Screening and validation of candidate genes involved in the regulation of egg yolk deposition in chicken

Junnan Zhang
China Agricultural University
Zhongyi Duan
China Agricultural University
Xiqiong Wang
China Agricultural University
Fengning Li
China Agricultural University
Jiajing Chen
China Agricultural University
Xingfu Lai
South China Agricultural University College of Economics and Management
Liang Qu
China Agricultural University
Congjiao Sun
China Agricultural University
Guiyun Xu (ncppt@cau.edu.cn)
China Agricultural University https://orcid.org/0000-0002-3301-9221

Research article

Keywords: RNA-seq, Ovary, Percentage of Yolk, QTL

Posted Date: January 28th, 2020

DOI: https://doi.org/10.21203/rs.2.22106/v1

Version of Record: A version of this preprint was published at Poultry Science on June 1st, 2021. See the published version at https://doi.org/10.1016/j.psj.2021.101077.

Abstract

Background: The eggs with the same egg weight but may have wide differences in yolk weight. Eggs with a high proportion of yolk have richer nutritional value, more flavor and taste, and are more favourite by parts of consumers. However, large proportion of yolk means relatively much dry matters in eggs, which reduces the feed efficiency to some extent. To elucidation the genetic factors for the amounts of yolk in eggs is of scientific and practical significance.

Results: Through RNA sequencing, we explored the transcriptome of ovarian tissue from twelve Wenchang chickens which includes six eggs with a high percentage of yolk (32%) and six low-percent eggs (25%).We identified 362 differential genes (P<0.01, log2foldchange<-1, log2foldchange>1), of which 220 were up-regulated and 142 were down-regulated in high-yolk percentage hens. According to GO term annotation and KEGG enrichment analysis, the differentially expressed genes (DEGs) were enriched into the regulation of various cell functions, cell differentiation and development. In addition, a large number of genes were associated with neuroactive ligand-receptor interactions, calcium signaling pathways and ubiquitin-mediated proteolysis. To further filter for genes that actually was involved in yolk accumulation, the chicken quantitative trait loci (QTL) database, genes within 100kb upstream and downstream of the yolk weight trait snp and intersection genes in protein interaction network diagrams were used for the detection of the genes that overlap with DEGs; we found that the 7 candidate genes in total, MNR2, AOX1, ANTXRL, GRAMD1C, EEF2, COMP, and JUND participate in affecting female reproductive performance and the growth and development of follicles, supporting cell transport, cell proliferation and differentiation, immune pathways, being combined with enzyme and cell secretion. All candidate genes and several random selected genes (a total of 24 genes) were verified by qRT-PCR and the trend was consistent with the results of RNA-Seq.

Conclusions: In conclusion, investigating the molecular mechanism of high-yolk percentage traits provides a sight to optimize the breeding strategies to alter the proportion of yolk in chicken eggs.

Background

Egg products are important for human diet all over the world because eggs are rich in excellent proteins, vitamins, minerals, and fats [1]. An egg is primarily made up of the eggshell, egg white and egg yolk. Comparing with other ingredients, yolk contains most content of dry matter, and the forming time of which is the longest around 10 days in general [3]. Studies have shown that the yolk comprises 15.7%–16.6% of proteins, 32%–35% of lipids, 0.2%–1% of carbohydrates, 1.1% of ash and roughly 50% of water; correspondingly nearly 86% of egg white is water [2]. The formation of yolk is regulated by a variety of factors such as diet, light, and neuroendocrine hormones. After yolk lipid precursors are synthesized in the liver, they are released into blood. They then enter the ovary, bind to receptors on the surface of the oocyte, and enter the follicle through endocytosis. Therefore, follicle development has a significant effect on yolk weight [4].

Yolk is the main source of egg flavor, the size of the yolk directly affects flavor and taste [5], and the absolute nutritional value of yolk is relatively high. Eggs with large yolks are more popular among the most of consumers[6], but it is not ruled out that some people think the opposite because of the high cholesterol in egg yolks. Also large proportion of yolk normally means more dry matter in eggs, which reduces the feeding efficiency to some extent. The correlation coefficient between proportion of yolk and feed conversion ratio is −0.25[7], thus hens with high percentage of yolk need more feed, which cut into profits of producers. To both satisfy the needs of consumers and protect the interests of producers, it is necessary to find a balance point. Using genetic strategy to alter yolk percentage is the most essential approach. Hence, it is important to investigate the genetic mechanism for yolk size.

In recent years, there are also many studies on follicular development and yolk weight. The initial growth and development of avian ovarian follicles at puberty and with each subsequent reproductive season in wild birds occurs in an orderly and progressive fashion with all stages of follicle development eventually present at the onset of egg production. Specifically, the initiation of lay is preceded by the activation of primordial follicles (initial recruitment) from within the ovarian cortex followed by the organization of primary follicles. This transition occurs via the development of a single inner layer of granulosa cells (GC) plus the incorporation of a multicellular theca layer. (Ovarian follicle selection and granulosa cell differentiation)

In the poultry breeding programs, the relationship between yolk percentage, hatching rate, and chick weight has also been studied. However, the genes involved in the regulation of yolk size have not been studied in detail, and the proportion of yolk has not been examined independent of egg weight and yolk weight. Because only focusing on the yolk weight causes the egg weight increasing, indirectly affecting the percentage of yolk, many methods of studying animal production and health have been developed and transcriptomic analyses have become important components of systems genomic or systems biology methods which provide a snapshot of all the gene expression profiles in a given tissue and insight into the gene functions pertaining to a particular trait[8]

Thus, in our study, hens with different percentages of yolk were selected for analysis after controlling the egg weight and percentage of eggshell. Using RNA-seq technology, we sequenced the ovaries of high- and low-yolk percentage hens to identify differentially expressed genes (DEGs). We then annotated these genes, and determined whether any were located on "yolk weight" quantitative trait locus (QTL).

