length of annotated lncRNAs, known ilncRNAs, known lincRNAs, novel ilncRNAs and novel lincRNAs were 1,100, 1,617, 1,995, 685, and 782, respectively. On average, the exon numbers associated with novel lncRNAs (1.18 on average) were clearly less than mRNAs (10.82 on average) (Fig. 2C). Exon numbers of annotated lncRNAs, known ilncRNAs, known lincRNAs, novel ilncRNAs and novel lincRNAs were 2.8, 1.6, 2.6, 1.14 and 1.3, on average, respectively. Most novel lncRNAs (571 of 587, 97%) contained no more than two exons, while above 83% mRNAs had no less than three exons. The expression levels of mRNAs and novel lncRNAs were further compared according to FPKM values, and the Fig. 2D shows that the expression levels of mRNAs in all samples were higher than lncRNAs (the average expression of annotated lncRNAs, known ilncRNAs, known lincRNAs, mRNAs, novel ilncRNAs and novel lincRNAs in different samples were 0.39, 1.41, 1.56, 1.67, 1.53, and 1.46, respectively).

In terms of chromosomal distribution, the novel lncRNAs were distributed across nearly all of the chicken chromosomes, except chromosome 16. Chromosome 1 has the greatest number of lncRNAs (97 lncRNAs), followed by chromosome 2 (72 lncRNAs), 3 (71 lncRNAs), 5 (51 lncRNAs) and Z (43 lncRNAs), whereas chromosomes 22, 23 and 25 contained only ilncRNAs (1, 1 and 3), respectively (Fig. 3).

Conservation Analysis

Out of 454 novel lincRNAs, six, nine and 12 lincRNAs were found to have homology to bovine, mouse and human lincRNAs, respectively. Also, three, five and four known lincRNAs had conserved sequences in bovine, mouse and human, respectively. Moreover, two and two novel ilncRNAs and one and two known ilncRNAs were evolutionary comparable with lncRNA sequences from bovine and mouse, respectively. No significant hit was found in comparison of ilncRNAs (known and novel) and human lncRNAs. In addition, of 16 novel and nine known conserved lincRNA transcripts, eight and two genes showed evidence of sequence homology among bovine, human or mouse (Supplementary File S2). Overall, transcript-level homology of our predicted lncRNAs with active transcribed lncRNAs in human, cow and mouse detected, and found that only 2% of known and 3% of novel lncRNAs can be aligned to investigated species. In agreement with our results, it was reported that less than 1% of chicken lincRNAs owned detectable sequence conservation with human or mouse lincRNAs. These findings confirm the previous reports that lncRNAs in vertebrates have low levels of interspecies short and highly conserved regions (Liu et al., 2017).

Synteny Analysis

The results of the synteny analysis indicated a similar structural architecture between the performed comparisons. Results of this analysis revealed that 62 known and 96 novel lincRNAs as well as 89 known and 116 novel lincRNAs were located in the neighborhood of respective orthologous protein coding genes in chicken/cow and chicken/human comparisons, respectively (Supplementary File S3). Out of 62 and 96 lincRNAs, two and three lincRNAs had conserved sequencing with bovine and human, respectively. Also, two known and four novel lincRNAs, out of 89 and 116 syntenic genes, had an orthologous sequence in bovine and human, respectively.

Moreover, 121 known lincRNAs and 212 novel lincRNAs were determined with conserved synteny among the investigated species. Of these, 26% (31 of 121) of known lincRNAs and 23% (48 of 212) of novel lincRNAs showed sequence homology with the investigated species. These findings were consistent with the previous study that reported higher conservation in synteny of the lncRNAs in comparison to their sequences (Bakhtiarizadeh and Salami, 2019).

Differential Expression Analysis

In total, 17,833 transcripts (including 398 annotated transcript lncRNAs, 10,624 known genes and 6,810 novel transcripts) were identified to be expressed in both breeds. Of these, 1,040 and 1,081 transcripts were expressed only in commercial and native breeds, respectively. Therefore, 15,712 transcripts were commonly expressed in both breeds. To identify potential feed efficiency related genes, differential expression analysis was applied and FPKM values of the genes were compared between the two chicken breeds. The analysis showed that 39 known lncRNAs (including one up-regulated ilncRNAs, eight up and 30 down-regulated lincRNAs) and 87 known mRNAs (including 63 up- and 24 down-regulated) were DEGs. In addition, 14 novel lincRNAs (four up and 10 down-regulated) tended to be differentially expressed. No annotated lncRNA and novel ilncRNAs were identified to be differentially expressed. The top five differentially expressed lncRNAs in each lncRNA classes are represented in Table 2 (the complete list of these genes is provided in Supplementary File S4).

Table 2
The top five differentially expressed genes lncRNAs classes between the two breeds and their cis target gene.

Source	ID	Closest left mRNA	Closest right mRNA	Expression in native Esfahan	Expression in commercial chicken	Adjusted p-value
Novel lincRNA	lincRNA.10349.1	KIAA0408	ECHDC1	2.86443	0	5.00E-05
	lincRNA12629.1	RTF1	ENSGALG00000008599	3.55402	17.0844	5.00E-05
	lincRNA.224.1	FAM19A5	ENSGALG00000039992	5159.49	1740.5	5.00E-05
	lincRNA15581.1	PRLR	SPEF2	0	1.87716	0.00015
	lincRNA.5050.1	TRAF1	C5	1.48732	4.27511	0.00065
Known lincRNA	lincRNA.1424.2	SH3BGR	B3GALT5	5.90444	15.7968	5.00E-05
	lincRNA.16016.1	RNF38	TRIM14	17.4264	6.9754	5.00E-05
	lincRNA.14534.1	LRRC8D	LRRC8C	4.24643	11.6967	5.00E-05
	lincRNA.3640.1	SLC26A2	ENSGALG00000001206	65.8459	32.8307	5.00E-05
	lincRNA.7295.1	COL22A1	KCNK9	7.77861	34.7036	0.00025
Known ilncRNA	ilncRNA.8900.7	ENSGALG00000043654	ENSGALG00000043654	21.2565	9.19081	0.00035

RNA-Seq Data Validation by RT-qPCR

To further confirm the accuracy of the sequencing data, RT-qPCR was performed on seven differentially expressed lincRNAs. As shown in Fig. 4a similar expression pattern (except linc.1424.1) was observed in RT-qPCR compared to RNA-Seq data. However, there were variations observed in these methods, which can be attributed to intrinsic features of these approaches.

LncRNA Target Genes and Functional Analysis

To understand the biological functions of the lncRNAs, cis-regulated genes of these genes were investigated. The genes transcribed in the 10/l00 kb upstream or downstream around the lincRNAs are generally considered to be cis target genes. All 52 (38 known and 14 novel) differentially expressed lincRNAs were detected to be located near of 70 mRNAs and these mRNAs were involved in "lipid metabolism", "carbohydrate metabolism", "growth" and "cell death".

Trans target genes of the known and novel lncRNAs were predicted through expression correlation analysis. In total, 53 lincRNAs (48 known and five novel lincRNAs) were corresponded to 619 target genes (Supplementary File S5). The identified co-expressed genes were similar for the most of known lincRNAs. Accordingly, lincRNAs with the similar target genes were clustered, which led to four clusters. The first, second, third, and fourth clusters included 4, 3, 15, and 24 lincRNAs, which were co-expressed with 179, 170, 133, and 130 target genes, respectively. Functional enrichment analysis of the target genes in cluster number one and two were significantly enriched in one and three GO terms and one and three KEGG pathways, respectively, which were related to biological processes such as “gamma-aminobutyric acid signaling pathway”, “ear morphogenesis”, “cochlea morphogenesis” and “inner ear morphogenesis”. Also, nine and 29 significant GO terms were found in cluster number three and four, respectively. The detailed information related to these clusters and their functional analysis are provided in Supplementary File S6. According to the previous studies, feed efficiency-related terms in cluster number three and four were detected including “monocarboxylic acid transport”, “potassium ion transport”, “chloride ion homeostasis”, “negative regulation of calcium ion trans membrane

transport” and “cellular response to cAMP”. KEGG pathway analysis was further performed and no significant pathways was identified.

In ilncRNAs, 10 known ilncRNAs were corresponded to 265 target genes (Supplementary File S5). Based on the co-expressed genes of the known ilncRNAs, which were similar among the ilncRNAs, two clusters were constructed. Three ilncRNAs were expressed in cluster number one with 130 target genes. Also, there were four ilncRNAs corresponding to 130 target genes in the second cluster. Interestingly, GO analysis of these target genes led to detection of 38 enriched terms (nine GO terms in cluster one and 29 GO terms in cluster two) (Fig. 5), which were similar to those observed in the cluster number three and four of lincRNAs (Supplementary File S6).

Next, PPI network construction was performed based on the target genes of lncRNAs. In total, two significant PPI networks were obtained including 1,320 and 398 genes and 5,129 and 440 interactions for lincRNAs and ilncRNAs, respectively. After combining PPI network and cis and trans target genes of lincRNAs, 1,029 lincRNAs (573 known and 454 novel) were connected to 1,514 mRNAs with 11,160 connection edges. Moreover, a network with 216 ilncRNAs (68 known and 168 novel) and 416 protein-coding genes along with 1,490 interactions were obtained. ClusterONE tool predicted one significant module in each of the networks with < 20 nodes including green module (P-value = 0.0007, Fig. 6) in lincRNA-based network and light-blue module (P-value = 6.147-7, Fig. 7) in ilncRNA-based network.

