

Identification of mRNA and Long Non-Coding RNA Expression Profiles in Developing Yorkshire Pig Spleens

Xinjian Li

Henan Agricultural University

Xuelei Han

Henan Agricultural University

Caixia Sun

Henan Agricultural University

Gaiying Li

Henan Agricultural University

Kejun Wang

Henan Agricultural University

Xiuling Li

Henan Agricultural University

Ruimin Qiao (✉ qrm480@163.com)

Henan Agricultural University <https://orcid.org/0000-0002-2648-6227>

Research

Keywords: mRNA, lncRNA, Yorkshire pigs, spleen

Posted Date: April 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-441442/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Epidemic diseases cause great economic loss in pig farms each year, some of which are characterized mainly in spleen. Yorkshire pig is the most popular used first dam in the commercial pork production system. But the mRNA and lncRNA expression networks in developing Yorkshire pig spleens remain obscure.

Results: Here, we profiled the systematic characters of mRNA and lncRNA repertoires in three groups of spleens from nine Yorkshire pigs, each three aged at 7 days, 90 days and 180 days. By using a precise mRNA and lncRNA identification pipeline, we identified 19,647 genes and 219 known and 3,219 putative lncRNA transcripts, 1,729 genes and 64 lncRNAs therein were found to express differentially in three groups. Gene expression characteristics of genes and lncRNAs were found to be basically fixed before 90 days after birth. Enrichment analysis of differentially expressed genes and potential target genes of differentially expressed lncRNAs both displayed crucial roles of up-regulation in immune activation and hematopoiesis and down-regulation in cell replication and division in 90 and 180 days compared to 7 days. The unregulated terms and their significance levels in 90 and 180 days both showed an extremely high degree of consistency. ENSSSCT0000001325 was the only lncRNA transcript that existed in three groups. *CDK1*, *PCNA* and *PLK* were detected to be hub genes that varied with age. BNIP3L, IL5, CD38 and TGF β 1 were found to be common top regulators from 7 to 90 and 180 days while ERAP1, NLRC5 and IL2RG were top regulators from 90 to 180 days.

Conclusions: This study provided the first mRNA and lncRNA expression profiles in Yorkshire spleens at three developmental stages. We established gene expression modules and networks in the spleen of pigs from immune system initiation to adulthood. Our results are helpful for the study of transcriptome and functional genomics of spleen tissue in farm animals.

Background

Epidemic disease in farms is an outstanding problem hampering pig production. As spleen plays a vital role in hemofiltration and antigen-specific immune and responses as a blood bank and the largest immune organ, many epidemic diseases in pig farms are associated with spleen damage, such as PRRS, PCV, Pseudorabies, Swine Fever and so on. To cope with those common epidemics, pig farms have a routine immunization program. Immunization activities of piglets usually start from 7 days after birth. And in general, hogs complete their routine immunization procedures at 70 days and are market in about 180 days after enough withdraw time. Therefore, it is of great significance to understand the time spectrum of immune system development in spleen of domestic pigs. But, only a few of reports on RNA expression of pig spleen were published [1–6].

Elucidating the development process of spleen immune system needs an in-depth understanding of RNA expression. In addition to the mRNA that encodes the protein; long non-coding RNA (lncRNA) as a group of non-coding RNA with a length longer than 200 bp, has various styles, bulky quantity and function diversity although the expression level of lncRNAs are lower than mRNAs [7]. lncRNA are known to affect in shearing, transduction and transcription and degradation of mRNA through combination with mRNA after transcription and function in epigenetic regulation [8, 9]. Up to 2,297 causative associations between lncRNAs and diseases have been detected [10, 11] (<http://www.manut.net/lncrnadisease/>). In other livestock such as cattle and chicken, lncRNAs have also been found to be associated with some common disorders, such as mastitis [12, 13], growth-retardation [13], cancer [14] and Marek's disease in chickens [15].

Yorkshire pig is the most widely used first dam in commercial pig production, but no temporal RNA sequencing study was reported about Yorkshire spleen and the likely molecular immune mechanisms changes as the age increases in Yorkshire spleen remains controversial. In order to understand gene and lncRNA repertoires in Yorkshire spleen, we now

appreciate the pervasive transcription of numerous mRNA and lncRNAs in Yorkshire spleens at 7 days, 90 days and 180 days after birth in order to help us in understanding pig immune system development better.

Materials And Methods

Ethics statement

All procedures involving animals are following the use guidelines of experimental animals established by the Ministry of Agriculture of China and the ethics committee of Henan Agricultural University.

Sample collection

Biospecimens were collected from nine healthy Yorkshire pigs, each three at the age of 7 days, 90 days and 180 days. All the animals were self-feeding in independent cages with normal epidemic prevention. Spleen tissues from these nine pigs were collected under the same condition according to the requirement of total RNA extraction.

Library preparation and RNA sequencing

Total RNA of the above nine spleen tissues were extracted using TRIZOL Reagent (Cat#15596-018, Life technologies, Carlsbad, CA, US) following the instructions. Each specimen was purified by RNeasy micro kit (Cat#74004, QIAGEN, GmbH, Germany) and RNase-Free DNase Set (Cat#79254, QIAGEN, GmbH, Germany), and qualified with Nanodrop ND-2000 and Agilent Bioanalyzer 2100 (Agilent technologies, Santa Clara, CA, US) to have RNA Integrity Numbers (RIN) of 9.8-10 and concentrations more than 100 ng/ μ L. Qualified RNA samples of three individuals at the same age were equally pooled together to form three RNA groups, Y-7, Y-90 and Y-180. These three RNA groups were sent to Shanghai Biotechnology Corporation for lncRNA and mRNA sequencing by Illumina HiSeq 2500 sequencer to produce paired reads (2 \times 125 nt).

RNA mapping and expressing quantification

Pair-end raw reads were filtered out to remove adaptors, low-quality reads that shorter than 25 bases, or having Q20 value less than 20 to get clean reads under the help of Seqtk software (<https://github.com/lh3/seqtk>). Trimmed clean reads were aligned to pig genome assembly (*Sus scrofa* 10.2) using TopHat (version: 2.0.13) [16] with default sets. Mapped transcripts were assembled individually with Cufflinks and Cuffquant [17]. lncRNA was identified by Cuffcompare (version 2.2.1) [18], Hmmscan [19], BLASTX and Coding Potential Calculator (CPC) [20]. Those transcripts that do not code a protein or a known protein-coding domain and with CPC score no more than 0 were considered as putative lncRNAs. RNA expression level of the genes, transcripts and lncRNAs was calculated using FPKM (fragments per kilobase of exon per million fragments mapped) [21]. Differential expressed genes and lncRNAs were identified by fold change Fisher-test. We used false discovery rate (FDR) of no more than 0.05 ($FDR \leq 0.05$) as the significant threshold.