Results

Phenotypic data

Table 1 Summary of (phenotypic data) sequencing chickens' phenotypic data in 43 weeks

Trait	H group(n=6) Mean±SD	L group(n=6) Mean±SD	Overall (n=390) Mean±SD
EW(27 weeks)	44.45±2.46	45.99±2.67	-
EW(33 weeks)	46.69±2.79	47.07±2.81	-
EW(43 weeks)	50.04±2.05	50.02±2.24	48.73±3.91
YW(g)	17.05±0.63[a]	13.51±0.7[c]	15.08±1.42[b]
PY	0.34±0.01[a]	0.27±0.01[c]	0.31±0.02[b]
ESW(g)	5.03±0.73	4.96±0.56	4.79±0.56
PS	0.1±0.01	0.1±0.01	0.1±0.01
ESS	2.75±0.36	3.24±0.67	3.09±0.75
EST(mm)	297.33±16.9	297.61±28.79	287.56±30.28
ESI	1.35±0.05	1.33±0.04	1.36±0.05
AH(mm)	4.99±0.73	5.36±0.76	5.32±0.95
YC	6.38±1.28	5.38±1.22	5.76±1.26
HU	72.29±6.37	75.22±6.35	75.35±7.48
BW(g)	1743.3±187.0[a]	1402.8±241.0[b]	-
OW(g)	52.62±7.08[a]	35.28±4.28[b]	-
ROW(g)	10.20±2.70	9.03±2.71	-
NOFR	5.33±1.03	4.33±0.82	-
NOF	54.00±13.91	45.50±12.11	-
LW(g)	47.09±10.51[a]	31.10±5.36[b]	-

EW:egg weight; YW:yolk weight; PY:percentage of yolk; ESW:egg shell weight; PS:percentage of shell; ESS: eggshell strength; EST: eggshell thickness; ESI:eggshell index; AH: albumen height; YC:yolk color; HU: haugh unit; BW:body weight; OW:ovary weight; NOFR: number of follicles removed; ROW: remaining ovary weight; NOF:number of follicle; LW:liver weight; Overall means all the population

All the phenotypic data in table 1 belong to 43-week-aged point, with the first two columns showing the high percentage of yolk and low percentage of yolk hens all the phenotypic traits, containing the egg weight , yolk weight, percentage of yolk, egg shell weight, percentage of shell, eggshell thickness, eggshell index, albumen height, yolk color, haugh unit, body weight, ovary weight, number of follicles removed, remaining ovary weight, number of follicle, liver weight and the third column with the whole population. Chickens selected for RNA-Seq were assigned to either a low (L) or high (H) percentage of yolk group and displayed no significant difference in egg weight and eggshell weight and any other egg quality traits without percentage of yolk and yolk weight at the studied age points. The sequencing chickens were slaughtered at this time point. Postmortem measurements further revealed no significant differences in the number of follicles, number of follicles (> 10 mm) shed, or number of small follicles (2 mm–10 mm) remaining. However, the body weight, intact ovarian weight, and liver weight of chickens in group H were significantly higher than those in group L. As there is a significant difference in yolk percentage between groups H and L, under the condition that the control egg weight does not differ, it follows that the weight of the organ that produces yolk precursor proteins—the liver—is also significantly different. In addition, as the liver is where fat is synthesized and metabolized, a heavier liver will also affect the weight of the chicken. Finally, we found that the overall number of follicles in the ovaries of chickens with high-yolk ratios was higher than in those of low-yolk ratios, but this trend was not significant.

RNA-seq of chicken ovary tissue and DEGs in the high- and low-yolk percentage groups

After removing adaptor sequence and low quality reads, we obtained 4.8×10^8 clean reads, with an average of 3.0×10^7 reads per sample (range: 2.39×10^7 to 3.48×10^7). The proportion of mapped reads was relatively high, ranging from 74.57% to 89.06%. The Q20 and Q30 quality values were 96.56% and 92.11%, respectively. The average effective rate value was 98.72%.

After mapping to the *Gallus gallus* genome, 14,591 genes were identified from all libraries. Among these genes, 7 were expressed only in group H, 29 were expressed only in group L, and 14,555 genes were expressed in both libraries. Among the common genes, 362 genes were found to be differentially expressed between these two groups, of which 142 and 220 were significantly upregulated and downregulated in the high-yolk percentage group respectively (Figure 1). Table 2 lists the 10 genes that were most highly up-regulated or down-regulated in group H.

Table 2. Detailed information about the top 10 up-regulated and down-regulated genes in the H group

Ensembl Gene ID	Gene	FoldChange	P-value	Up/Down	Gene description
Down-regulated genes					
ENSGALG00000029011	SLC35F1	0.0424959	0.00518268	Down	SoluteCarrierFamily35MemberF1
ENSGALG00000030032	SLC4A10	0.0270718	0.000457706	Down	Solute Carrier Family 4 Member 10
ENSGALG00000009645	ESRRG	0.0303217	0.000548399	Down	Estrogen Related Receptor Gamma
ENSGALG00000039023	NIPBLL	0.0511349	1.71588E-05	Down	Nipped-B homolog-like
ENSGALG00000043080	CCDC85A	0.0232416	0.002943567	Down	Coiled-Coil Domain Containing 85A
Up-regulatedgenes					
ENSGALG00000002907	MYL1	13.436006	0.002334536	Up	Myosin Light Chain 1
ENSGALG00000025958	LOC112530098	26.028801	0.002391896	Up	small nucleolar RNA U3
ENSGALG00000031211	LBX2	66.24974	2.72767E-05	Up	Ladybird Homeobox 2
ENSGALG00000037051	ODF3L15	62.085633	0.004124084	Up	outer dense fiber protein 3-like 15
ENSGALG00000020160	ADH6	11.431479	0.009049071	Up	Alcohol Dehydrogenase 6

GO and KEGG analysis of DEGs

We used the clusterProfiler package in R project to perform KEGG enrichment analysis on the 362 DEGs to identify their biological functions and associated pathways. Gene ontology (GO) enrichment analysis was also used to determine the molecular functions, cellular components and biological processes related to the DEGs we identified (Figure 2). We found that the list of DEGs were enriched for extracellular matrix organization, chondrocyte proliferation, extracellular structure organization, platelet activation, collagen-activated signaling pathway, ceramide biosynthetic process, ceramide metabolic process, positive regulation of cell migration, negative regulation of cellular component movement, negative regulation of alpha-beta T cell differentiation. GO analysis further showed that the DEGs are mainly responsible for cell proliferation, differentiation, and movement, and also participate in neuro-metabolic regulation.

The ten most-enriched KEGG categories related to the following pathways: neuroactive ligand-receptor interaction, focal adhesion, regulation of actin cytoskeleton, ECM-receptor interaction, glycerophospholipid metabolism, platelet activation, relaxin signaling pathway, ubiquitin digestion and absorption, adipocytokine signaling pathway, and folate biosynthesis.

