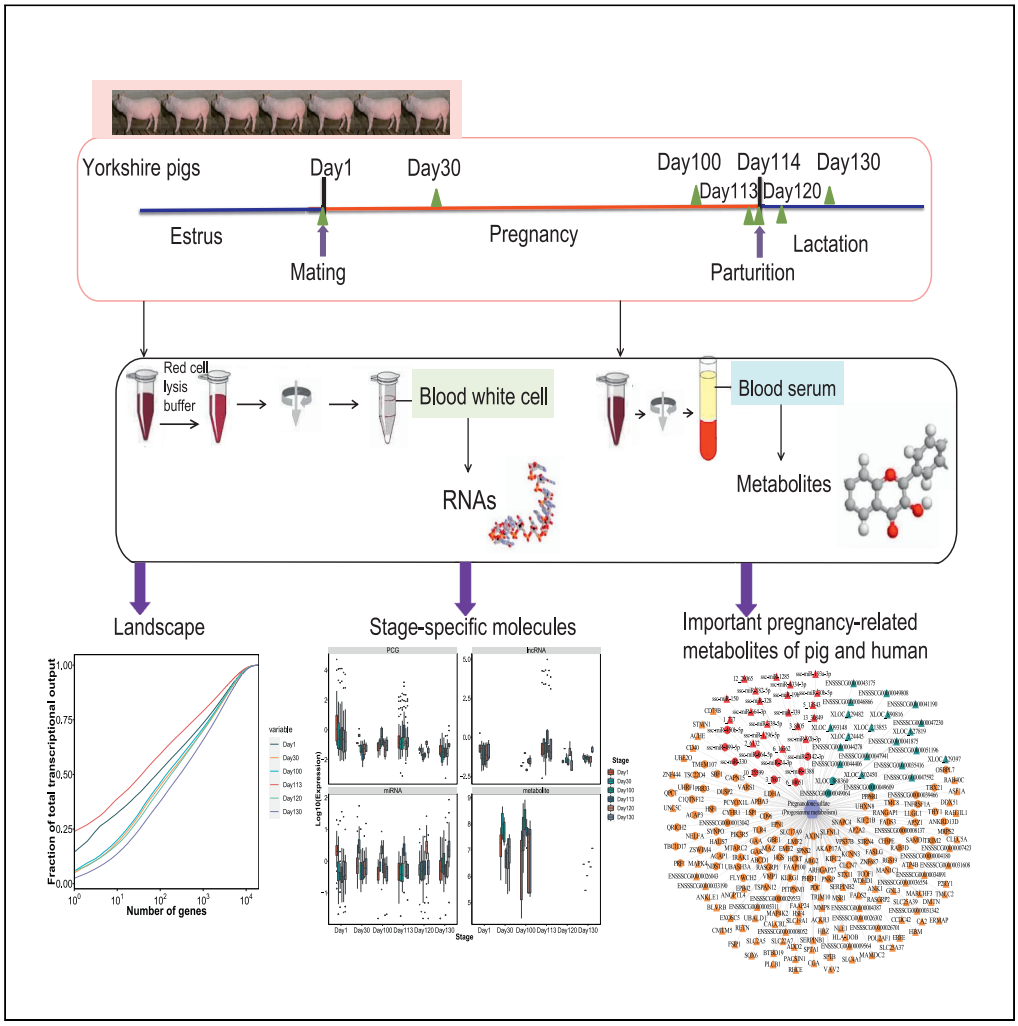
iScience

CellPress
OPEN ACCESS

Article

Integrated analysis of transcriptome and metabolome revealed biological basis of sows from estrus to lactation



Lijun Shi, Huihui Li, Xiaoyu Huang, ..., Ligang Wang, Hua Yan, Lixian Wang

shilijun01@caas.cn (L.S.)
iaswlx@263.net (L.W.)