Enrichment analysis upon differential expressed genes and potential target genes of lncRNAs

In order to evaluate the mRNA expression level and involving pathways in different development period of Yorkshire spleen, genes were first converted into their corresponding H. sapiens Entrez gene IDs using the latest version of the Metascape database (<https://cytoscape.org/>, updated on 2021-02-01). The enrichment analysis of GO term, KEGG pathway and protein-protein interaction upon the differential expressed mRNAs and potential targeting genes of lncRNAs were performed using Metascape. All genes in the genome have been used as the enrichment background. Terms with a p -value < 0.01, a minimum count of 3, and an enrichment factor > 1.5 (the enrichment factor is the ratio between the

observed counts and the counts expected by chance) were collected and grouped into clusters based on their membership similarities.

Gene network analysis

The gene expression profile at three different developmental timing were clustered to find out genes with the same expression model using STEM-Short Time-series Expression Miner software [22]. Protein and protein interaction (PPI) network for all the differential expressed genes in the spleen was analyzed using STRING (<https://string-db.org/>) and displayed using Cytoscape (<https://cytoscape.org/>). Top gene networks and regulators in Y-90 vs Y-7, Y-180 vs Y-90 and Y-180 vs Y-7 were obtained through IPA (<http://www.ingenuity.com/products/ipa>).

Real time PCR

LncRNA and mRNA quantification was performed with using real time PCR. CDNA from Yorkshire spleen aged at 7 days, 90days and 180 days were synthesized with random primers using the First Strand cDNA Synthesis Kit (Takara, Japan) according to the instruction. The PCR reaction system contained 5 uL 2× Taq Master Mix (Takara, Japan), 3.6 uL sterilized distilled water, 0.2 uL forward primers (10 μM), 0.2 uL reverse primers (10 μM) and 1 uL cDNA. The PCR reaction procedure was as follows: pre-denaturation at 94°C for 5min, followed by 40 cycles of pre-denaturation at 94°C for 30s, annealing for 30s, extension at 72°C for 30s and extension for 8 min, stored at 12°C. Primer sequences used for the RT-PCR experiments were listed in Table S1. The expression level of RNA was calculated using the $2^{-\Delta\Delta Ct}$ method allowing standard error no more than 0.2.

Results

Statistics of clean reads alignment

After RNA sequencing and reads quality control, we obtained over 98 million 2×125 bp paired-end trimmed clean reads from each group, corresponding to an average of 12.3 Gb in size, of which over 79.4% were mapped to the *Sus scrofa* 10.2 assembly genome (Table 1). Approximately, 75.3%, 36.8%, 2.4%, 24.9%, 13.5%.24.7% of the clean reads were mapped to gene, coding, splicing, intron, no-coding and intergenic regions in Y-3 while 71.5%, 27.6%, 1.3%, 29.6%, 14.3%, 28.5% were mapped in Y-90 and 71.7%, 27.9%, 1.4%, 29.3%, 14.5%, 28.3% were mapped in Y-180 (Fig. S1). After aligning, we obtained 19,647 genes in nine Yorkshire spleens at 7, 90 and 180 days old in total.

Profile of distinct expressed genes

In order to understand how genes varied in spleen tissues of Yorkshire pigs at different development stages, we compared the expression levels of genes in the three age groups, Y-7, Y-90 and Y-180. Our ranking analysis (see Materials and Methods) yielded a list of 1,729 differential expressed genes, among which 1490 (617↑, 837↓) genes were in Y-90 vs Y-7, 1160 (556↑, 604↓) genes were in Y-180 vs Y-7 and 124 (58↑, 66↓) genes were in Y-180 vs Y-90. We found that the gene expression pattern in Y-90 and Y-180 showed a high degree of consistency (Fig. 1). Of the 1160 differently regulated genes in Y-180 vs Y-7, 957 were also detected in Y-90 vs Y-7. Only 27 differential expressed genes were at in common at trees stages, of which 11 gene were protein coding genes with characters, including *IGKV-3*, *IGKV-7*, *IGKV-11*, *IGLV-2*, *IGLV-3*, *SLA-3*, *SLA-10*, *CXCL9*, *HDAC9*, *HSPA1A* and *MRC2*. What was interesting was that some extremely differential expressed miRNAs were detected. For example, ssc-mir-199a-2 expressed in an extremely high level in Y-7 but was almost absent in Y-90 and Y-180 ($\downarrow 4.7E-10$, $\downarrow 4.7E-10$) whereas ssc-mir-4332 expressed in an extremely high level in Y-7 and Y-180, but was absent in Y-90 ($\downarrow 1.1E-8$). In addition, ssc-mir-451 was down regulated both in Y-90 and Y-180 ($\downarrow 0.2$, $\downarrow 0.2$).

GO and KEGG enrichment analysis of the distinct genes

After converting all the genes in pig into their corresponding H. sapiens Entrez gene IDs, 413, 27 and 497 differential expressed genes were remaining in Y-90 vs Y-7, Y-180 vs Y-90 and Y-180 vs Y-7. Go and KEGG enrichment analysis of these differential expressed genes were performed using Cytoscape. The overlapped up and down regulated genes and genes that shared the same enriched GO and KEGG ontology term(s) were shown in Circos plots (Fig. 2). Both up and down regulated gene expression profiles in Y-90 and Y-180 showed great similarity. Comparing with the up regulated genes, only a small number of down regulated genes that colored in light orange were unique to each of the three groups of samples.

In order to understand these enriched GO and KEGG ontology terms in detail, we extracted top 20 up and down regulated GO and KEGG clusters in the three groups and displayed them in the heatmaps (Fig. 3). No significant up or down regulated enriched term was found in Y-180 vs Y-90. As the age increased, various GO terms related to the immune system were accumulated in large numbers (Fig. 3a), such as GO: 0046649 and GO: 0002253 whereas the significant down regulated genes were almost enriched in cell replication and division (Fig. 3c), such as GO: 0006260 and GO: 0051301. For the up regulated genes enriched GO terms, 14 out of the top 20 were almost at equally significant level in both Y-90 vs Y-7 and Y-180 vs Y-7, among which the activation of lymphocytes and B cells and the regulation of various immune responses were at a relative higher significant level. The other six GO terms of the top 20 were all at a more significant level in Y-180 vs Y-7 than in Y-90 vs Y-7, which enriched in negative regulation of immune response and phagocytosis, such as GO:0050777 (negative regulation of immune response) and GO:0002683 (negative regulation of immune system process). For the down regulated GO terms, 18 of the top 20 were at almost the same level in Y-180 and Y-90.