Functional enrichment analysis indicated that genes in green module (lincRNA-based network) were significantly enriched in 165 GO terms and seven KEGG pathways, which were related to glucose homeostasis (GO:0042593), calcium ion transport (GO:0006816), insulin receptor signaling pathway (GO:0008286), lipid phosphorylation (GO:0046834). The light-blue module (ilncRNA-based network) were enriched in 132 metabolic pathways and four KEGG pathways including “positive regulation of ion transport”, “positive regulation of sodium ion trans membrane transporter activity”, “negative regulation of lipid metabolic process” and “carbohydrate homeostasis”. The functional enrichment analysis of the two modules showed that both are related to feed efficiency (Supplementary File S7).

QTL Analysis

To further investigate the potential of the lncRNAs in feed efficiency related traits, a co-localization analysis was performed with feed efficiency associated QTLs. The current version of the chicken QTLdb includes 11,340 QTLs representing 347 different traits. Of these, eight hundred and thirty-four QTLs that were associated with six traits related to feed efficiency were considered. Of these, 29 QTLs were found to be overlapped with the novel lncRNAs. In this regard, out of 454 and 133 novel lincRNAs and ilncRNAs, 47 and 11 lncRNAs were located in 19 and 10 QTLs, respectively, including 36 lncRNAs in “feed intake” and 23 lncRNAs in “residual feed intake” (Supplementary File S8). Of these, ilncRNA.2932 and lincRNA.5494, which were located in feed intake related QTLs, were specifically expressed in commercial breed. Out of 47 novel lincRNAs, one differentially expressed lincRNA (lincRNA.10349) was overlapped with a QTL related to residual feed intake (RFI). Moreover, seven lincRNAs (lincRNA.1964, lincRNA.2201, lincRNA.2215, lincRNA.5649, lincRNA.2920, lincRNA.2178, and lincRNA.2186) were located in more than one QTL related to feed efficiency (Supplementary File S8). Interestingly, several cis target genes of these lincRNAs, including KCNRG, DLEU7, CTSC and RAB30, were reported to be involved in feed efficiency and RFI or has a known role the growth rate (Zhu et al., 2012; Wolc et al., 2013; Gondret et al., 2017). These genes can be considered as promising candidate genes responsible for feed efficiency in chicken (Table 3).

Table 3
.The list of identified novel lincRNAs in more than one QTL related to feed efficiency.

lncRNA	lncRNA position	Number of QTL	QTL Id	Closest left mRNA	Closest right mRNA
lincRNA.2178	1:186194986186196686	3	64551,64552,64553	SLC36A4	-
lincRNA.1964	1:169421116–169445490	3	64842,64815, 64927	KCNRG	DLEU7
lincRNA.2186	1:186422650–186423467	3	64551, 64552, 64553	-	FAT3
lincRNA.2201	1:188331186–188331518	3	64551, 64552, 64553	CTSC	RAB38
lincRNA.2215	1:189167401–189168015	3	64552, 64553, 64555	RAB30	-
lincRNA.2920	11: 6752337–6753098	2	64559, 64560	PAPD5	HEATR3
lincRNA.5494	19: 6298902–6299353	2	64567, 64568	GOSR1	ABR

Discussion

Recently, lncRNAs have received much consideration and a growing number of studies have shown that lncRNAs play key roles in various physiological and pathological processes (Kern et al., 2018; Alexandre et al., 2020). Whereas many of the lncRNAs and their functions are known in different species, such as humans and mouse, lncRNA research is in its infancy in domestic animals, especially in chickens (Zhang et al., 2017). In addition, the role of lncRNAs in the liver of chicken in regulating feed efficiency-related traits is unclear. The most common way of identifying relevant lncRNAs is by differential expression between contrasting conditions (Bakhtiarizadeh and Salami, 2019). Hence, several potential known and novel lncRNAs were identified in the present study by comparing the gene expression profile of the liver of two extremely different chicken breeds (Iranian native chicken vs Ross breed). Comparing the features of the identified novel lncRNAs to protein coding genes showed consistency between our results and the previous studies, as lncRNAs had a lower expression, shorter transcript length (Zhu et al., 2017), fewer exons (Wang et al., 2018) and lower GC content relative to mRNAs (Bakhtiarizadeh et al., 2016; Bakhtiarizadeh and Salami, 2019). These findings indicate that the used bioinformatic pipeline is reliable.

Moreover, in agreement with our previous study, synteny and conservation analysis of the novel lncRNAs emphasized that synteny of lncRNAs are more conserved than their cross-species sequence conservation (Bakhtiarizadeh and Salami, 2019). Therefore, syntenic analysis can be considered as a useful approach for improving the prediction of novel lncRNAs in the genomes that are incompletely annotated. In brief, 1,110 annotated lncRNAs, 525 known lincRNAs, 68 known ilncRNAs, 454 novel lincRNAs and 133 novel ilncRNAs were found. Of these, 38, 1, and 14, known lincRNAs, known ilncRNAs and novel ilncRNAs, respectively, were identified as DEGs. Although differentially expressed genes can be considered as potential candidates related to feed efficiency, however further investigations are needed to be ensured about their potential roles in these biological processes. Hence, different functional analysis were applied to further understand the functions of these genes in regulating feed efficiency in chicken.

Recent studies suggested that the function of lncRNAs can be deduced by analyzing their co-expressed mRNAs or neighboring protein-coding genes in genome wide (Jandura and Krause, 2017). Accordingly, in the cis mode, the identified target genes of the novel lincRNAs were mainly involved in lipid, carbohydrate and growth metabolism, which are related to feed efficiency and make these novel genes as ideal candidates to investigate the regulatory mechanism of feed efficiency in chicken. The relationship between lipid metabolism and feed efficiency has been reported in previous studies, as animals with lower feed efficiency have higher fat deposition and cholesterol levels (Nafikov and Beitz, 2007; Karisa et al., 2014). One of these target genes was ethyl malonyl-CoA decarboxylase (ECHDC1, targeted by lincRNA.103491.1), which is a new metabolite proof reading enzyme. ECHDC1 can eliminate ethyl malonyl-CoA by converting it to butyryl-CoA (Linster et al., 2011). Interestingly, ECHDC1

was a DEG in our study that was down-regulated in Ross breed and showed an opposite expression pattern with its neighboring lincRNA.103491.1. Moreover, these genes were located in the QTL regions associated with feed efficiency, which reinforce their potential function in this subject. These findings make these genes (lincRNA and its target gene) as interesting candidates for a follow-on experiment to assess their impact on feed efficiency in chicken. The other important target gene was VPS13C that was predicted to be targeted by a lincRNA.2454.1. VPS13C, a member of the VPS13 family of proteins (VPS13A, B, C, and D), regulates galectin-12 stability in adipocytes (Yang et al., 2016). In this regard, the important roles of galectin-12 in adipocyte differentiation and lipolysis have been reported (Yang et al., 2016). Also, this gene is suggested to play an important role in glucose homeostasis for high milk production in dairy cow (Lemley et al., 2008). In the present study, higher expression of this gene and the relevant lincRNA in the commercial breed can be considered as their importance in regulating fat and glucose homeostasis, which might be the cause of higher feed efficiency in commercial chicken than the native breed. This finding also emphasized the functional diversity of lncRNAs, which may contribute to widespread regulatory roles in the liver tissue of chicken. The membrane-bound PRLR, as a member of the cytokine receptor family, is closely related to the growth hormone receptor. Over 300 separate biological activities have been attributed to PRL including endocrine signaling, metabolism, control of water and electrolyte balance, growth and development (Bu et al., 2013). PRLR is predicted to be a target gene of lincRNA.15581.1 and its higher expression in chicken can increase feed efficiency by increasing growth rate and decreasing storage energy (Hou et al., 2020).

The genes that were identified as target of differentially expressed known lincRNAs and also were associated with feed efficiency were included B3GALT5, ACAA2, CDCA7L, CHST7, LRRC8D, LRRC8C, LMAN1, CTDSPL, EBP, and AGPAT3. Accordingly, ACAA2 that was predicted to be targeted by lincRNA.15325 is a gene involved in mitochondrial fatty acid oxidation. ACAA2 encodes an enzyme that catalyze the cleavage of 3-ketoacyl CoA to yield acetyl-CoA and acyl-CoA, the final step of the mitochondrial fatty acid beta-oxidation spiral (Abasht et al., 2019). In this study, lincRNA.15325 and its related target gene were up-regulated in the commercial breed. Taking into account its importance in fatty acid oxidation, it might be the cause of higher feed efficiency in commercial chicken than the native breed. It is reported that an increase in the expression of EBP, a key gene functioning in cholesterol biosynthesis, exhibited greater rates of gain and feed efficiency (Connor et al., 2010). This is supported by our result that lincRNA.2988 and its neighboring gene (EBP) showed higher expression in Ross breed and may impact feed digestion. In this context, B3GALT5 (β-1, 3-Galactosyltransferase) was predicted as a cis target gene of lincRNA.1424. Zeng et al. (2017) conducted a study to identify duodenum genes and pathways through transcriptional profiling in two extreme RFI phenotypes of the duck population. B3GALT was DEG between the two duck breeds. B3GALT5 is involved in glycosylation that is a metabolic pathway consisting of the enzymatic modification of proteins and lipids through the stepwise addition of sugars (Trinchera et al., 2014). Therefore, higher expression of lincRNA.1424 and its target gene in higher efficient chicken can increase feed efficiency with enzymatic modification of proteins and lipids. Also, lincRNA.14534, lincRNA.313, lincRNA.15325, lincRNA.5912, lincRNA.6167, lincRNA.1460 and lincRNA.2104 were up-regulated in the liver of high efficient breed (Ross) and their expression were positively correlated with LRRC8D, LRRC8C (Brunes et al., 2021), LMAN1 (Reyer et al., 2017), CTDSPL (Wolc et al., 2013), CDCA7L (Ramayo-Caldas et al., 2019), CHST7 (Dawson et al., 2006) and AGPAT3 (Zarek et al., 2017). Previous studies reported these target genes to be located within the most significant SNPs associated with RFI or feed efficiency.