Highlights

Landscape of the pig transcriptome and metabolome from estrus to lactation

Stage-specific PCGs, lncRNAs, miRNAs, and metabolites

Integrated analysis of metabolome and transcriptome in development stages

Key metabolite markers both human and pig pregnancy

Shi et al., iScience 26, 105825
January 20, 2023
https://doi.org/10.1016/j.isci.2022.105825

iScience

CellPress
OPEN ACCESS

Article

Integrated analysis of transcriptome and metabolome revealed biological basis of sows from estrus to lactation

Lijun Shi,[1,*] Huihui Li,[1] Xiaoyu Huang,[1] Ze Shu,[1] Jingna Li,[1] Ligang Wang,[1] Hua Yan,[1] and Lixian Wang[1,2,*]

SUMMARY

Characterization of molecular mechanisms underlying pregnancy development of sows is important for the genetic improvement of pig breeding traits, and also provides resources for biomedical research on human pregnancy diseases. However, the transcriptome and metabolome across multiple developmental stages of sow pregnancy were still lacking. In this study, we obtained 84 distinct RNA sequencing and 42 metabolome datasets of pig blood across six development stages from estrus to lactation. We confirmed the initial sequence and exonic structural features, stage-specific molecules, expression or accumulation pattern of molecules, the regulatory mechanism of transcriptome and metabolome, and important pregnancy-related metabolites both in pigs and humans. In conclusion, we proposed the key differences among the stages of sows from estrus to lactation in RNAs and metabolites and put forward key markers. These data results were expected to provide essential resources for pig breeding and biomedical research on human pregnancy disease.

INTRODUCTION

It is helpful to employ animal models for validating medicinal medications and better understanding the development of disease. In-depth research has been done using the pig as a model for neurological, cardiovascular, and metabolic problems.[1] Animal reproductive features have economic significance and are largely sex specific.[2] Many crucial events leading to the phenotypes of reproductive traits occur during the pregnancy and breastfeeding periods. In the life of a pregnant sow, each day is never the same, and every day marks the piglet's growth. The endocrine status of the pregnant female and her nutritional status are critical for successfully establishing and maintaining pregnancy.[3] A dynamic, carefully planned process, the metabolism, during pregnancy can have disastrous effects on both the mother and fetus if it fails.[4] Before giving birth to a litter of piglets, a sow goes through a gestation period that lasts roughly three months, three weeks, and three days. Pig pregnancies are complicated to establish and maintain and depend on the conceptus and sow's uterine endometrium communicating effectively, in which, the stages of sows from estrus to lactation are often divided into eight parts, including days 0–9 (estrus and mating period), days 10–35 (implantation period), days 36–110 (embryonic stage), and days 111–114 (last few days before farrowing), 0–7 days postpartum (primary stage of breastfeeding), 8–20 days after delivery (mid-lactation period), and after 21 days after delivery (late lactation).[5–8]

The significant changes in the transcriptome at different life stages may result in age-specific susceptibilities to disease, reproduction, or exposure to contaminants. Several community-wide research studies have been conducted to create a database of genes expressed throughout mice, humans, rats, and cattle normal development,[9–13] such efforts are less advanced for the pig, especially sows in gestation. Metabolites are the end product of gene transcription and protein expression, and also have the power to control these processes.[14] Analysis of the combined transcriptome and metabolome provides the information to interpret key molecules from genes to phenotypes. In pigs, some combined transcriptome and metabolome studies for important economic traits or drug reactions have been reported, such as feed efficiency,[15] hypoxia-related signature,[16] and the responses of pigs orally inoculated with amoxicillin solution,[17] while studies on metabolites in various stages of pregnancy are still quite a few. Therefore, characterizing the molecular mechanisms of the reproduction and lactation process of pigs based on the

[1]Institute of Animal Science, Chinese Academy of Agricultural Sciences, No. 2 Yuanmingyuan West Road, Haidian District, Beijing 100193, China

[2]Lead contact

*Correspondence: shilijun01@caas.cn (L.S.), iaswlx@263.net (L.W.)