With the age increased, disease infection and immune cell signaling pathways were enriched (Fig. 3b). Almost all terms had different expression levels in Y-90 and Y-180. In addition to Epstein-Barr virus infection and tuberculosis, the concentration of rheumatoid, staphylococcus and cytomegalovirus infection were all higher in Y-180. The concentration of antigen presentation and protein in endoplasmic reticulum and protein export were higher in Y-90, while the enrichment of chemokines, natural killer cells, T cells and other immune cells and intestinal immunity were higher in Y-180. The down-regulated pathway was basically the same between Y-90 and Y-180 (Fig. 3d). Cell cycle, alcoholism and DNA replication pathways were highly down-regulated in both groups. Signaling pathways including AMPK, mitophagy, adenine ribonucleotide biosynthesis, pyrimidine metabolism, insulin related PI3K-Akt and foxo, anemia, ferroptosis and fatty acid metabolism were less down-regulated. Down regulation of pyrimidine and fatty acid metabolism were higher in Y-180 vs Y-7 while PI3K-Akt signaling pathway was higher in Y-90 vs Y-7. Hematopoietic cell lineage and phagosome were only found in Y-90 vs Y-7.

Network analysis of known differentially expressed genes

To understand the differentially expressed gene patterns and the gene interactions involved in the significant expression modules at 7 days, 90 days and 180 days after birth in the Yorkshire spleens, we first used STEM to perform gene expression model analysis on the 1,729 genes and the results were shown in Fig. S2. A total of eight modules were obtained, of which three were significant and contained 1175 genes. Within the three significant modules, one included 666 genes that were up regulated from 7 days to 90 days and maintained the up regulation in 180 days, the other one involved 850 genes that were down regulated from 7 to 90 days and maintained the down regulation in 180 days, and the last one contained 52 genes that were up regulated from 7 days to 90 days and up regulated from 90 days to 180 days. Then, we used STRING to conduct PPI analysis on the 1175 module genes. A total of 756 interactions with combined scores no less than 0.97 were filtered to display using Cytoscape. Networks that combined no less than three edges, having average shortest path length no less than three and with no less than three neighbors within distance two were finally

displayed in Fig. 4. There was a total of 165 nodes and the center nodes genes were turned out to be *CDK1*, *PCNA* and *PLK* in Yorkshire spleen up and down regulation gene network.

To further capture how the relationships of the functional genes varied from one week to 180 days after birth, differentially expressed genes in Y-90 vs Y-7, Y-180 vs Y-90 and Y-180 vs Y-7 had been selected to analyze network enrichment using IPA (<http://www.ingenuity.com/products/ipa>) under default settings. The significant enriched networks of differentially expressed genes further supported the view of the similarity between Y-90 and Y-180 groups. Top one significant gene network of Y-90 vs Y-7, Y-180 vs Y-7 and Y-180 vs Y-90 were exhibited in Fig. 5. These networks were related to cellular function and maintenance, hematological system development and function, cell death and survival. Center genes of the networks of Y-90 vs Y-7 and Y-180 vs Y-7 were almost the same, such as *CD40*, *REL* and *BNIP3L*. Only *SPP1* and *HLA-A* were involved in the gene network of Y-180 vs Y-90. In addition to that, top regulators in three groups were obtained, including *BNIP3L*, *IL5*, *CD38*, *TGFβ1* and *BCL3* in Y-90 vs Y-7, while *BNIP3L*, *IL5*, *CD38*, *TGFβ1* and *IL4* in Y-180 vs Y-7, and *ERAP1*, *NLRC5* and *IL2RG* in Y-180 vs Y-90.

Expression profile of LncRNAs

After genome aligning, transcripts assembling and FPKM calculating (see Materials and Methods), we totally detected 219 known and 3219 putative lncRNAs, including 22 exonic_sense, 294 exonic_antisense, 232 intronic_sense, 232 intronic_antisense, 2435 intergenic, and 295 bidirectional lncRNAs (Fig. S3a). Using the same criteria for mRNA, our ranking analysis yielded a list of 64 differentially expressed lncRNAs that had RPKM value more than one, among which 51 (16↑, 35↓) were in Y-90 vs Y-7, 9 (4↑, 5↓) were in Y-180 vs Y-90, and 56 (20↑, 36↓) were found in Y-180 vs Y-7. We found that the lncRNA expression pattern in Y-90 and Y-180 also showed a high degree of consistency (Fig. S3b). Of the 51 differentially regulated lncRNAs in Y-90 vs Y-7, 43 were also found in Y-90 vs Y-7. This similarity for genes between Y-90 and Y-180 was also existed for lncRNAs. The 64-differentially expressed lncRNAs were located on 13 genes, of which only three encode proteins with characters, those were *SLA-3*, *SLA-8* and *SLA-9*. *SLA-3* was up regulated in Y-90 vs Y-7 and Y-180 vs Y-7, then down regulated from the age of 90 days to 180 days. *SLA-8* was up regulated in Y-180 vs Y-7. *SLA-9* was up regulated in Y-180 vs Y-7 and Y-180 vs Y-90. The only one differentially expressed lncRNA transcript in common for three groups was ENSSSCT00000001325 on *SLA-3* gene.

In order to understand the characteristics of these 64 differentially expressed lncRNAs in Yorkshire spleen tissues at three time points and the biological processes involved, we predicted the target genes of the 64 lncRNAs and carried out KEGG pathways enrichment analysis of those genes at each time point. Enriched pathways were obtained from Y-90 vs Y-7 and Y-180 vs Y-7 which only involved in six genes including *RRP7A*, *BMP6*, *NUP205*, *DDO*, *POLR2D*, and *COL5A2*. Each of the pathways contained only one gene. In detail, from Y-90 vs Y-7, four genes enriched in six pathways, including *RRP7A*-hsa03008 (ribosome biogenesis), *BMP6*-hsa04340 and hsa04350 (hedgehog and TGF-β signaling pathway), *NUP205*-hsa03013 (RNA transport), *DDO*-hsa00250 and hsa04146 (alanine, aspartate and glutamate metabolism and peroxisome). From Y-180 vs Y-7, five genes were enriched in six pathways. Except for *RRP7A*, *BMP6* and *DDO* genes and pathways in the above, *POLR2D* functions in five pathways including hsa00240 (pyrimidine metabolism), hsa00230 (purine metabolism), hsa03020 (RNA polymerase), etc, and *COL5A2* functions in four pathways including hsa04512 (ECM receptor interaction), hsa04974 (protein digestion and absorption), hsa04510 (focal adhesion) and hsa05146 (amoebiasis) were detected.