Sixteen lincRNAs were found to be co-expressed significantly with the MCHR1 gene. This gene is related to chemical synaptic transmission and regulates energy homeostasis and body weight. The function of this gene in reducing feed intake, body weight and body fat have been reported in mice (Zhang et al., 2014). Interestingly, MCHR1 was up-regulated in the commercial breed, which is in agreement with its function to reduce the feed intake and increase feed efficiency. It is worth to note that animals with higher feed efficiency require less energy for metabolism (McKenna et al., 2019; Guinguina et al., 2020). Moreover, 15 lincRNAs were predicted as potential regulators of ADRA2A, in trans mode. Previous studies have shown that activation of ADRA2A, through inhibition of hormone-sensitive lipase, leads to inactivation of the adipocytes lipolysis receptor and prevents adipocyte accumulation (Sawczuk et al., 2013). The other known lincRNA (lincRNA.14916) was found to target GAD2, which is associated with stimulating food intake. LincRNA.14916 and its co-expression gene were up-regulated in the commercial breed. Accordingly, it appears to increased feed efficiency in commercial poultry by reducing maintenance energy (Boutin et al., 2003). As mentioned above, the predicted target genes for the known lincRNAs were similar to the known ilncRNAs, which indicate a

synergistic effect between lincRNAs and ilncRNAs for regulating a common biological process. LincRNA.13441 (as a novel lincRNA) was predicted as a potential regulator of 133 genes. Its target genes were enriched in various biological categories including “calcium ion transport” and “muscle contraction regulation”. It is well documented that calcium signaling is an important modulator of lipid metabolism (Xue et al., 2001). Hence, the function of these lincRNAs (such as 537 ADRA2A and MCHR1) could be closely related to lipid metabolism as well as feed efficiency development due to their co-expressed targeted mRNAs in the commercial breed affect lipid metabolism To better understand how lncRNAs cooperate with their target genes, integrated networks (separately for each of lincRNAs and ilncRNAs) were constructed and one significant module was found in each of the networks. Interestingly, member genes of both modules were significantly enriched in functional categories related to feed efficiency including “lipid phosphorylation”, “glucose homeostasis”, “carbohydrate homeostasis”, “insulin receptor signaling pathway” and “calcium ion transport”. Energy metabolism is an important factor affecting the feed efficiency of livestock and poultry (Kaewpila et al., 2018). ADAR2A encodes adrenoceptor alpha 2A and is a regulator of catecholamines, which have been introduced to be associated with energy metabolism and fat metabolism. Catecholamine-stimulated whole body lipolysis and lipolysis in subcutaneous adipocytes are blunted in obesity (Blaak et al., 1994), thereby limiting lipid mobilization and favoring fat accumulation. Notably, ADAR2A was predicted to be target of 24 lncRNAs (lincRNA and ilncRNA) that might be regarded as key regulators. On the other hand, the other members of these module have been reported to be involved in calcium signaling pathways including MCHR1 (Pissios et al., 2003), P2RY1 (Choi et al., 2001) and F2R (Marchesi et al., 2019) (Fig. 6–7). According to a previous study calcium participated in the expression of genes related to lipid metabolism and prohibition of fat deposition (Cao et al., 2017).

QTL analysis revealed that 11 novel ilncRNAs and 47 novel lincRNAs as putative effective lncRNAs in feed efficiency related processes, as they were overlapped with the potential regions associated with RFI in genome wide. Of these, the predicted target genes of the four lincRNAs including lincRNA.10349 (target ECHDC1) (Linster et al., 2011), lincRNA.10336 (target SMAD3) (Yadav et al., 2011), lincRNA.2986 (target EBP) (Connor et al., 2010) and lincRNA.10372 (target TPD52L1) (Kamili et al., 2015) were related to lipid metabolism. For example, SMAD3 is a multifaceted regulator in adipose physiology, pathogenesis of obesity and type 2 diabetes (Yadav et al., 2011). As discussed above, animals with lower feed efficiency have higher fat deposition and cholesterol levels (Nafikov and Beitz, 2007; Karisa et al., 2014). In addition, six lincRNAs (linRNA.1964, linRNA.2201, linRNA.2215, linRNA.5649, linRNA.2920, linRNA.2178) were located in more than one QTL related feed efficiency. CTSC, a predicted cis target gene of lincRNA.2201, encodes a lysosomal cysteine proteinase and play a central role in bacterial killing and immune regulation in T lymphocytes. The effects of this gene on the observed difference between the pigs with low and high RFI have been reported (Gondret et al., 2017). The closest protein coding gene to lincRNA.1964 was DLEU7, which encodes a protein containing 221 amino acids. Fibroblast growth factor (FGF) regulates the expression of DLEU7 during early embryogenesis (Zhu et al., 2012). Moreover, association of this gene with human height has been reported (Weedon et al., 2008; Sovio et al., 2009; Kang et al., 2010), which can be suggested to play an important role in chicken growth and feed efficiency. The other gene related to growth rate was RAB30 (cis target of lincRNA.2215). Claire D’Andre et al. (2013) conducted a study on the identification and characterization of genes that control fat deposition in chicken. RAB30 was appeared to be down-regulated in slow-growing Xinghua chickens. lincRNAs that were located in QTL regions related to feed efficiency compared with other lincRNAs are more likely to be truly related to feed efficiency. Cis target genes of these lincRNAs that were involved in feed efficiency and RFI or has a known role in the growth rate and lipid metabolism, were interesting functional candidate genes responsible for feed efficiency in chicken. Our findings supported this hypothesis that lncRNAs may affect feed efficiency mainly through regulating lipid metabolism, glucose homeostasis, growth rate, immune system, modification of proteins and energy homeostasis. However further experiments still require to validate the suggested functions of these lncRNAs.

Conclusion

The present study provided a valuable resource to clarify the genetic basis of feed efficiency and further experiments will corroborate the function of all the RNAs the reported here. To do this end, RNA-Seq data along with bioinformatics approaches were applied and a series of lncRNAs and target genes were identified. Our results indicated that some of the lncRNAs might regulate locally their neighboring genes in cis mode as well as in trans mode. Functional enrichment analysis showed that the identified lncRNAs had enough potential to be related to feed efficiency, as their predicted target genes were significantly

involved in the biological processes and KEGG pathways associated with feed efficiency including lipid metabolism, carbohydrate metabolism, and growth. Moreover, several lncRNAs were identified that were overlapped with QTLs controlling feed efficiency such as RFI, which highlighted their importance in this context. Although little is known about the functions of lncRNAs in chicken, our results provided the initial step for studying how changes in lncRNA expression affect the regulation of mechanisms involved in chicken feed efficiency.

Declarations

Conflict of Interests

The authors declare no conflict of interests regarding the publication of this paper.

Author contributions

M.R.B conceived the ideas. M.R.B and P.K designed study and analyzed the data. M.R.B, P.K, A.S and H.I interpreted the data and wrote the main manuscript text. All authors read and approved the final manuscript.

Ethics declarations

All methods were carried out in accordance with relevant guidelines and regulations. All experimental protocols were approved by the Animal Care and Use Committee at the Animal Science Research Department, Safiabad Agricultural.

References

1. Abasht, B., Zhou, N., Lee, W. R., Zhuo, Z., & Peripolli, E. The metabolic characteristics of susceptibility to wooden breast disease in chickens with high feed efficiency. *Poultry science*, **98**(8), 3246-3256. (2019).
2. Alexandre, P. A. *et al*. "Exploring the Regulatory Potential of Long Non-Coding RNA in Feed Efficiency of Indicine Cattle." *Genes* **11**(9), 997. (2020).
3. Alsaffar, A. A., & Khalil, F. Livestock Management LM-589 Why poultry welfare in Kuwait is an obstacle to trade?. *Sustainable Animal Agriculture for Developing Countries*, 711. (2015).
4. Bakhtiarizadeh, M. R., Hosseinpour, B., Arefnezhad, B., Shamabadi, N., & Salami, S. A. In silico prediction of long intergenic non-coding RNAs in sheep. *Genome*, **59**(4), 263-275. (2016).
5. Bakhtiarizadeh, M. R., & Salami, S. A. Identification and expression analysis of long noncoding RNAs in fat-tail of sheep breeds. *G3: Genes, Genomes, Genetics*, **9**(4), 1263-1276. (2019).
6. Bateman, A. *et al*. "The Pfam protein families database." *Nucleic acids research* **32(**suppl_1),.D138-D141. doi: 10.1093/nar/gkh12 (2004).
7. Blaak, E. E. *et al*. "Beta-adrenergic stimulation of energy expenditure and forearm skeletal muscle metabolism in lean and obese men." *American Journal of Physiology-Endocrinology And Metabolism* **267** (2), E306-E315. doi: 10.1152/ajpendo.1994.267.2.E306 (1994).
8. Bolger, A. M., Lohse, M., & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**(15), 2114-2120. doi: 10.1093/bioinformatics/btu170 (2014).
9. Boutin, P. *et al.* "GAD2 on chromosome 10p12 is a candidate gene for human obesity." *PLoS biology* **1**(3), e68. doi: 10.1371/journal.pbio.0000068 (2003).
10. Brunes, L.C. *et al*. "Weighted single-step genome-wide association study and pathway analyses for feed efficiency traits in Nellore cattle." *Journal of Animal Breeding and Genetics* **138**(1), 23-44. doi: 10.1111/jbg.12496 (2021).
11. Bu, G. *et al*. "Characterization of the novel duplicated PRLR gene at the late-feathering K locus in Lohmann chickens." *Journal of molecular endocrinology* **51**(2), 261-276. doi: 10.1530/JME-13-0068 (2013).
12. Cao, C. *et al*. "Impact of exudative diathesis induced by selenium deficiency on LncRNAs and their roles in the oxidative reduction process in broiler chick veins." *Oncotarget* **8**(13), 20695. doi: 10.18632/oncotarget.14971(2017).