Gene networks

To investigate the DEGs interacting with each other, we analyzed the major DEGs(p.adjust < 0.05) using the STRING 11 database. Several interacting genes were identified from differential expressed genes. From figure 3, we found that in addition to the candidate gene, some of the nodes in the figure have spiral structures inside, which means that the three-dimensional structure of the protein is known. If not, the nodes are empty. Line thickness indicates the strength of data support and hide disconnected nodes in the network. EEF2 may have a significant influence on the size of egg yolk which interacts with RPS4X, C12orf66, TOP2A, ATAD2B. In addition ,we also add more nodes to current network, to predict RPS8 RPS11, RPS3A RPL3, RPL8 might have synergy to EEF2 gene, which is involved in metabolic regulation of cells and we classified it as a candidate gene(Table 3).

Candidate genes

To identify candidate genes, we further examined the 362 differentially expressed genes (Additional file 1： Table S1) between the H and L groups. The DEGs with the highest fold changes were SLC35F1, LBX2, SLC4A10, ESRRG, and CCDC85A. These genes have not been previously studied, and our RNA-Seq data suggest that they may be involved in the regulation of yolk size; however, these data alone are not enough to propose them as candidate genes for yolk deposition. To further identify candidate yolk-related genes, the genomic regions associated with the traits “yolk weight” and “ovary weight” (Additional file 2： Table S2) were selected for analysis on the basis of the location of QTLs (https://www.animalgenome.org/cgi-Bin/QTLdb/GG/index). We then limited the DEG list to include only those genes with an adjusted P-value of less than 0.05, and combined this with the GO and KEGG pathway results. The following four genes—MNR2, AOX1, ANTXRL, and GRAMD1C—emerged as promising candidate genes for yolk synthesis, transport and metabolism during the egg laying process. Details of the candidate genes identified in the comparison of groups H and L are listed in Table 3. Because all these genes correspond to metabolic pathways, signaling, and cellular process pathways, the level of their expression may indirectly regulate the growth and development of yolk. We also searched for SNPs that were reported in a genome-wide association study (GWAS) as relating to yolk weight, and examined all genes located up to 100 kb upstream and downstream of these SNPs. Comparing these results with our list of DEGs, we found two genes were common to both datasets: COMP and JUND. At last, we found that the EFF2 gene also has potential as a candidate gene through the protein interaction of the STRING website.

Table 3. Summary of candidate genes involved in Ovary weight and Yolk weight, based on differential expression in ovary tissue samples between the H group and L group and genome-wide association analyses for yolk weight

Symbol	CHR	Ensembl gene ID	Log2 fold change	Gene name	padj	QTL
MNR2	7	ENSGALG00000011349	-1.27	homeodomain protein	0.04	Ovary weight
AOX1	7	ENSGALG00000008185	-1.26	aldehyde oxidase 1	0.04	Ovary weight
ANTXRL	6	ENSGALG00000005969	-1.39	anthrax toxin receptor-like	0.04	Ovary weight
GRAMD1C	1	ENSGALG00000035629	-1.23	GRAM domain containing 1C	0.04	Yolk weight
EEF2	28	ENSGALG00000033884	1.48	eukaryotic translation elongation factor 2	0.0000794	-
Symbol	**CHR**	**Ensembl gene ID**	**Log2 fold change**	**Gene name**	**Pvalue**	**Snp position**
COMP	28	ENSGALG00000003283	1.65	cartilage oligomeric matrix protein	0.001	rs312474469
JUND	28	ENSGALG00000043641	1.15	JunD proto-oncogene	0.001	rs315213484

Quantitative RT-PCR validation

The expression of all the candidate genes were validated by quantitative RT-PCR, including 5 down-regulated genes (MNR2, AOX1, ANTXRL, GRAMDIC and EEF2) and 2 up-regulated genes (COMP and JUND). Besides, to verify the accuracy of our RNA-Seq results, we also randomly selected 18 DEGs (ADH6, MYL1, LOC112530098, ODF31L15, LBX2, CCDC85A, ESRRG, NIPBLL, SLC4A10, SLC35F1, MFAP2, ALPL, ADNP, FPGT, CITED2, SRPX, FKBP14 and CACNA1H) for RT-PCR validation. The β-Actin gene was used as a reference gene to normalize the expression levels of the 25 DEGs. The results showed that the gene expression detected by these two methods were generally consistent, except for ADNP gene, which were less consistent (Figure 4).

Discussion

In this experiment, the egg quality of Wenchang laying hens was measured over a complete laying period, and the relationship between percentage of yolk and egg weight and yolk weight was investigated. The majority of the health-promoting ingredients in eggs are present in the yolk, compared with egg white—which is 88% water. Therefore the percentage of yolk is an important indicator for measuring the nutrition of eggs[9], with a higher yolk percentage indicating a higher nutritional level. As we known that the percentage of yolk is regulated by the egg weight and yolk weight. According to previous reports, in normal flocks there is a positive relationship between egg weight and body weight [10]. In a study of White Leghorn chicken, the genetic correlation between the percentage of yolk, and yolk weight, was about 0.52[11]. In comparison, Rodda et al. (1977) reported the genetic correlation between yolk percentage and egg weight in the poultry population of the Canadian Institute of Agricultural Animals were between −0.28 and −0.1[12]. From this it can be seen that there is a significant negative correlation between yolk percentage and egg weight. The reason why yolk percentage decreases with an increase in

egg weight, may relate to excessive eggshell weight or the excessive secretion of egg white. During continuous egg production, the size of yolk is almost unchanged, while the secretion of other proteins can be quite different. Here, we corrected for egg weight by selecting chickens with consistent egg weight production, before performing RNA-Seq on ovarian tissue. This allowed us to narrow the range of factors affecting yolk percentage, and thus better identify candidate genes relating to the production of high-yolk percentage eggs. Compared with traditional cDNA microarray technologies, RNA-Seq has many advantages—such as a greater dynamic range, lower bias, lower frequency of false-positives, and higher repeatability[13]. Moreover, results from RNA-Seq transcriptomes have a high correlation with RT-PCR results. In general, including more biological replicates in an experiment increases the reliability of the results. Accordingly, we performed six biological replicates in each of our trials to ensure confidence in our data.

Egg production mainly depends on the number of undifferentiated pre-stage follicles in the ovary. The follicles of birds have unique developmental characteristics; a follicle is selected from a pool of small yellow follicles during the development of the fractional follicles in the follicles prior to grading[14]. The development of follicles plays a crucial role in various physiological processes in poultry, and has an impact on the economics of production. Yolk size and percentage mainly depends on follicle development, and as the ovary plays an important role in regulating the size of follicles[15], the ovary is also important for the formation of yolk. In chickens, the oocyte is deposited about 10 days before ovulation, after which yolk is deposited rapidly. In addition to immunoglobulins, which are rare components of yolk, yolk is mainly derived from yolk precursors that are synthesized by the liver. A hen that produces one egg per day, generates about 19 g of yolk precursors. Yolk precursors are gradually synthesized by the liver, under the action of estrogen, transported to the developing follicle via blood circulation, and deposited in the developing egg by a specific receptor-mediated mechanism. In this study, we have found that participating in the liver lipid metabolism, growth and development, involved in hormone secretion regulation, which may affect some genes in the process of yolk formation; meanwhile we identified genes that may regulate the size of yolk, by comparing gene expression in the ovaries of hens with different yolk percentages. From this dataset we predicted candidate yolk-related genes.