Taking advantage of the mRNA expression profile in this study, we further analyzed the potential target differentially expressed genes of the 64 lncRNAs. As a result, 32 of the 64 lncRNAs targeted 31 differentially expressed Ensembl genes, of which 23 genes had characters. In detail, 24, 24 and 9 lncRNAs potentially targeted 24, 21 and 6 Ensembl genes in Y-90 vs Y-7, Y-180 vs Y-7 and Y-180 vs Y-90, of which 18, 21 and 4 genes had characters (Table S2). Within the 23 characteristic genes, 17 were overlapped in Y-90 vs Y-7 and Y-180 vs Y-7 and with the same up & down regulated

state. The four genes in Y-180 vs Y-90 were not unique. More than one lncRNA was found to target more than one gene, such as the target gene *BMP6* and *TRBC1*, which were both targeted by two new lncRNAs in Y-90 vs Y-7 and Y-180 vs Y-7. Among the 23 target genes, *SLA-3* and *SLA-8* gene were acquired immunity related genes, *SCARNA2* and *SCARNA6* gene were members of small Cajal body-specific RNA family, *NCAPD3* (Non-SMC Condensin II Complex Subunit D3) is a subunit of condensin II, a large protein complex involved in chromosome condensation, *BMP6* (Bone morphogenetic protein 6) is a member of the *TGFβ* gene coded protein, *ANP32B* (Acidic Nuclear Phosphoprotein 32 Family Member B) functions in RNA polymerase binding, *AHSP* (Alpha-hemoglobin-stabilizing protein) is involved in hemoglobin assembly.

Real time PCR

Totally, six differential expressed genes and six lncRNA transcripts were randomly selected to perform RT-PCR in order to validate the sequencing data using *GAPDH* as the internal control gene for respectively. The sequencing and experiment result of these genes and lncRNAs were presented in Fig. 6. It was showed that there was a relatively good consistency between the sequencing and quantitative experiment results except for *ALB* gene and lncRNA transcript ENSSSCT0000022100.

Discussion

In this study, we identified 19,647 genes and 3,438 lncRNA transcripts in nine spleen tissues from 7-day-old, 90-day-old and 180-day-old Yorkshire pigs, among which 1,729 genes and 64 lncRNAs were found to be differentially expressed among three age groups. Compared with the genes, most lncRNAs detected were new lncRNAs and the differentially expressed lncRNA transcripts among three groups were much fewer, indicating that the study on lncRNAs in spleen still needs to be strengthened.

As the largest immune organ and blood bank, dynamic changing of the gene expression with the age in the spleen of Yorkshire pigs confirmed its function. On the time line from 7 days to 90 days after birth, immune system related GO terms and KEGG pathways enrichment were presented in the up regulated genes while cell replication and division GO terms and KEGG pathways were outstanding in the down regulated genes. This was consistent with the research results of lncRNA-mRNA profiling in Chinese local pig breed Rongchang during multiple time periods [1]. This suggested that the function of spleen in T cell and lymphocytes activation was not species-specific in farm animals [2], and the changing of gene expression with age was relatively conservative among pig breeds. Combined with the aggregation of genes from spleen tissues of Rongchang pigs at birth to 30 days after birth, we thought that the gene expression significant change in spleen tissue of pigs appeared between 30 days to 90 days after birth, that is, the growing period. And there will be very little gene expression change after 90 days. In particular, compared with 7 days, the GO term and pathway enriched in the down-regulated genes and their enrichment levels were basically the same in the spleen tissue at 90 and 180 days.

There are some similarities and also some differences for the enriched terms and pathways of up regulated genes between 90 days and at 180 days. For example, they both had highly enriched GO terms about lymphocyte activation, B cell activation, acquired immune response and immune response, but the enrichment degree of negative regulation of immune system process, immune response and phagocytosis were much higher at 180 days. At the same time, the enrichment levels of KEGG pathways were more different between 90 days and 180 days. They only had the same enrichment level in B cell signal receptor pathway.

In summary, our study provided the first mRNA- lncRNA profiles in Yorkshire spleens at three developmental stages. Our results are helpful for the study of transcriptome and functional genomics of spleen tissue in farm animals. To understand the developing spleen more clearly, stages between 30 days and 90 days after birth should be added in the future.

Declarations

Conflicts of Interest

The authors declare no competing interests.

Author Contributions

Conceived and designed the experiments: X.L., R.Q.; Performed the experiments: X.H., C.S.G.L.; Analyzed the data: R.Q., C.S.; Contributed to the writing of the manuscript: R.Q., X.L; All authors reviewed the manuscript.

Funding

This work was supported by the Pig Industry Technology System Innovation Team Project of Henan Province (S2012–06-G03).

Data Availability Statement

All the sequencing data has been uploaded in GenBank with the accession number PRJNA719591.