13. Choi, R.C. *et al*. "Expression of the P2Y1 nucleotide receptor in chick muscle: its functional role in the regulation of acetylcholinesterase and acetylcholine receptor." *Journal of Neuroscience* **21**(23) 9224-9234. doi: 10.1523/JNEUROSCI.21-23-09224 (2001).
14. D'Andre, H.C. *et al*. "Identification and characterization of genes that control fat deposition in chickens." *Journal of animal science and biotechnology* **4**(1), 1-16. doi: 10.1186/2049-1891-4-43 (2013).
15. Connor, E.E. *et al*. "Enhanced mitochondrial complex gene function and reduced liver size may mediate improved feed efficiency of beef cattle during compensatory growth." *Functional & integrative genomics* **10**(1), 39-51. doi: 10.1007/s10142-009-0138-7 (2010).
16. Dawson, P. A., Gardiner, B., Grimmond, S., & Markovich, D. Transcriptional profile reveals altered hepatic lipid and cholesterol metabolism in hyposulfatemic NaS1 null mice. *Physiological genomics*, **26**(2), 116-124. doi: 10.1093/bioinformatics/bts635 (2006).
17. Foissac, S. *et al*. "Livestock genome annotation: transcriptome and chromatin structure profiling in cattle, goat, chicken and pig." *bioRxiv*. 316091. doi: 10.1101/316091 (2018).
18. Gondret, F. *et al*. "A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs." *BMC genomics* **18**(1), 1-17. doi: 10.1186/s12864-017-3639-0 (2017).
19. Guinguina, A., Yan, T., Bayat, A. R., Lund, P., & Huhtanen, P. The effects of energy metabolism variables on feed efficiency in respiration chamber studies with lactating dairy cows. *Journal of Dairy Science*, **103**(9), 7983-7997. doi: https://doi.org/10.3168/jds.2020-18259 (2020).
20. Herd, R. M., & Arthur, P. F. Physiological basis for residual feed intake. *Journal of animal science*, **87**(suppl_14), E64-E71. doi: https://doi.org/10.2527/jas.2008-1345 (2009).
21. Hezroni, H. *et al*. "Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species." *Cell reports* **11**(7), 1110-1122. doi: https://doi.org/10.1016/j.celrep.2015.04.023 (2015).
22. Hou, X. *et al*. "Transcriptome analysis of skeletal muscle in pigs with divergent residual feed intake phenotypes." *DNA and cell biology* **39**(3), 404-416. doi: 10.1089/dna.2019.4878 (2020).
23. Hu, Z. L., Park, C. A., & Reecy, J. M. Developmental progress and current status of the Animal QTLdb. *Nucleic acids research*, **44**(D1), D827-D833. doi: 10.1093/nar/gkv1233 (2016).
24. Iyer, M.K. *et al*. "The landscape of long noncoding RNAs in the human transcriptome." *Nature genetics* **47**(3) 199-208. doi: 10.1038/ng.3192 (2015).
25. Izadnia, H. R., Tahmoorespur, M., Bakhtiarizadeh, M. R., Nassiri, M., & Esmaeilkhanien, S. Gene expression profile analysis of residual feed intake for Isfahan native chickens using RNA-SEQ data. *Italian Journal of Animal Science*, **18**(1), 246-260. doi: 10.1080/1828051X.2018.1507625 (2019).
26. Jandura, A., & Krause, H. M. The new RNA world: growing evidence for long noncoding RNA functionality. *Trends in Genetics*, **33**(10), 665-676. doi: https://doi.org/10.1016/j.tig.2017.08.002 (2017).
27. Kaewpila, C., Sommart, K., & Mitsumori, M. Dietary fat sources affect feed intake, digestibility, rumen microbial populations, energy partition and methane emissions in different beef cattle genotypes. *animal*, **12**(12), 2529-2538. doi: https://doi.org/10.1017/S1751731118000587 (2018).
28. Kamili, A. *et al*. "TPD52 expression increases neutral lipid storage within cultured cells." *Journal of cell science* **128**(17), 3223-3238. doi: 10.1242/jcs.167692 (2015).
29. Kang, S.J. *et al*. "Genome-wide association of anthropometric traits in African-and African-derived populations." *Human molecular genetics* **19**(13), 2725-2738. doi: 10.1093/hmg/ddq154 (2010).
30. Kang, Y.J. *et al*. "CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features." *Nucleic acids research* **45**(W1), W12-W16. doi: 10.1093/nar/gkx428 (2017).
31. Karisa, B., Moore, S., & Plastow, G. Analysis of biological networks and biological pathways associated with residual feed intake in beef cattle. *Animal Science Journal*, **85**(4), 374-387. doi: https://doi.org/10.1111/asj.12159 (2014).
32. Kern, C. *et al*. "Genome-wide identification of tissue-specific long non-coding RNA in three farm animal species." *BMC genomics* **19**(1), 1-14. doi: 10.1186/s12864-018-5037-7 (2018).

33. Korostowski, L., Sedlak, N., & Engel, N. The Kcnq1ot1 long non-coding RNA affects chromatin conformation and expression of Kcnq1, but does not regulate its imprinting in the developing heart. *PLoS Genet* **8**(9)**,** e1002956. doi: 10.1371/journal.pgen.1002956 (2012).
34. Kuang, J., Yan, X., Genders, A. J., Granata, C., & Bishop, D. J. An overview of technical considerations when using quantitative real-time PCR analysis of gene expression in human exercise research. *PloS one*, **13**(5), e0196438. doi: 10.1371/journal.pone.0196438 (2018).
35. Kuleshov, M.V. *et al*. "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update." *Nucleic acids research* **44**(W1), W90-W97. doi: 10.1093/nar/gkw377 (2016).
36. Kuo, Richard I., *et al*. "Normalized long read RNA sequencing in chicken reveals transcriptome complexity similar to human." *BMC genomics* **18**(1), 1-19. doi: 10.1186/s12864-017-3691-9 (2017).
37. Lemley, C. O., Butler, S. T., Butler, W. R., & Wilson, M. E. Insulin alters hepatic progesterone catabolic enzymes cytochrome P450 2C and 3A in dairy cows. *Journal of dairy science*, **91**(2), 641-645. doi: https://doi.org/10.3168/jds.2007-0636 (2008).
38. Li, A., Zhang, J., & Zhou, Z. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC bioinformatics*, **15**(1), 1-10. doi: 10.1186/1471-2105-15-311(2014).
39. Li, H. *et al*. "Transcriptome profile analysis reveals an estrogen induced LncRNA associated with lipid metabolism and carcass traits in chickens (Gallus Gallus)." *Cellular Physiology and Biochemistry* **50**(5), 1638-1658. doi: 10.1159/000494785 (2018).
40. Li, W. *et al*. "Analysis of four complete linkage sequence variants within a novel lncRNA located in a growth QTL on chromosome 1 related to growth traits in chickens." *Journal of animal science* **98**(5), skaa122. doi: 10.1093/jas/skaa122 (2020).
41. Li, Z. *et al*. "Integrated analysis of long non-coding RNAs (LncRNAs) and mRNA expression profiles reveals the potential role of LncRNAs in skeletal muscle development of the chicken." *Frontiers in physiology* **7,** 687. doi: 10.3389/fphys.2016.00687 (2017).
42. Linster, C.L. *et al*. "Ethylmalonyl-CoA decarboxylase, a new enzyme involved in metabolite proofreading." *Journal of Biological Chemistry* **286**(50), 42992-43003. doi: https://doi.org/10.1074/jbc.M111.281527 (2011).
43. Luiting, P., Schrama, J. W., Van der Hel, W., & Urff, E. M. Metabolic differences between White Legho-rns selected for high and low residual food consumption. *British Poultry Science*, **32**(4), 763-782. doi: https://doi.org/10.1080/00071669108417402 (1991).
44. Liu, H., Wang, R., Mao, B., Zhao, B., & Wang, J. Identification of lncRNAs involved in rice ovule development and female gametophyte abortion by genome-wide screening and functional analysis. *BMC genomics*, **20**(1), 1-16. doi: 10.1186/s12864-019-5442-6 (2019).
45. Liu, W. *et al*. "A genome-wide SNP scan reveals novel loci for egg production and quality traits in white leghorn and brown-egg dwarf layers." *PloS one* **6**(12), e28600. doi: 10.1371/journal.pone.0028600 (2011).
46. Liu, Y. *et al*. "Analyses of long non-coding RNA and mRNA profiling using RNA sequencing in chicken testis with extreme sperm motility." *Scientific reports* **7**(1), 1-8. doi: 10.1038/s41598-017-08738-9 (2017).
47. Marchesi, J. A. P.*et al*. "Whole transcriptome analysis of the pectoralis major muscle reveals molecular mechanisms involved with white striping in broiler chickens." *Poultry science* **98**(2), 590-601. doi: https://doi.org/10.3382/ps/pey429 (2019).
48. McKenna, C. *et al*. "Residual feed intake phenotype and gender affect the expression of key genes of the lipogenesis pathway in subcutaneous adipose tissue of beef cattle." *J. Anim. Sci. Biotechnol* **10** 212-221. doi: 10.1186/ s40104-018-0300-y (2019).
49. Meale, S.J. *et al*. "Exploration of biological markers of feed efficiency in young bulls." *Journal of agricultural and food chemistry* **65**(45), 9817-9827. doi: https://doi.org/10.1021/acs.jafc.7b03503 (2017).
50. Mering, C.V. *et al*. "STRING: a database of predicted functional associations between proteins." *Nucleic acids research* **31**(1), 258-261. doi: 10.1093/nar/gkg034 (2003).