We identified three DEGs located on a QTL relating to ovary weight, namely MNR2, ANTXRL, and AOX1. Notably, the ectopic expression of MNR2 in neural cells initiates a program of somatic motor neuron differentiation that is characterized by the expression of homeodomain proteins [16]. Indeed, neural networks are inextricably linked to the development of follicles in the ovaries [17]. Specifically, the maturity and ovulation of the yolk follicle is profusely innervated by both adrenergic and cholinergic fibers [18].

The aldehyde oxidase 1 (AOX1) gene encodes a homodimeric protein that produces hydrogen peroxide. In mouse, it is involved in myogenesis[19], and it may also play a role in the muscle development of cattle[20]. The avian genome contains two aldehyde oxidase genes, AOX1 and AOH. These are both located on chromosome 7 and are structurally very similar. The function of AOX1 in mice is analogous to its role in chicken, and the prediction of AOX1 as a candidate yolk-related gene has convincing evidence.

AOX1 promotes cell proliferation and can be highly expressed in the liver and lungs[21]; yolk precursors are also synthesized by the liver. This may indirectly affect the formation of yolk precursors in the follicles. Therefore, we have speculated that AOX1 indirectly regulates the growth and development of follicles, and thereby influences the size of yolk. There is also evidence that the inhibition of AOX1 receptor activation protects the placenta against oxidative damage, alterations in protein expression, and mitochondrial dysfunction.

MNR2 belongs to the Mnx-class of homeodomain proteins[16], which takes part in transcriptional repression and the specification of cell identity[22]. Pathways involving MNR2 include: the lipid bilayer surrounding the vacuole and separating its contents from the cytoplasm of the cell, and biological processes involved in the maintenance of chemical and cellular internal steady states. Moreover, the MNR2 gene can induce ectopic expression during embryonic development, indicating that it can indirectly affect female reproductive performance and the growth and development of follicles[23]. Embryonic expression of zebrafish mnx genes is very dynamic; one of these, mnr2a, shows expression in an endodermal cell population that is initially intermingled with insulin-positive cells and that later becomes restricted to the exocrine pancreas[24]. Our results confirm that this gene has an effect on cell secretion, synthesis, and possibly follicular development. During yolk formation, the development of follicles in the ovary plays a crucial role in the yolk deposition, which may in turn affect the size and proportion of the yolk.

A GWAS examining egg yolk weight, follicle weight, and ovarian weight traits in chicken, identified the loci of 12 yolk weight single nucleotide polymorphisms (SNPs), 3 follicle weight SNPs, and 31 ovary weight SNPs[17]. According to the variant effect predictor (VEP; Ensembl), 12 SNP loci were found to predict all genes within 100 kb upstream and downstream. This corresponded to 362 DEGs, including two genes (COMP and JUND), at two sites (rs312474469 and rs315213484). COMP can be enriched in focal adhesion and in the extracellular matrix (ECM)-receptor interaction pathway. Cell-matrix adhesion is highly significant in animal biology, as it supports cell transport, cell proliferation and differentiation, as well as gene regulation. ECM is a complex mixture of structural and functional macromolecules that plays an important role in the morphogenesis of tissues and organs, and in the maintenance of cell and tissue structure and function. JUND is involved in the MAPK signaling pathway. It is a highly conserved module, and this pathway can be involved in a variety of cellular biological functions, including cell proliferation, differentiation and migration. Therefore, we speculate that JUND may affect the number of follicles, and their growth and development in chickens.

The output from the STRING 11 website(The latest version of STRING (11.0) more than doubles the number of organisms it covers, to 5090. The most important new feature is an option to upload entire, genome-wide datasets as input, allowing users to visualize subsets as interaction networks and to perform gene-set enrichment analysis on the entire input) analysis shows that among the significantly DEGs[25], the EEF2 gene has interactions with RPS4X, C12orf66, TOP2A, ATAD2B. This interaction depicts the importance of the EFF2 gene in this output. The function of the EEF2 gene is to participate in the immune pathways, enzyme that catalyzes the transfer of a phosphate group, tissue and cell growth

and development process, combined with enzyme and cell secretion[26]. In addition, the genes interacting with EFF2 gene are all negative regulatory genes, which suggests that EEF2 gene may play a synergistic role with four other genes in regulating the yolk of egg yolk, which has a great influence on the size of egg yolk

By integrating information obtained from DEGs, pathway analysis and published data. This study provides a list of candidate genes whose function and expression levels are closely related to egg yolk size. Most of these candidate genes are located on the Yolk Weight, Ovary weight QTL and some are located around the SNP site that is significantly associated with egg yolk weight. This study used relatively few samples, but they were all in a highly controlled environment.

Conclusion

The study investigated the transcriptome of ovarian tissue in laying hens with different percentage of yolk. The results of this experiment provide an important understanding of the yolk deposition biology in poultry and clarify the biological functions and mechanisms of differentially expressed genes. The research also has the following innovations: On the one hand, on the experimental subjects, we have undergone strict quality control, the group of chickens maintains the characteristics of high or low percentage of yolk at different age points and has a certain stability. On the other hand, as far as we know, this study is the first to use RNA-seq to explore the proportion of Chinese local varieties of egg yolk. RNA-seq is known as the most accurate tecnology for genome-wide gene expression studies. The results revealed differences in biological mechanisms related to egg yolk size. A total of 362 differentially expressed genes were identified and 7 of them were identified as candidate genes. Provide a reference for us to further study the proportion of egg yolk.

Methods

Animal husbandry and ovary collection

In this study we used a pure line of Wenchang chickens with complete pedigrees, which were provided from the farm Yangzhou Poultry Research Institute, Jiangsu province. At 26 weeks, according to appearance, growth and egg production, 20 families were selected (from 64 families), giving a total of 530 chickens. All chickens were kept under a light/dark cycle of 16-hours light and 8-hours dark (16L: 8D), and had free access to feed and water. Eggs were collected over 3–4 successive days at 27, 33, and 43 weeks of age. After egg collection at 10:00, all other measurements were taken on the same day. The changes in egg weight, yolk weight, and yolk percentage of each chicken were analyzed at each age point.