References

1. Che T, Li D, Jin L, Fu Y, Liu Y, Liu P, Wang Y, Tang Q, Ma J, Wang X et al: **Long non-coding RNAs and mRNAs profiling during spleen development in pig**. *PLoS One* 2018, **13(3):e0193552**.
2. Kern C, Wang Y, Chitwood J, Korf I, Delany M, Cheng H, Medrano JF, Van Eenennaam AL, Ernst C, Ross P et al: **Genome-wide identification of tissue-specific long non-coding RNA in three farm animal species**. *BMC Genomics* 2018, **19(1):684**.
3. Huang ZY, Li L, Li QQ, Liu XD, Li CC: **[The effect of lncRNA TCONS_00815878 on differentiation of porcine skeletal muscle satellite cells]**. *Yi Chuan* 2019, **41(12):1119–1128**.
4. Long K, Feng S, Ma J, Zhang J, Jin L, Tang Q, Wang X, Mai M, Xiao W, Liu L et al: **Small non-coding RNA transcriptome of four high-altitude vertebrates and their low-altitude relatives**. *Sci Data* 2019, **6(1):192**.
5. Zhang J, Xu X, Chen H, Kang P, Zhu H, Ren H, Liu Y: **Construction and analysis for dys-regulated lncRNAs and mRNAs in LPS-induced porcine PBMCs**. *Innate Immun* 2021, **27(2):170–183**.
6. Yang Y, Zhou R, Zhu S, Li X, Li H, Yu H, Li K: **Systematic Identification and Molecular Characteristics of Long Noncoding RNAs in Pig Tissues**. *Biomed Res Int* 2017, **2017:6152582**.
7. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG et al: **The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression**. *Genome Res* 2012, **22(9):1775–1789**.
8. Fatica A, Bozzoni I: **Long non-coding RNAs: new players in cell differentiation and development**. *Nat Rev Genet* 2014, **15(1):7–21**.
9. Taniguchi M, Guan LL, Basarab JA, Dodson MV, Moore SS: **Comparative analysis on gene expression profiles in cattle subcutaneous fat tissues**. *Comp Biochem Physiol Part D Genomics Proteomics* 2008, **3(4):251–256**.
10. Bao Z, Yang Z, Huang Z, Zhou Y, Cui Q, Dong D: **LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases**. *Nucleic Acids Res* 2019, **47(D1):D1034-D1037**.
11. Chen X, Yan GY: **Novel human lncRNA-disease association inference based on lncRNA expression profiles**. *Bioinformatics* 2013, **29(20):2617–2624**.

12. Szyda J, Mielczarek M, Fraszczak M, Minozzi G, Williams JL, **Wojdak-Maksymiec K: The genetic background of clinical mastitis in Holstein-Friesian cattle.** *Animal* 2019, **13(10):2156–2163.**
13. Ozdemir S, Altun S: **Genome-wide analysis of mRNAs and lncRNAs in Mycoplasma bovis infected and non-infected bovine mammary gland tissues.** *Mol Cell Probes* 2020, **50:101512.**
14. Sabara PH, Jakhesara SJ, Panchal KJ, Joshi CG, Koringa PG: **Transcriptomic analysis to affirm the regulatory role of long non-coding RNA in horn cancer of Indian zebu cattle breed Kankrej (Bos indicus).** *Funct Integr Genomics* 2020, **20(1):75–87.**
15. Han B, He Y, Zhang L, Ding Y, Lian L, Zhao C, Song J, Yang N: **Long intergenic non-coding RNA GALMD3 in chicken Marek's disease.** *Sci Rep* 2017, **7(1):10294.**
16. Kim D, **Salzberg SL: TopHat-Fusion: an algorithm for discovery of novel fusion transcripts.** *Genome Biol* 2011, **12(8):R72.**
17. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L: **Differential analysis of gene regulation at transcript resolution with RNA-seq.** *Nat Biotechnol* 2013, **31(1):46–53.**
18. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.** *Nat Biotechnol* 2010, **28(5):511–515.**
19. **Eddy SR: A new generation of homology search tools based on probabilistic inference.** *Genome Inform* 2009, **23(1):205–211.**
20. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, **Gao G: CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine.** *Nucleic Acids Res* 2007, **35(Web Server issue):W345-349.**
21. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5(7):621–628.**
22. **Ernst J, Bar-Joseph Z: STEM: a tool for the analysis of short time series gene expression data.** *BMC Bioinformatics* 2006, **7:191.**

Tables

Table 1. Genomic region distribution of the output of clean reads.

Group	Total clean reads	Total mapped reads	Unique mapped	Region distribution					
				Gene	Intergenic	Coding	Splicing	Intron	Non-coding [†]
Y-7	123,134,202	98,937,321	85,814,748	74475278	24,462,043	36445772	2343397	24659691	13369815
Y-90	98,709,782	78,279,022	68,422,219	55965083	22,313,939	21591787	1048206	23152959	11220337
Y-180	129,080,634	102,646,848	89,742,389	73586197	29,060,651	28632107	1426424	30036857	14917233

[†] Non-coding includes 5'UTR, 3'UTR and noncoding RNA regions.