51. Meseure, D. *et al*. "Prognostic value of a newly identified MALAT1 alternatively spliced transcript in breast cancer." *British journal of cancer* **114**(12), 1395-1404. doi: 10.1038/bjc.2016.123 (2016).
52. Muret, K. *et al*. "Long noncoding RNAs in lipid metabolism: literature review and conservation analysis across species." *BMC genomics* **20**(1), 1-18. doi: 10.1186/s12864-019-6093-3 (2019).
53. Muret, K. *et al*. "Long noncoding RNA repertoire in chicken liver and adipose tissue." *Genetics Selection Evolution* **49**(1), 1-17. doi: 10.1186/s12711-016-0275-0 (2017).
54. Nafikov, R. A., & Beitz, D. C. Carbohydrate and lipid metabolism in farm animals. The Journal of nutrition, **137**(3), 702-705. doi: 10.1093/jn/137.3.702 (2007).
55. Nepusz, T., Yu, H., & Paccanaro, A. Detecting overlapping protein complexes in protein-protein interaction networks. *Nature methods*, **9**(5), 471-472. doi: 10.1038/nmeth.1938 (2012).
56. Ning, C. *et al*. "Long Non-coding RNA and mRNA profile of liver tissue during four developmental stages in the chicken." *Frontiers in genetics* **11**, 574. doi: 10.3389/fgene.2020.00574 (2020).
57. Pertea, M. *et al*. "StringTie enables improved reconstruction of a transcriptome from RNA-seq reads." *Nature biotechnology* **33**(3), 290-295. doi: 10.1038/nbt.3122 (2015).
58. Pissios, P., Trombly, D. J., Tzameli, I., & Maratos-Flier, E. Melanin-concentrating hormone receptor 1 activates extracellular signal-regulated kinase and synergizes with Gs-coupled pathways. *Endocrinology*, **144**(8), 3514-3523. doi: 10.1210/en.2002-0004 (2003).
59. Qiu, L. *et al*. "Discovery of novel long non-coding RNAs induced by subgroup J avian leukosis virus infection in chicken." *Developmental & Comparative Immunology* **76,** 292-302. doi: https://doi.org/10.1016/j.dci.2017.06.015 (2017).
60. Ramayo-Caldas, Y. *et al*. "Integrating genome-wide co-association and gene expression to identify putative regulators and predictors of feed efficiency in pigs." *Genetics Selection Evolution* **51**(1) 1-17. doi: 10.1186/s12711-019-0490-6 (2019).
61. Rekaya, R., Sapp, R. L., Wing, T., & Aggrey, S. E. Genetic evaluation for growth, body composition, feed efficiency, and leg soundness. *Poultry science*, **92**(4), 923-929. doi: doi.org/10.3382/ps.2012-02649 (2013).
62. Ren, T. *et al*. "Sequencing and characterization of lncRNAs in the breast muscle of Gushi and Arbor Acres chickens." *Genome* **61**(5), 337-347. doi: 10.1139/gen-2017-0114 (2018).
63. Reyer, H. *et al*. "Strategies towards improved feed efficiency in pigs comprise molecular shifts in hepatic lipid and carbohydrate metabolism." *International journal of molecular sciences* **18**(8). 1674. doi: 10.3390/ijms18081674 (2017).
64. Sawczuk, M., Maciejewska-Karłowska, A., & Cięszczyk, P. A single nucleotide polymorphism rs553668 in the ADRA2A gene and the status of Polish elite endurance athletes. (2013).
65. Shannon, P. *et al*. "Cytoscape: a software environment for integrated models of biomolecular interaction networks." *Genome research* **13**(11), 2498-2504. doi: 10.1101/gr.1239303 (2003).
66. Sovio, U. *et al*. "Genetic determinants of height growth assessed longitudinally from infancy to adulthood in the northern Finland birth cohort 1966." *PLoS genetics* **5**(3), e1000409. doi: 10.1371/journal.pgen.1000409 (2009).
67. Sui, Y., Han, Y., Zhao, X., Li, D., & Li, G. Long non-coding RNA Irm enhances myogenic differentiation by interacting with MEF2D. *Cell death & disease,* **10**(3), 1-13. doi: 10.1038/s41419-019-1399-2 (2019).
68. Sun, L. *et al*. "Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts." *Nucleic acids research* **41**(17), e166-e166. doi: 10.1093/nar/gkt646 (2013).
69. Tang, R. *et al*. "Comprehensive analysis of lncRNA and mRNA expression changes in Tibetan chicken lung tissue between three developmental stages." *Animal Genetics* **51**(5), 731-740. doi: https://doi.org/10.1111/age.12990 (2020).
70. Trapnell, C. *et al*. "Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation." *Nature biotechnology* **28**(5), 511-515. doi: 10.1038/nbt.1621(2010).
71. Trinchera, M., Zulueta, A., Caretti, A., & Dall'Olio, F. Control of glycosylation-related genes by DNA methylation: the intriguing case of the B3GALT5 gene and its distinct promoters. *Biology*, **3**(3), 484-497. doi: 10.3390/biology3030484 (2014).
72. Ulitsky, I. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nature Reviews Genetics*, **17**(10), 601-614. doi: 10.1038/nrg.2016.85 (2016).

73. Untergasser, A. *et al*. "Primer3Plus, an enhanced web interface to Primer3." *Nucleic acids research* 35 (suppl_2) W71-W74. doi: 10.1-093/nar/gkm306 (2007).
74. Vance, K. W., & Ponting, C. P. Transcriptional regulatory functions of nuclear long noncoding RNAs.*Trends in Genetics*, **30**(8), 348-355. doi: https://doi.org/10.1016/j.tig.2014.06.001 (2014).
75. Wang, L. *et al*. "CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model." *Nucleic acids research* **41**(6), e74-e74. doi: 10.1093/nar/gkt006 (2013).
76. Wang, Li. *et al*. "Genome-wide identification and characterization of novel lncRNAs in Ginkgo biloba." *Trees* **32**(5), 1429-1442. doi: 10.1007/s00468-018-1724-x (2018).
77. Weedon, M. N. *et al*. "Genome-wide association analysis identifies 20 loci that influence adult height." *Nature genetics* **40**(5), 575-583. doi: 10.1038/ng.121 (2008).
78. Wolc, A. *et al*. "Pedigree and genomic analyses of feed consumption and residual feed intake in laying hens." *Poultry science* **92**(9), 2270-2275. doi: https://doi.org/10.3382/ps.2013-03085 (2013).
79. Wucher, V. *et al*. "FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome." *Nucleic acids research* **45**(8), e57-e57. doi: 10.1093/nar/gkw1306 (2017).
80. Xue, B., Greenberg, A.G., Kraemer, F.B., & Zemel, M.B. Mechanism of intracellular calcium ([Ca2+]i) inhibition of lipolysis in human adipocytes. *The FASEB Journal* **15**(13)**,** 2527-2529. doi: https://doi.org/10.1096/fj.01-0278fje (2001).
81. Yadav, H. *et al*. "Protection from obesity and diabetes by blockade of TGF-β/Smad3 signaling." *Cell metabolism* **14**(1), 67-79. doi: https://doi.org/10.1016/j.omtn.2019.08.003 (2011).
82. Yang, L. *et al*. "Identification of key genes and pathways associated with feed efficiency of native chickens based on transcriptome data via bioinformatics analysis." *BMC genomics* **21**(1), 1-18.doi: https://doi.org/10.1186/s12864-020-6713-y (2020).
83. Yang, R. Y. *et al*. "Identification of VPS13C as a galectin-12-binding protein that regulates galectin-12 protein stability and adipogenesis." *PloS one* **11**(4), e0153534. doi: 10.1371/journal.pone.015353(2016).
84. Yousefi Zonuz, A., Alijani, S., Mohammadi, H., Rafat, A., & Daghigh Kia, H. Estimation of genetic parameters for productive and reproductive traits in Esfahan native chickens. *Journal of Livestock Science and Technologies*, **1**(2), 34-38. doi: 10.22103/JLST.2013.550 (2013).
85. Zarek, C. M., Lindholm-Perry, A. K., Kuehn, L. A., & Freetly, H. C. Differential expression of genes related to gain and intake in the liver of beef cattle. *BMC research notes*, **10**(1), 1-8. doi: 10.1186/s13104-016-2345-3(2017).
86. Zeng, T. *et al*. "Gene expression profiling reveals candidate genes related to residual feed intake in duodenum of laying ducks." *Journal of animal science* **95**(12), 5270-5277. doi: 10.2527/jas2017.1714 (2017).
87. Zhang, L.N. *et al*. "Effects of a specific MCHR1 antagonist (GW803430) on energy budget and glucose metabolism in diet-induced obese mice." *Obesity* **22**(3) 681-690. doi: https://doi.org/10.1002/oby.20418 (2014).
88. Zhang, T. *et al*. "Analysis of long noncoding RNA and mRNA using RNA sequencing during the differentiation of intramuscular preadipocytes in chicken." *PloS one* **12**(2). e0172389. doi: 10.1371/journal.pone.0172389 (2017).
89. Zhu, B., Xu, M., Shi, H., Gao, X., & Liang, P. Genome-wide identification of lncRNAs associated with chlorantraniliprole resistance in diamondback moth Plutella xylostella (L.). *BMC genomics*, **18**(1), 1-11. doi: 10.1186/s12864-017-3748-9 (2017).
90. Zhu, X. *et al*. "Characterization and expressional analysis of Dleu7 during Xenopus tropicalis embryogenesis." *Gene* **509**(1)**,** 77-84. doi: https://doi.org/10.1016/j.gene.2012.08.024 (2012).