At 43 weeks of age,euthanasia was performed by cervical dislocation, with all effort made to minimize suffering,six chickens from group H and six from group L were sacrificed. Chickens selected for RNA-Seq were assigned to either a low (L) or high (H) percentage of yolk group and displayed no have significant

differences in egg weight at the studied age points(27, 33, 43weeks of age). Groups H and L represented laying hens with high or low-yolk percentages, respectively. For RNA isolation, ovarian tissue was quickly removed from each chicken and placed on dry ice, then stored at −80 ° C.

Total RNA extraction, cDNA library preparation, and sequencing

Twelve cDNA libraries were sequenced from the ovaries of the low (L) and high (H) percentage of yolk groups. Data from the 12 remaining chickens was then analyzed.

Each ovarian tissue sample was placed in a mortar, frozen in liquid nitrogen, and finely ground. Total RNA was extracted from 70–100 mg of each sample using TRIzol reagent (Tiangen Biotech Co., Beijing, China). The quality of the total RNA was checked using the Agilent 2100 Bioanalyzer System (Santa Clara, CA, USA) and by 1% agarose gel electrophoresis (180 v, 16 min).

After the RNA samples were quantified, eukaryotic mRNA was enriched using magnetic beads conjugated to oligo(dT). Then fragmentation buffer was added, breaking the mRNA into short pieces. This mRNA was used as template, with random hexamer primers, for cDNA synthesis. Buffer, dNTPs, DNA polymerase I, and RNase H, were added, and cDNA was synthesized. cDNA was then purified using AMPure XP beads. Purified double-stranded cDNA was first used to repair ends, then A-tails were added and joined to the sequencing adaptors. Finally PCR amplification and AMPure XP beads were used to purify the PCR products. After library construction was complete, Qubit 2.0 was used for preliminary quantification, and to determine library dilution. Agilent 2100 was used to detect the size of inserted fragments in the libraries. After inserts were found to meet the expected size, the effective concentration of each library was quantified using qPCR, to ensure the quality of the library.

To complete library preparation, cDNAs were amplified using PCR, and the cDNA libraries were paired-end sequenced using the Illumina Hiseq Xten platform.

Differential expression analysis

Raw reads were cleaned by removing adaptor sequences, and low quality reads were discarded. Processed reads were then mapped to the chicken reference genome (http://ftp.ensembl.org/pub//release-94/fasta/gallus_gallus/dna/) using Hisat2 software. This process has higher alignment efficiency than TopHat, and can be roughly divided into three steps: establishing a genome index, alignment to genome, and sorting. This process also uses StringTie, which is faster and more accurate than Cufflinks[27], and results in more complete transcripts. DEGs were identified using the DESeq2 software package[28].

Bioinformatic analysis of DEGs and Identify candidate genes

Differential expression tables were imported into the GO and KEGG databases for enrichment analysis. GO terms and KEGG pathways (the clusterProfiler package in R) showing P-values of less than 0.05 were considered as significantly enriched among the DEGs[29]. DEGs were then mapped to the chicken QTL

database. Genes corresponding to yolk weight and ovary weight QTLs were selected, and then the corresponding pathways were examined to select candidate genes for further analysis. By comparative analysis and imperative validation, we detected DEGs mapped to QTL regions—the yolk weight QTL, and the ovary weight QTL. The 100kb genes upstream and downstream of the reported yolk heavy SNP sites were also identified

Confirmation of RNA-Seq results via qRT-PCR

We randomly selected 13 DEGs for RT-PCR to verify the accuracy of our RNA-Seq data. Total RNA was reverse-transcribed into cDNA using the PrimeScript RT reagent kit with gDNA Eraser (Tiangen, Beijing), according to manufacturer's instructions. Primers were designed using primer5 software, based on relevant gene sequences retrieved from NCBI. qRT-PCR was performed in a final reaction volume of 20 μL using the SYBR® Green PCR Master Mix Kit (Roche) in a LightCycler® 480 Real-Time PCR Detection System. The following protocol was used: 95 °C for 10 min; 45 cycles of 95 °C for 10 s, 60 °C for 10 s, and 72 °C for 10 s. Three biological replicates were performed for each sample, and GAPDH and β-Actin were used as reference genes. Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method. A trend analysis was performed on the log2-fold change values from RNA-Seq and qRT-PCR to determine consistency.

Abbreviations

DEG: Differentially Expressed Gene ; QTL: Quantitative Trait Loci ; GO: Gene Ontology; GWAS: Genome Wide Association; KEGG: Kyoto Encyclopedia of Genes and Genomes; mRNA: Messenger RNA; padj: Adjusted p value; RNA-Seq: RNA-Sequencing; SNP: Single Nucleotide Polymorphism

Declarations

Ethics approval and consent to participate

All protocols for collection of the ovary samples of Wenchang chicken were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at China Agricultural University. Ovary samples were collected specifically for this study following standard procedures with the full agreement of the Yangzhou Poultry Research Institute who owned the animals.

Consent to publication

Not applicable.

Availability of data and materials

All supporting data can be found within the additional files

Competing interests

The authors declare that they have no competing interests.

Funding

The current study was funded in part by the China Agriculture Research System (CARS-40).

Authors' contributions

JNZ performed bioinformatics and statistical analysis, and also was a major contributor to manuscript preparation. FNL, XQW and XFL performed experiments and sample collection. CJS participated in result interpretation, wrote, revised and approved the manuscript. GYX and CJS commented the manuscript and were major contribution to manuscript revision. GYX conceived and designed the experiments and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We would like to thank field staff and laboratory technicians who helped carry out this study.

Author details

Key Laboratory of Animal Genetics and Breeding of Ministry of Agriculture, National Engineering Laboratory of Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China.

References

1. Munger LH, Garcia-Aloy M, Vazquez-Fresno R, Gille D, Rosana A, Passerini A, Soria-Florido MT, Pimentel G, Sajed T, Wishart DS et al. Biomarker of food intake for assessing the consumption of dairy and egg products. Genes Nutr. 2018, 13:26.
2. Kovacs-Nolan J, Phillips M, Mine Y. Advances in the value of eggs and egg components for human health. J Agric Food Chem. 2005, 53(22):8421-8431.
3. Marza VD. The formation of the hen's egg.V.Histochemistry of yolk-formation. Proteins. Quarterly Journal Of Microscopical Science. 1935, 78(310):191-249.
4. Arukwe A, Goksøyr A. Eggshell and egg yolk proteins in fish: hepatic proteins for the next generation: oogenetic, population, and evolutionary implications of endocrine disruption. Comp Hepatol. 2003, 2(1):4.
5. Liu X, Zhang Y, Yan P, Shi T, Wei X. Effects of conjugated linoleic acid on the performance of laying hens, lipid composition of egg yolk, egg flavor, and serum components. Asian-Australas J Anim Sci. 2017, 30(3):417-423.
6. Fernandez ML, Andersen CJ. The good egg, the forgotten benefits: protein, carotenoids, choline and glycemic index. Handbook of eggs in human function. 2015: 17-34.