Figures

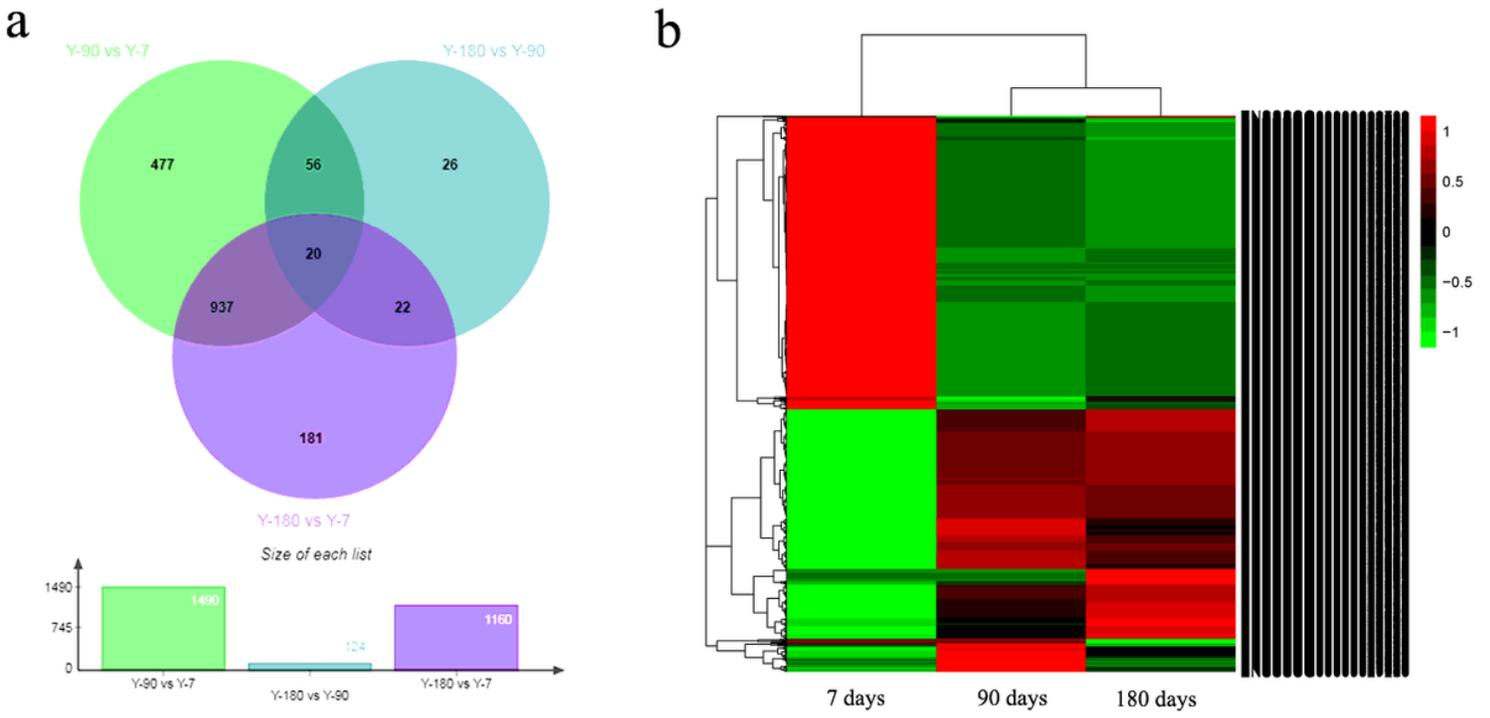


Figure 1

Venn diagram and heat map of 1729 differentially expressed genes. a Venn diagram; b Heat map.

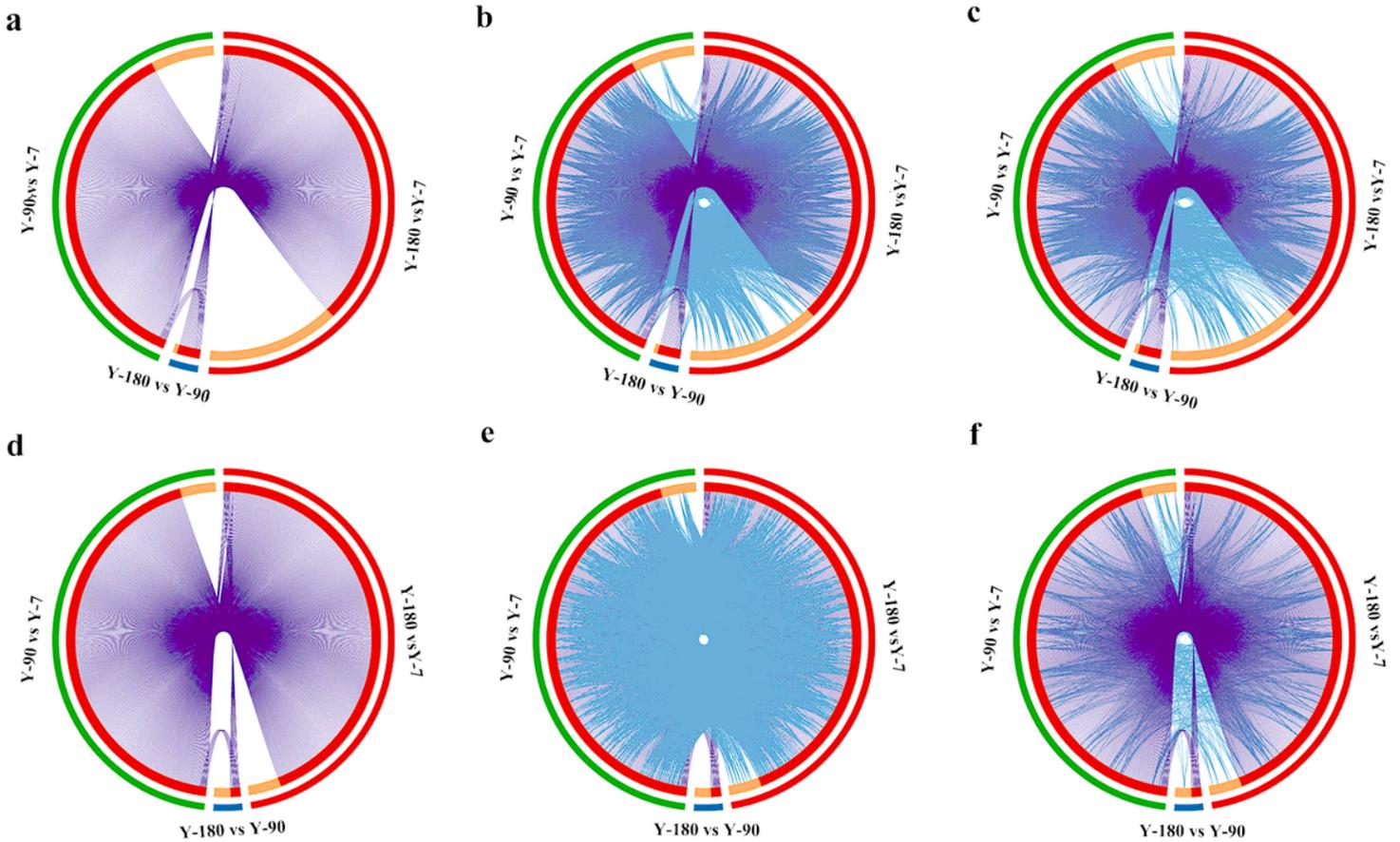


Figure 2

Circos plot of how 1729 differentially expressed genes overlapped among three groups. Purple curves link identical up (a) and down (d) regulated genes at the shared gene level. At the shared term level, blue curves link up and down regulated genes that belong to the same enriched up (b) and down (e) regulated GO terms and the same enriched up (c) and down (f) regulated KEGG terms. The inner circle represents up and down regulated gene lists, where hits are arranged along the arc. Genes that hit multiple lists are colored in dark orange, and genes unique to a list are shown in light orange.