Figures

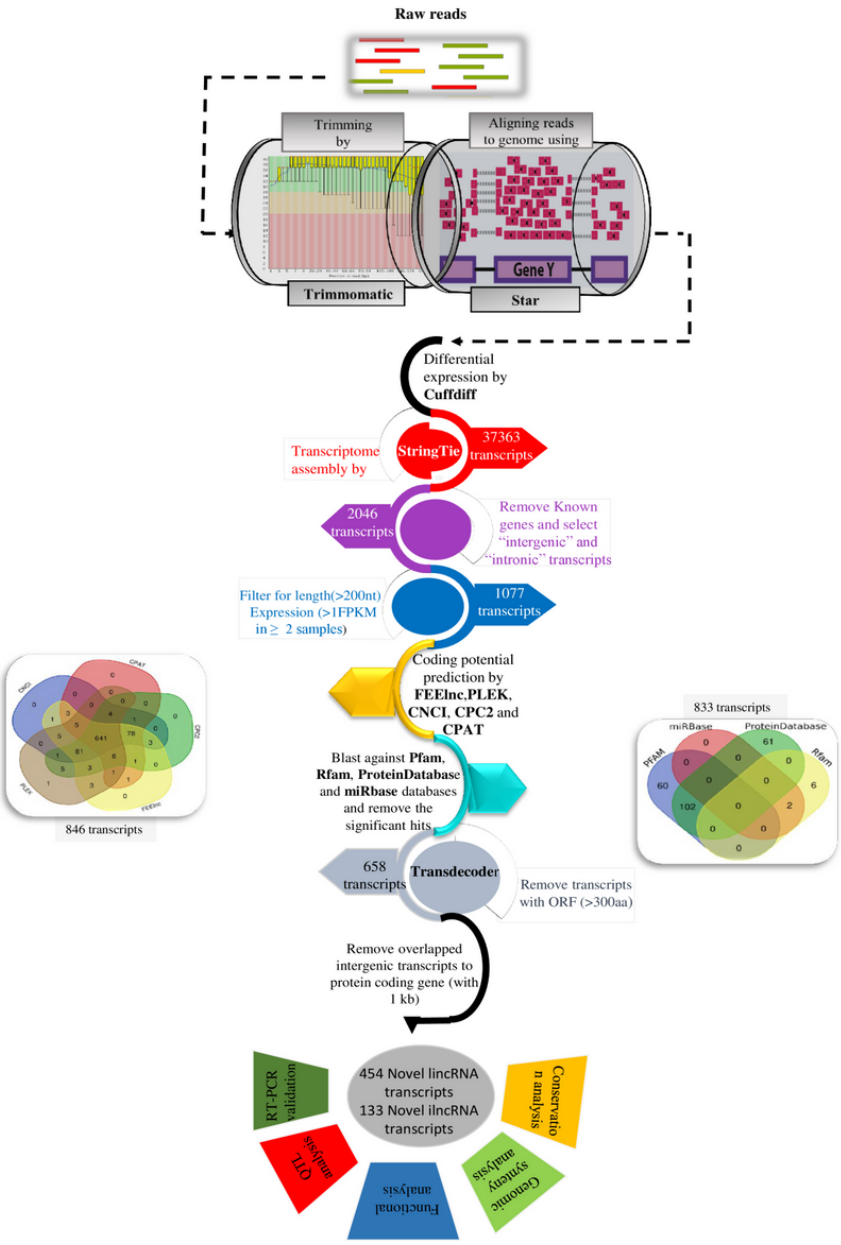


Figure 1

The bioinformatics pipeline for identifying annotated, known and novel lncRNAs. The right and left Venn diagrams illustrate the results of blasting the transcripts against four different databases and the potential coding ability of the transcript using five software, respectively.

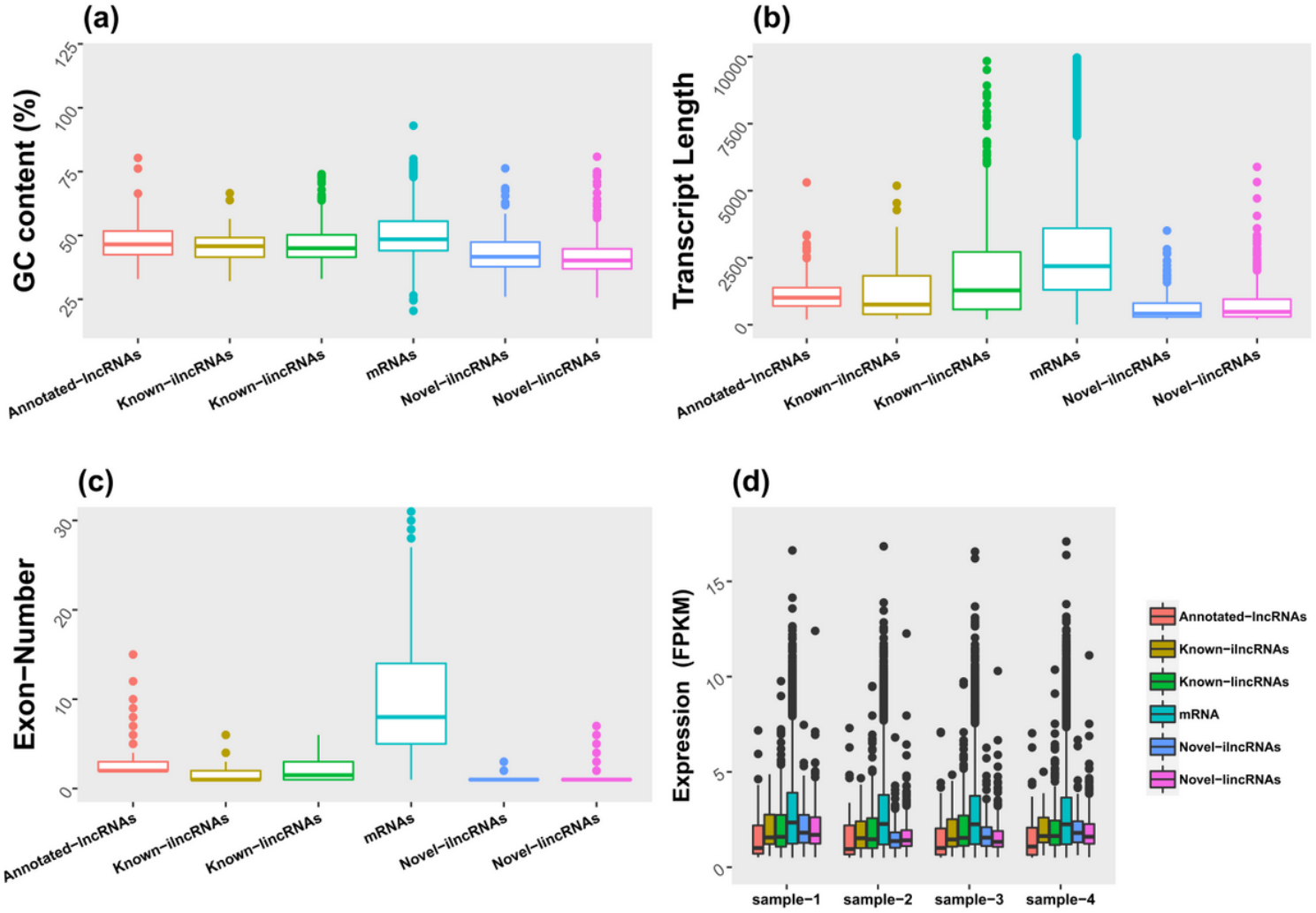


Figure 2

Comparison of (a) GC contents, (b) transcript length , (c) exon number and (d) gene expression levels of the annotated, novel and known lncRNAs with mRNAs.