7. Hartmann C, Strandberg E, Rydhmer L, Johansson K. Genetic relations of yolk proportion and chick weight with production traits in a White Leghorn line. Br Poult Sci. 2003, 44(2):186-191.
8. Salleh MS, Mazzoni G, Hoglund JK, Olijhoek DW, Lund P, Lovendahl P, Kadarmideen HN. RNA-Seq transcriptomics and pathway analyses reveal potential regulatory genes and molecular mechanisms in high- and low-residual feed intake in Nordic dairy cattle. BMC Genomics. 2017, 18(258).
9. Hartmann C, Wilhelmson M. The hen's egg yolk: a source of biologically active substances. World's Poultry Science Journal. 2001, 57(1):13-28.
10. Rahn H, Paganelli CV, Ar A. Relation of Avian Egg Weight to Body Weight. The Auk 1975, 92(4):750-765.
11. Hartmann C, Johansson K, Strandberg E, Wilhelmson M. One-generation divergent selection on large and small yolk proportions in a White Leghorn line. Br Poult Sci. 2000, 41(3):280-286.
12. Rodda DD, Friars GW, Gavora JS, Merritt ES. Genetic parameter estimates and strain comparisons of egg compositional traits. Brit Poultry Sci. 1977, 18(4):459-473.
13. Wilhelm BT, Landry JR. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. Methods. 2009, 48(3):249-257.
14. Hernandez AG, Bahr JM. Role of FSH and epidermal growth factor (EGF) in the initiation of steroidogenesis in granulosa cells associated with follicular selection in chicken ovaries. Reproduction. 2003, 125(5):683-691.
15. Kang L, Cui XX, Zhang YJ, Yang CH, Jiang YL. Identification of miRNAs associated with sexual maturity in chicken ovary by Illumina small RNA deep sequencing. BMC Genomics. 2013, 14.
16. Tanabe Y, William C, Jessell TM. Specification of Motor Neuron Identity by the MNR2 Homeodomain Protein. Cell. 1998, 95(1):67-80.
17. Sun C, Lu J, Yi G, Yuan J, Duan Z, Qu L, Xu G, Wang K, Yang N. Promising Loci and Genes for Yolk and Ovary Weight in Chickens Revealed by a Genome-Wide Association Study. Plos One. 2015, 10(9):e137145.
18. Odlind V, Rosen M, Stenbeck M. FEMALE GENITAL ORGANS. Acta Oncol. 1995, 344:53-59.
19. Kamli MR, Kim J, Pokharel S, Jan AT, Lee EJ, Choi I. Expressional studies of the aldehyde oxidase (AOX1) gene during myogenic differentiation in C2C12 cells. Biochem Biophys Res Commun. 2014, 450(4):1291-1296.
20. Guillocheau GM, El HA, Meersseman C, Esquerre D, Rebours E, Letaief R, Simao M, Hypolite N, Bourneuf E, Bruneau N et al. Survey of allele specific expression in bovine muscle. Sci Rep. 2019, 9(1):4297.
21. Terao M, Kurosaki M, Saltini G, Demontis S, Marini M, Salmona M, Garattini E. Cloning of the cDNAs coding for two novel molybdo-flavoproteins showing high similarity with aldehyde oxidase and xanthine oxidoreductase. J Biol Chem. 2000, 275(39):30690-30700.
22. William CM, Tanabe Y, Jessell TM. Regulation of motor neuron subtype identity by repressor activity of Mnx class homeodomain proteins. Development. 2003, 130(8):1523-1536.

23. Wang Y, Li J, Feng C, Zhao Y, Hu X, Li N. Transcriptome analysis of comb and testis from Rose-comb Silky chicken (R1/R1) and Beijing Fatty wild type chicken (r/r). Poultry Sci. 2017, 96(6):1866-1873.
24. Wendik B, Maier E, Meyer D. Zebrafish mnx genes in endocrine and exocrine pancreas formation. Dev Biol. 2004, 268(2):372-383.
25. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019, 47(D1):D607-D613.
26. Kamel NN, Ahmed AMH, Mehaisen enGMK, Mashaly MM, Abass AO. Depression of leukocyte protein synthesis, immune function and growth performance induced by high environmental temperature in broiler chickens. Int J Biometeorol. 2017, 61(9):1637-1645.
27. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015, 33(3):290.
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014, 15(12).
29. Yu GC, Wang LG, Han YY, He QY. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS. 2012, 16(5):284-287.