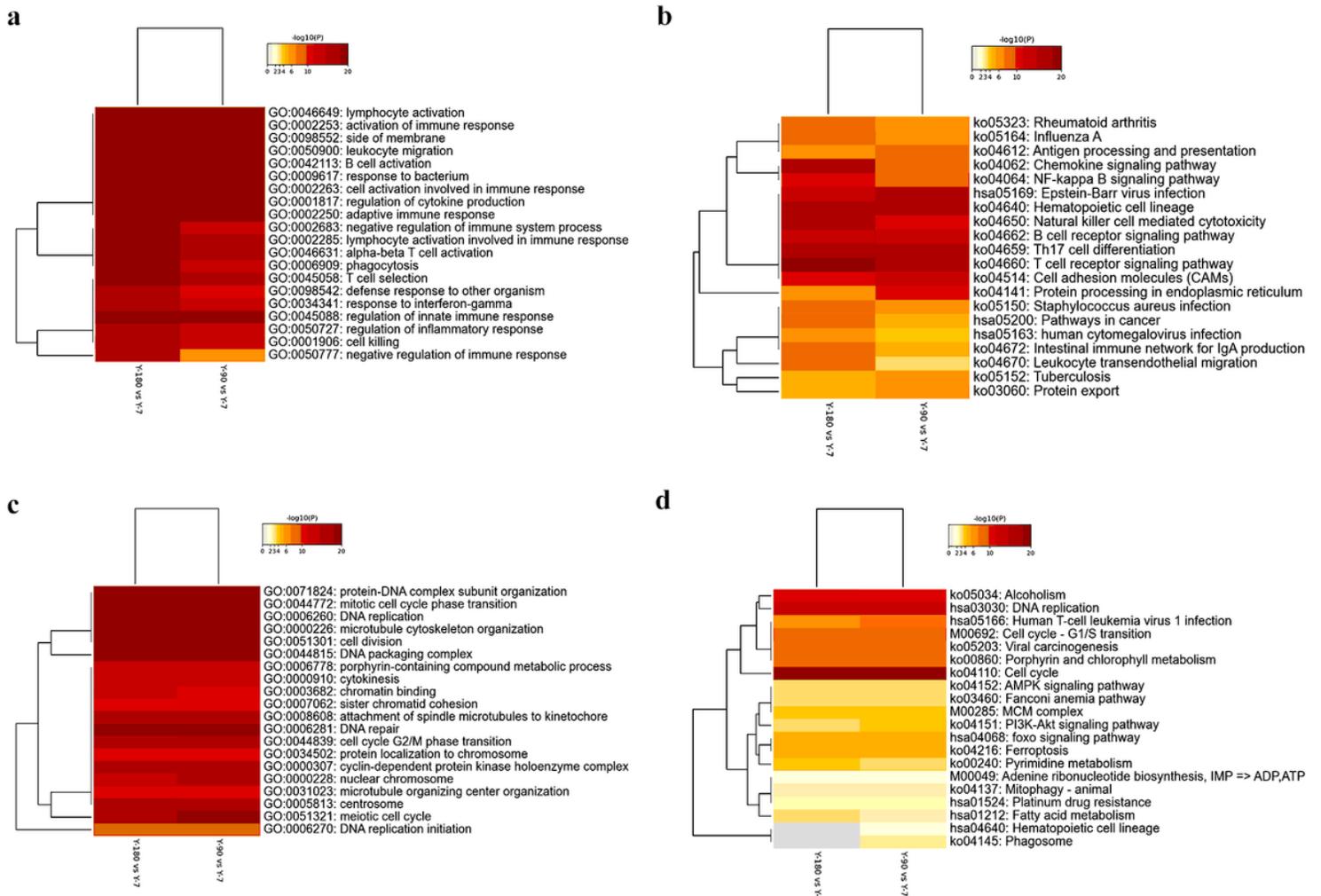


Figure 3

Heatmap of top 20 enriched up and down regulated GO terms and KEGG pathways. Up regulated GO (a) and KEGG (b), down regulated GO (c) and KEGG (d). The heatmap cells are colored by their p-value; white cells indicate the lack of enrichment for that term in the corresponding gene list.

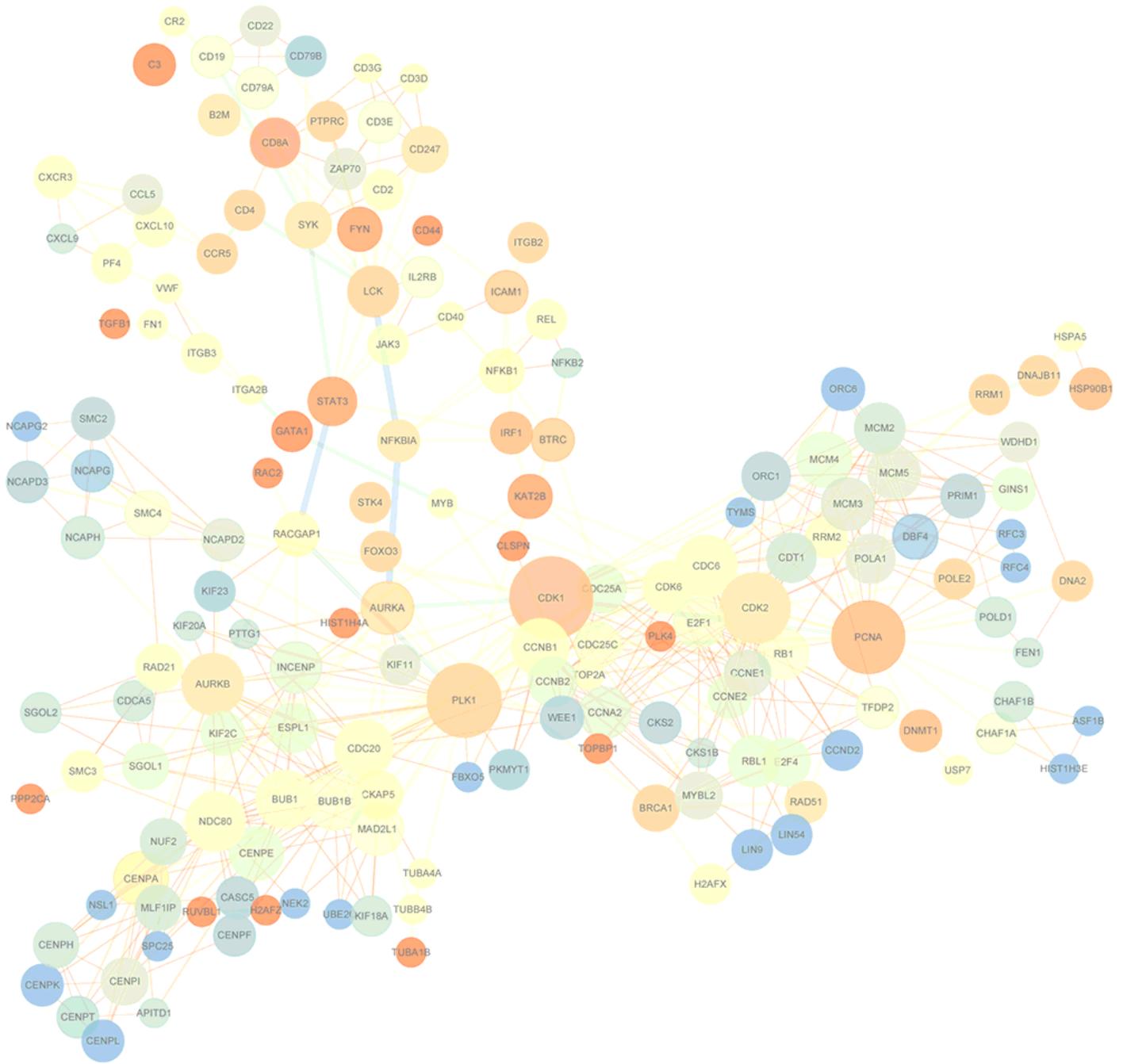


Figure 4

Protein-protein interaction network of differentially expressed module genes.