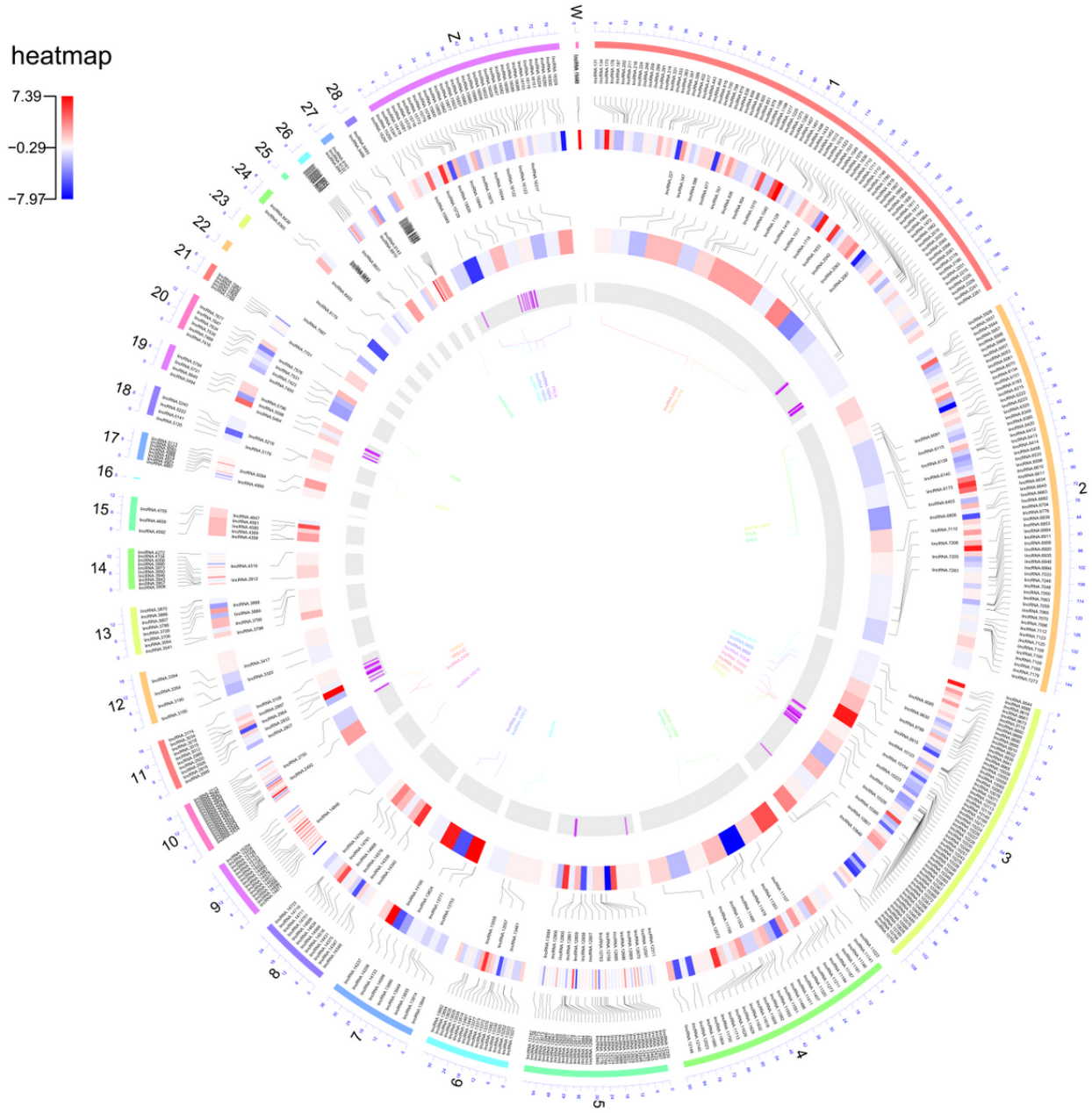
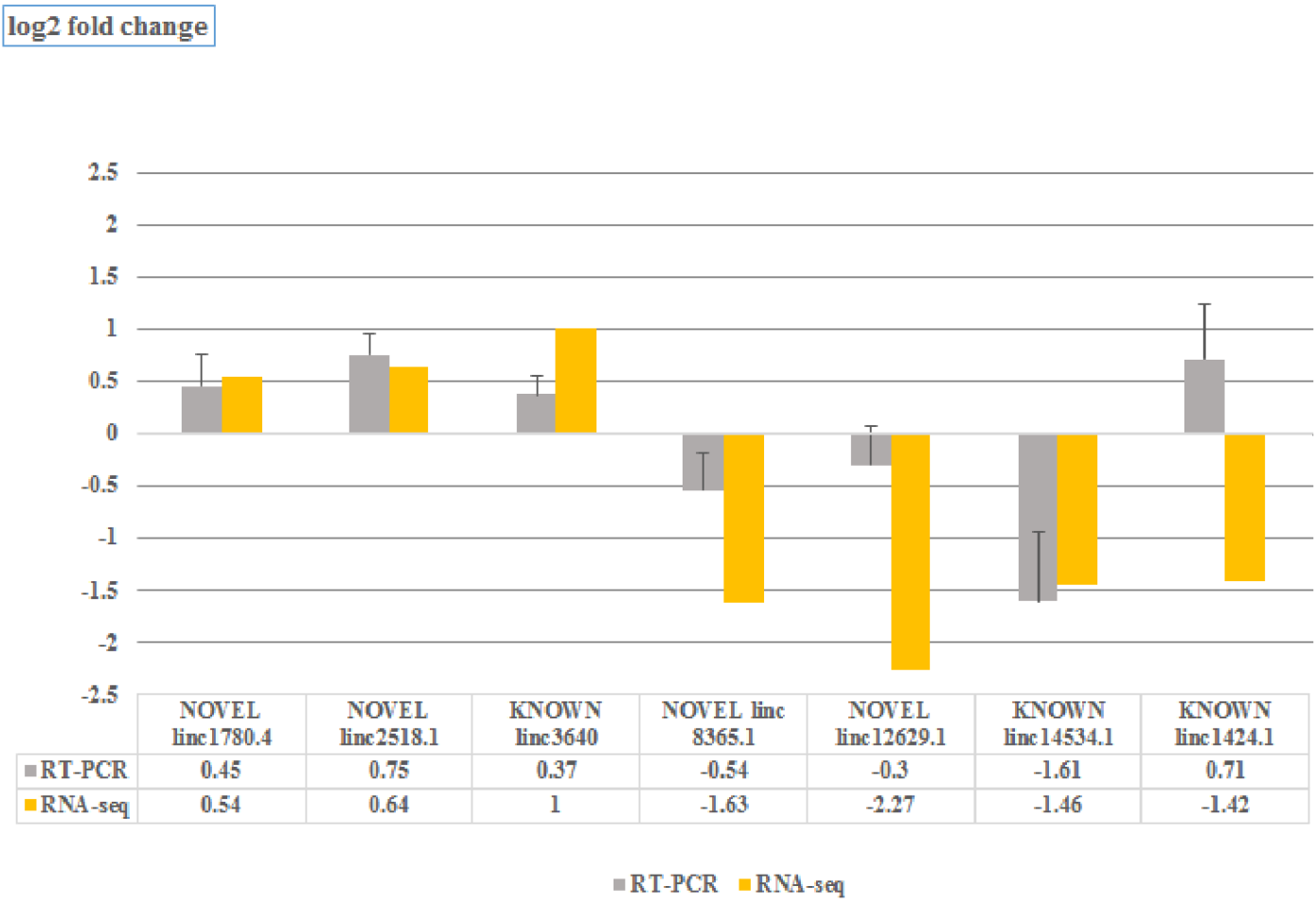


Figure 3

Circos plot represents the genome-wide distribution density of all the identified novel lncRNAs (in clockwise order). From outside to inside the plot shows the chromosomes, and the next ring the position of the novel lincRNAs, the log2 fold change (Esfahani against Ross) of the novel lincRNAs, position of the novel ilncRNAs, the log2 fold change (Esfahani against Ross) of the novel ilncRNAs, the location of Residual feed intake QTLs from animal QTL database and the names of promising lncRNAs and mRNAs pairs associate with Residual feed intake. Red heatmap colors represent the higher expression of the gene in Esfahan than Ross breed.



	NOVEL linc1780.4	NOVEL linc2518.1	KNOWN linc3640	NOVEL linc 8365.1	NOVEL linc12629.1	KNOWN linc14534.1	KNOWN linc1424.1
RT-PCR	0.45	0.75	0.37	-0.54	-0.3	-1.61	0.71
RNA-seq	0.54	0.64	1	-1.63	-2.27	-1.46	-1.42

Figure 4

Validation of seven randomly selected novel lincRNAs by RT-qPCR. The fold change represents the ratio of average expression of native samples relative to that of commercial breed.

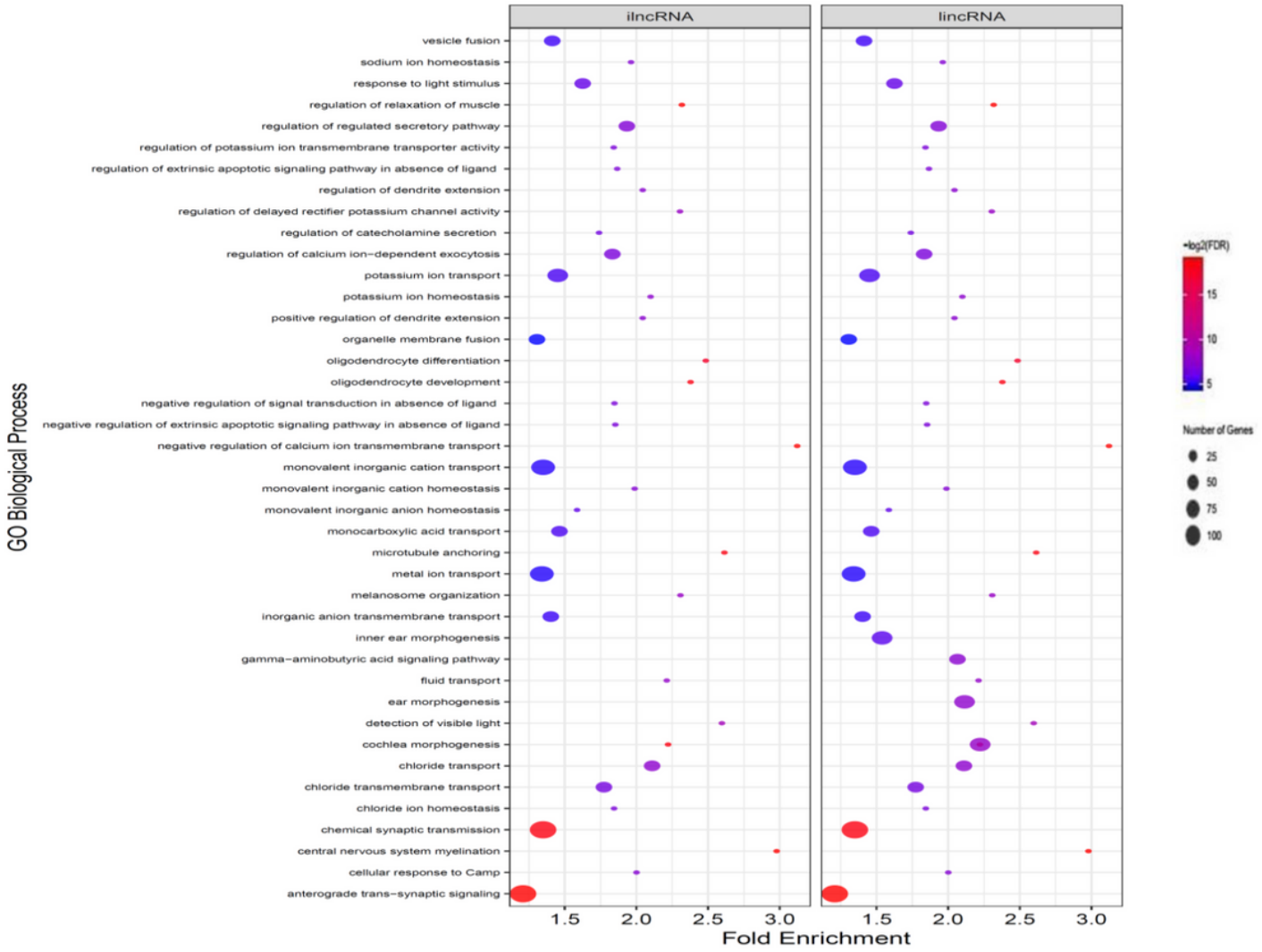


Figure 5

Functional enrichment analysis of the trans target genes of predicted known lincRNAs and ilncRNAs.