Figures

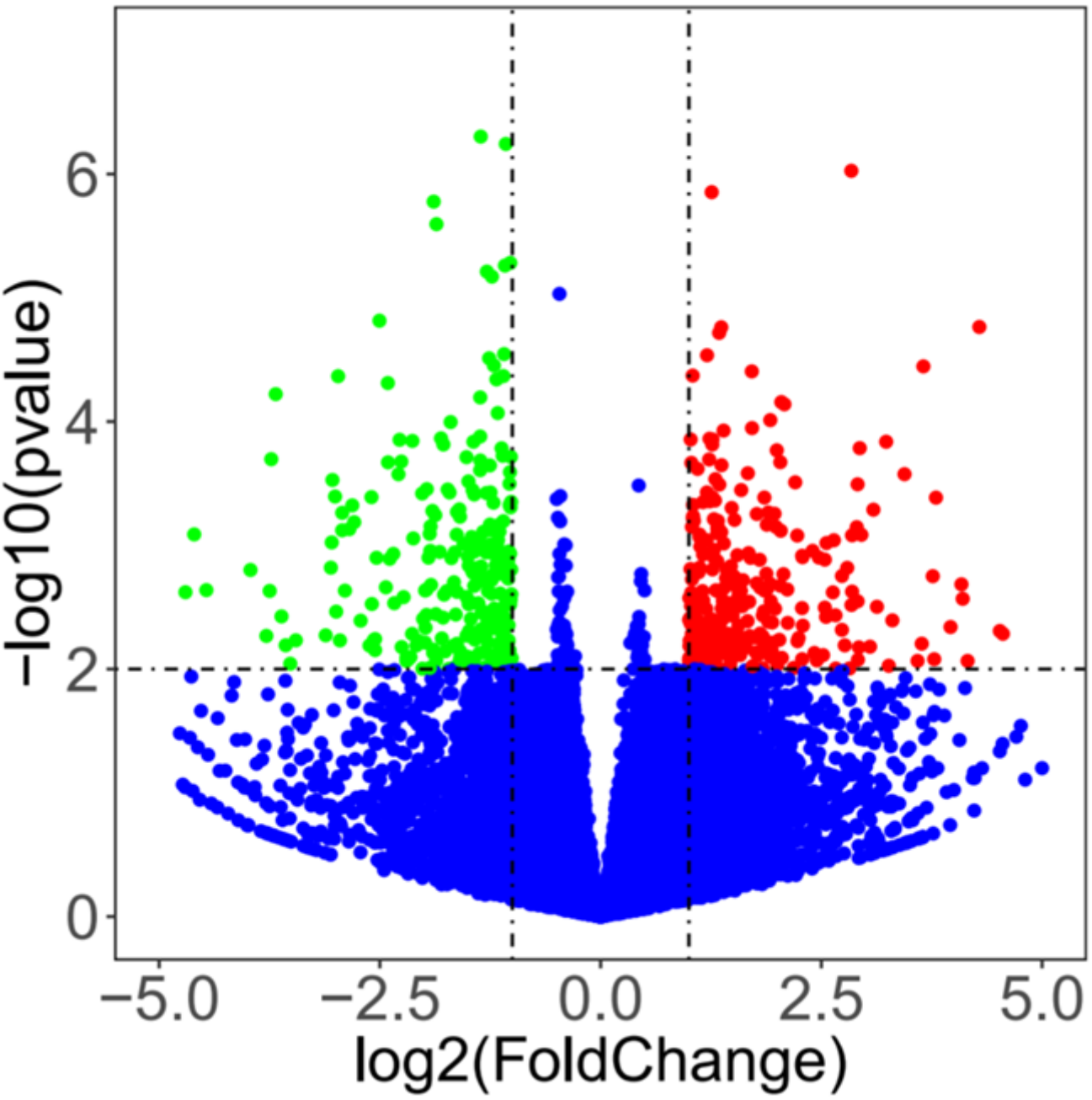

Figure 1

Scatter plot of differentially expressed genes (H Group vs L Group). Red points represent up-regulated genes with a log2 (fold change) > 1 and pvalue < 0.01. Green points represent down-regulated genes with a log2 (fold change) <−1 and padj < 0.01. Blue points represent genes showing no significant difference. Fold change = normalised gene expression in the H Group / normalized gene expression in the L Group.

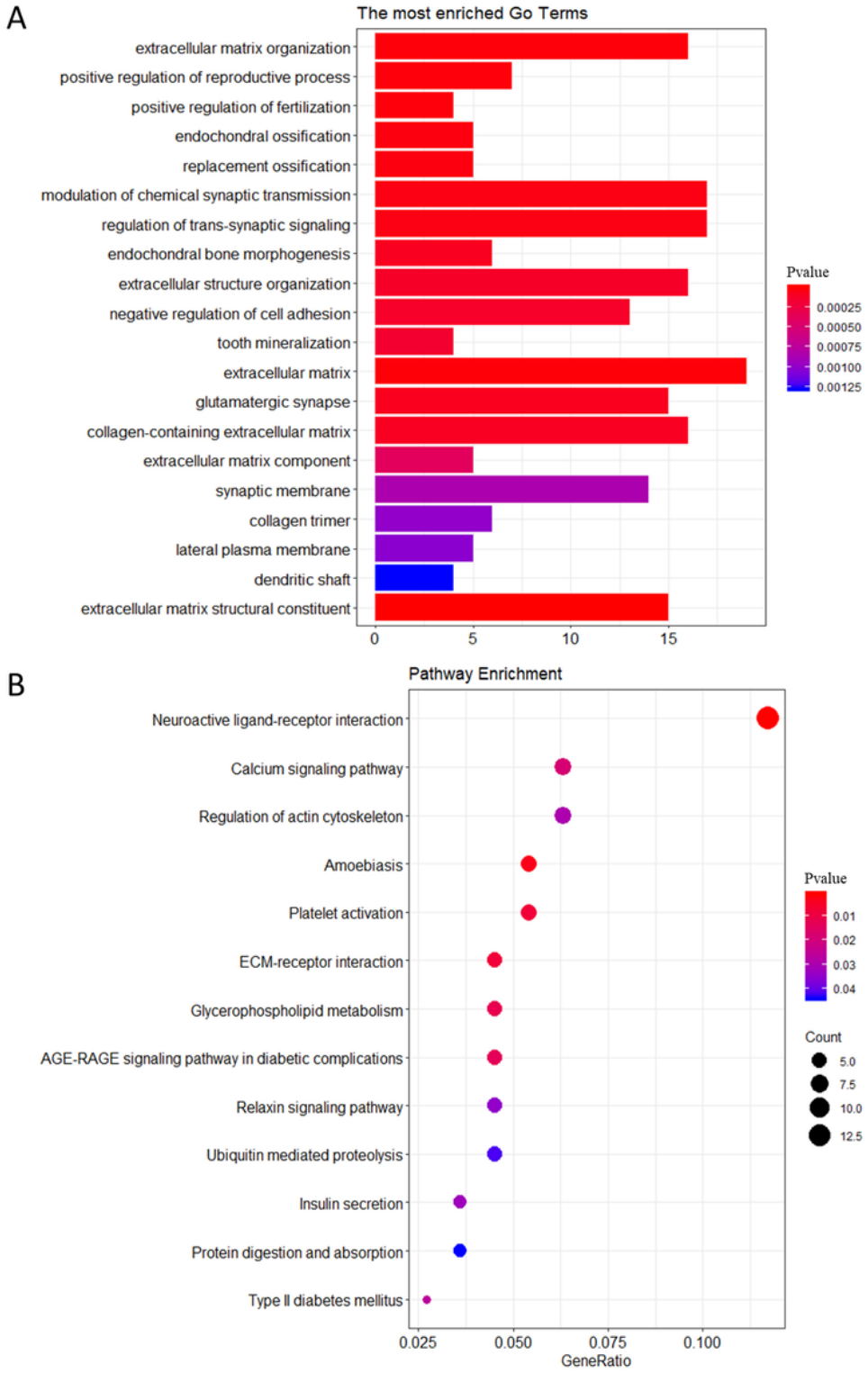

Figure 2

Go-enriched map of and Pathway enrichment scatter ploy of differentially expressed genes. (A)Different colors represent different degrees of significance(Pvalue□0.05),and the abscissa value means the number of enriched genes. (B) Different colors also represent different degrees of significance (P.value < 0.05).The ordinate indicates the path name, and the abscissa value means the ratio of the number of

enriched genes to the total number of genes. The size of the point indicates how many differentially expressed genes are in the pathway, and the color of the point corresponds to a different P.value range.