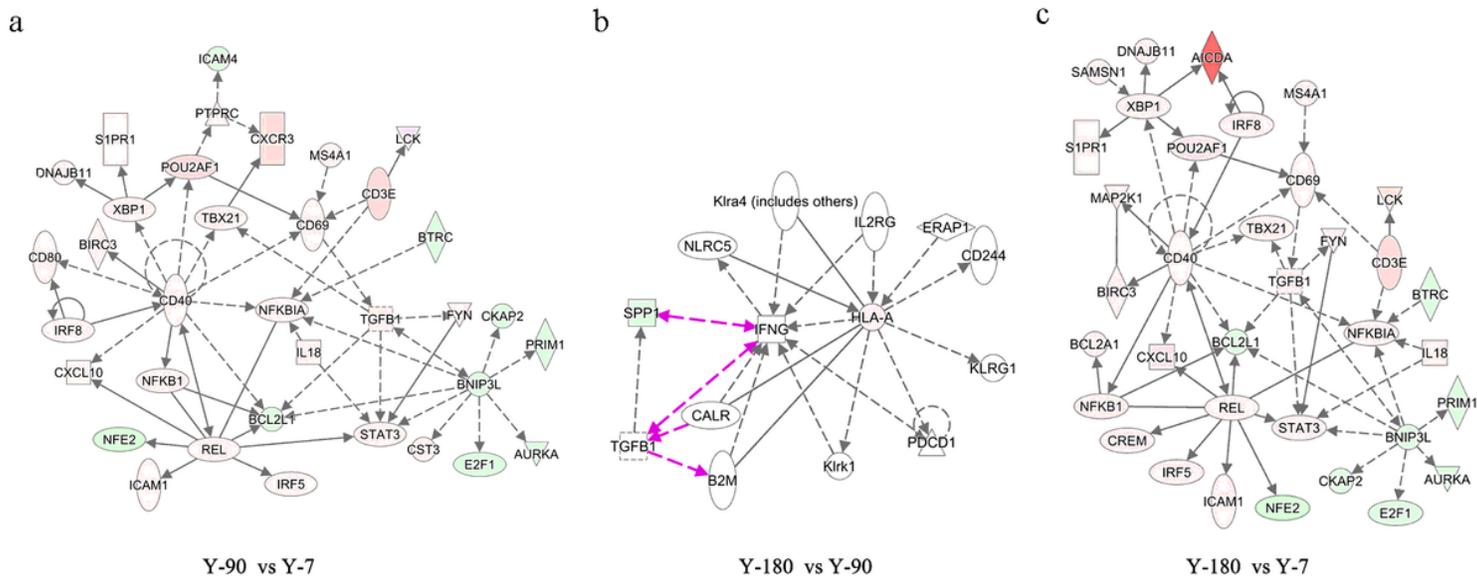


Figure 5

Top one significant IPA network of differentially expressed genes in three groups.
Sequencing (E) -Quantitative (O)

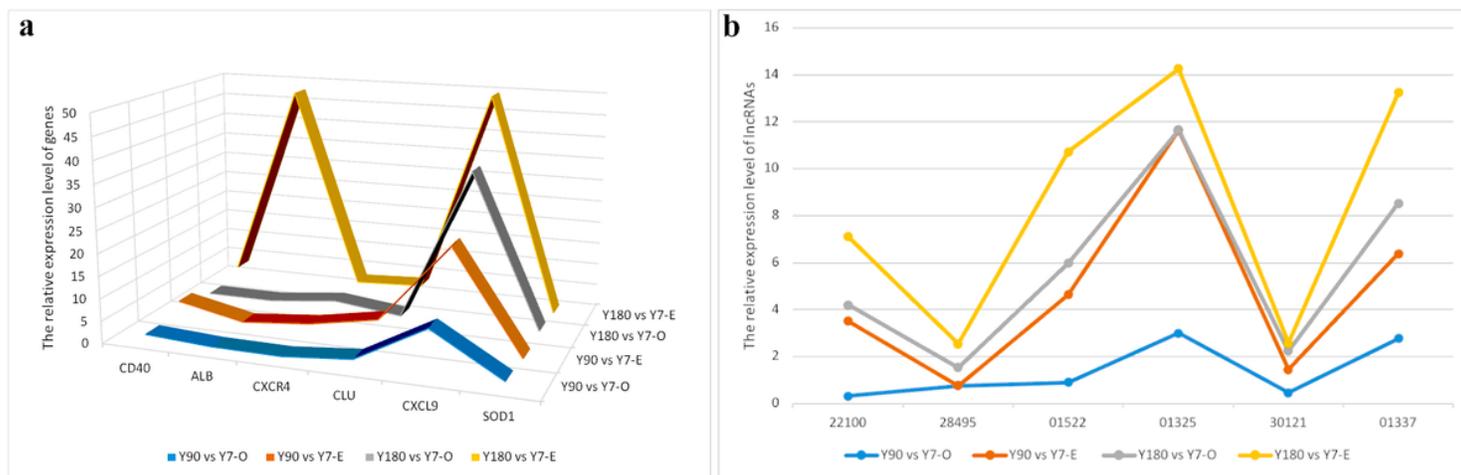


Figure 6

Comparison of RNA sequencing and RT-PCR results for six genes and six lncRNA transcripts. a gene; b lncRNA transcripts (ENSSSCT00000001522, -01325, -22100, -30121, -28495).

Supplementary Files

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

- [FigureS1.tif](#)
- [FigureS2.tif](#)
- [FigureS3.tif](#)
- [TableS1.docx](#)
- [TableS2.docx](#)