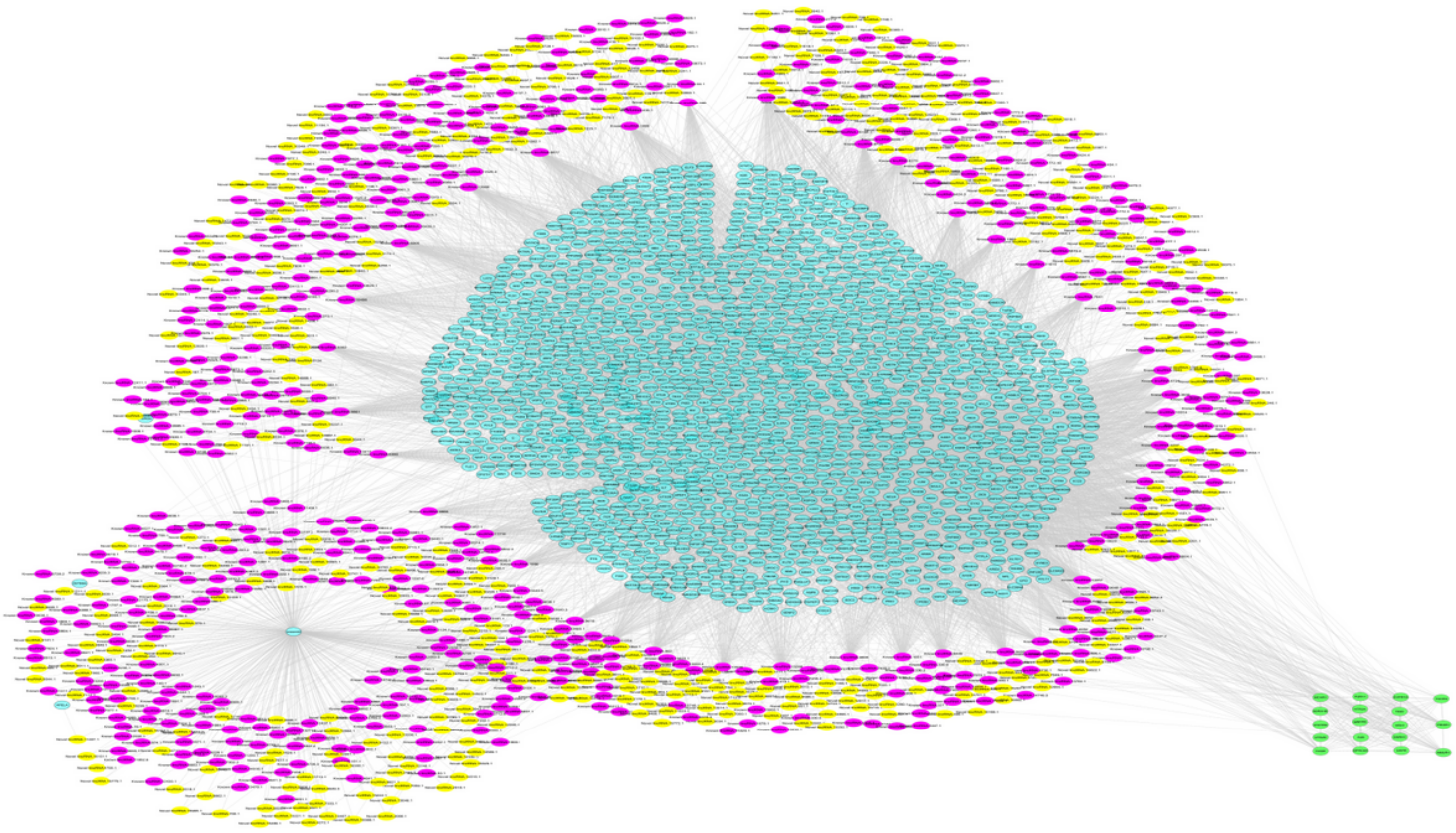


Figure 6

Final constructed network based on the known and novel lincRNAs. Purple nodes: Known lincRNAs, yellow nodes: Novel lincRNAs, Sky blue nodes: protein coding genes, green nodes: predicted module (P-value=0.007).

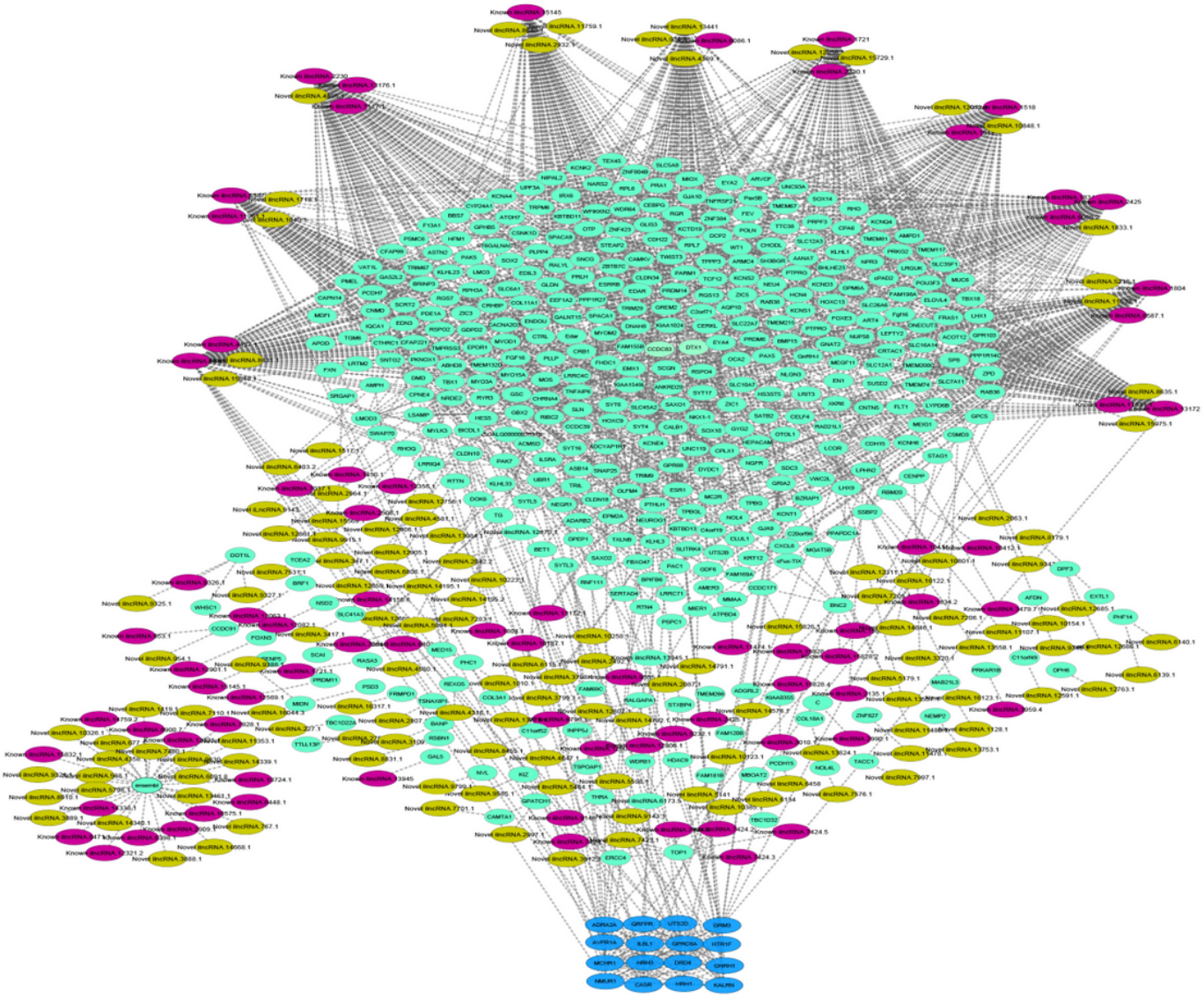


Figure 7

Final constructed network based on the known and novel ilncRNAs. Purple nodes: Known ilncRNAs, Light Olive nodes: Novel ilncRNAs, Medium Spring Green nodes: protein coding genes, light-blue nodes: predicted module (P-value=6.147-e7).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- S1.xlsx
- S2.xlsx
- S3.xlsx
- S4.xlsx
- S5.xlsx
- S6.xlsx
- S7.xlsx

- S8.xlsx