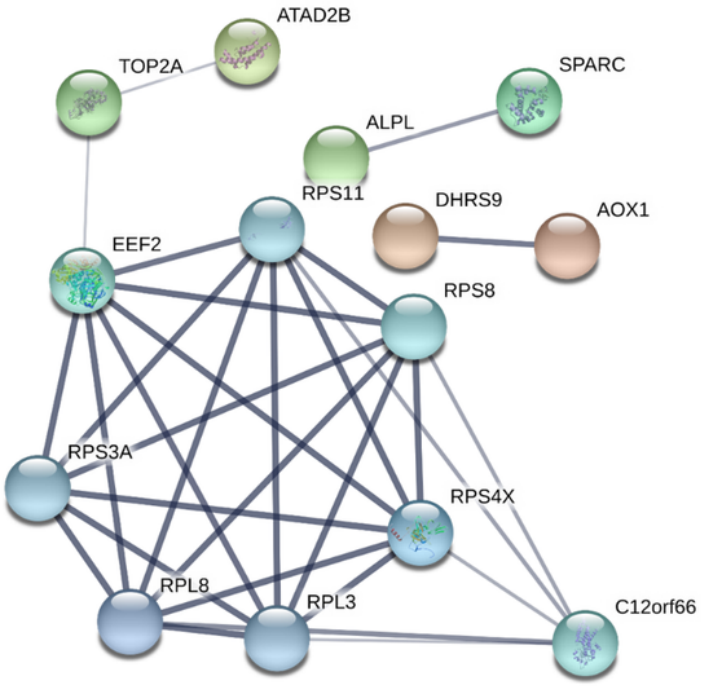

Figure 3

The relationship between 42 DEGs in network. Each circle represents a protein. The helix in the protein represents the known structure of the protein and the thickness of the lines between proteins represents the degree of interaction between proteins.

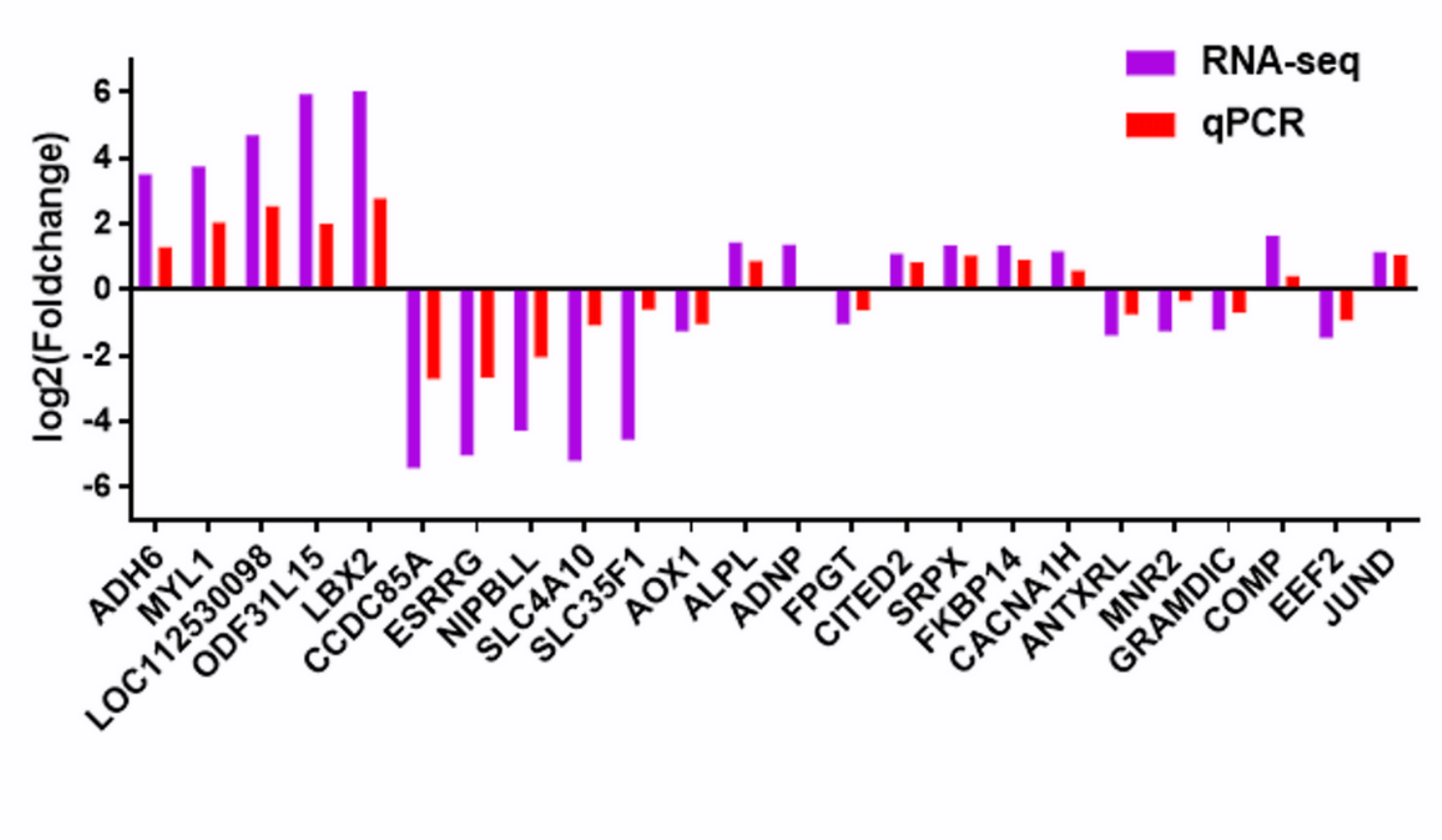

Figure 4

Expression of differentially expressed genes. Correlations of the mRNA expression levels of 23 differentially expressed genes (including all candidate genes). The purple column and red column show the log2 (ratio of mRNA levels) values measured via RNA-Seq and qRT-PCR, respectively.

Supplementary Files

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

- Additionalfile2.xlsx
- Additionalfile1.xlsx