https://doi.org/10.1016/j.isci.2022.105825

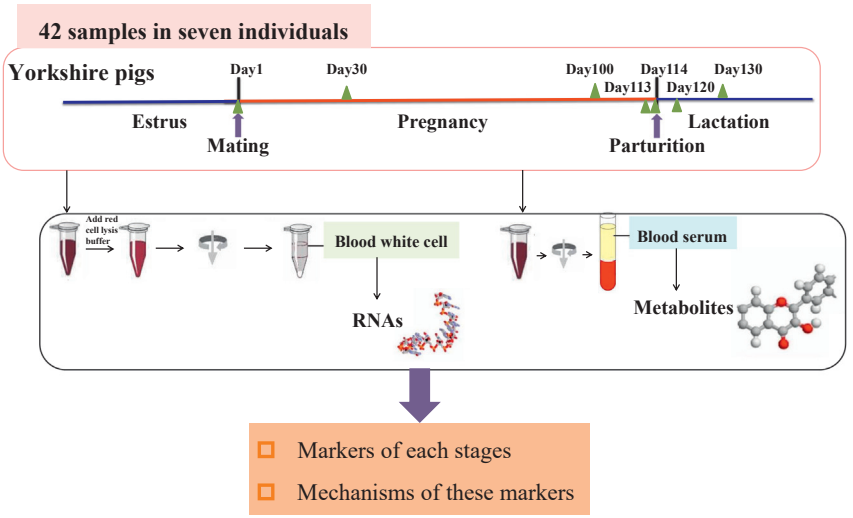


Figure 1. Details of the study design

Blood samples were collected from six development (day 1, 30, 100, and 113 of pregnancy, and day 120 and 130 after parturition) of seven pigs, and used for transcriptome and metabolome sequencing.

transcriptome and metabolome data will be very necessary and can be a useful model for human pregnancy disorders.

In this study, to comprehensively characterize the key molecular underpinning the sow developmental stage from estrus to lactation, we sequenced the whole transcriptome and metabolome in seven Yorkshire pigs. We obtained 126 sequence datasets (Figure 1), and perform the bioinformatics analyses for the RNAs and metabolites. These data results will be a very useful resource to research pig breeding and human pregnancy disease.

RESULTS

Overview of the landscape of the pig transcriptome and metabolome from estrus to lactation

We obtained 84 RNA-seq datasets of mRNAs, lncRNAs, and miRNAs throughout six stages in seven Yorkshire pigs, which produced 3,910,124,178 clean reads (Table S1). A total of 14,810 protein-coding genes (PCGs), 5,058 lncRNAs (2,701 known and 2,357 novel), and 1,889 miRNAs (340 known and 1,549 novel) were expressed in the pig blood samples from estrus to lactation (Table S2). For each stage, the molecules were expressed with expression values $\geq$0.1 in at least two samples, and six stages observed differences in the numbers of transcriptomic molecules (Table S3 and Figure 2A). In detail, stage day 113, the last few days before farrowing, had the least amount of PCGs, lncRNAs, and miRNAs that were expressed in comparison with the other five stages (Figure 2A). In all six stages, 13,295 genes, 4,209 lncRNAs, and 1,242 miRNAs were commonly expressed molecules. Additionally, we obtained 42 metabolome datasets for metabolites and identified 629 metabolites including 225 untargeted lipidomic metabolites (Table S2). Hierarchical clustering analysis (HCA, Figure 2B) of metabolites of all 42 samples demonstrated that different stages exhibited unique metabolite profiles. Morphologically or functionally similar stages, such as d ay 1 versus (vs) day 30, day 100 vs day 113, and day 120 vs day 130, were clustered more closely to each other, indicating that these stages shared a more consistent metabolite profile with the other stages (Figure 2B). In addition, day 120 and day 130 were the stages of lactation, and they were distinguished from the other stages of pregnancy (Figure 2B). We also calculated the correlation coefficient for every two groups and found that day 1 and day 30, day 100 and day 113, and day 120 and day 130 had higher correlations (0.9721, 0.9567, and 0.9885) than other groups (Table S4). While, the HCA results of PCG, lncRNA, and miRNA did not show obvious clustering distinctions between different stages (Figures S1A–S1C).

For PCGs, lncRNAs, miRNAs, and metabolites identified in all stages, the principal component analysis (PCA) was performed, and the results of metabolites revealed the similarity between biological replicates and differences between every two stages (Figure S2), while the observation of PCGs, lncRNAs, and miRNAs did not (Figures S3–S5). Furthermore, the PCA of PCGs, lncRNAs, and miRNAs between any two

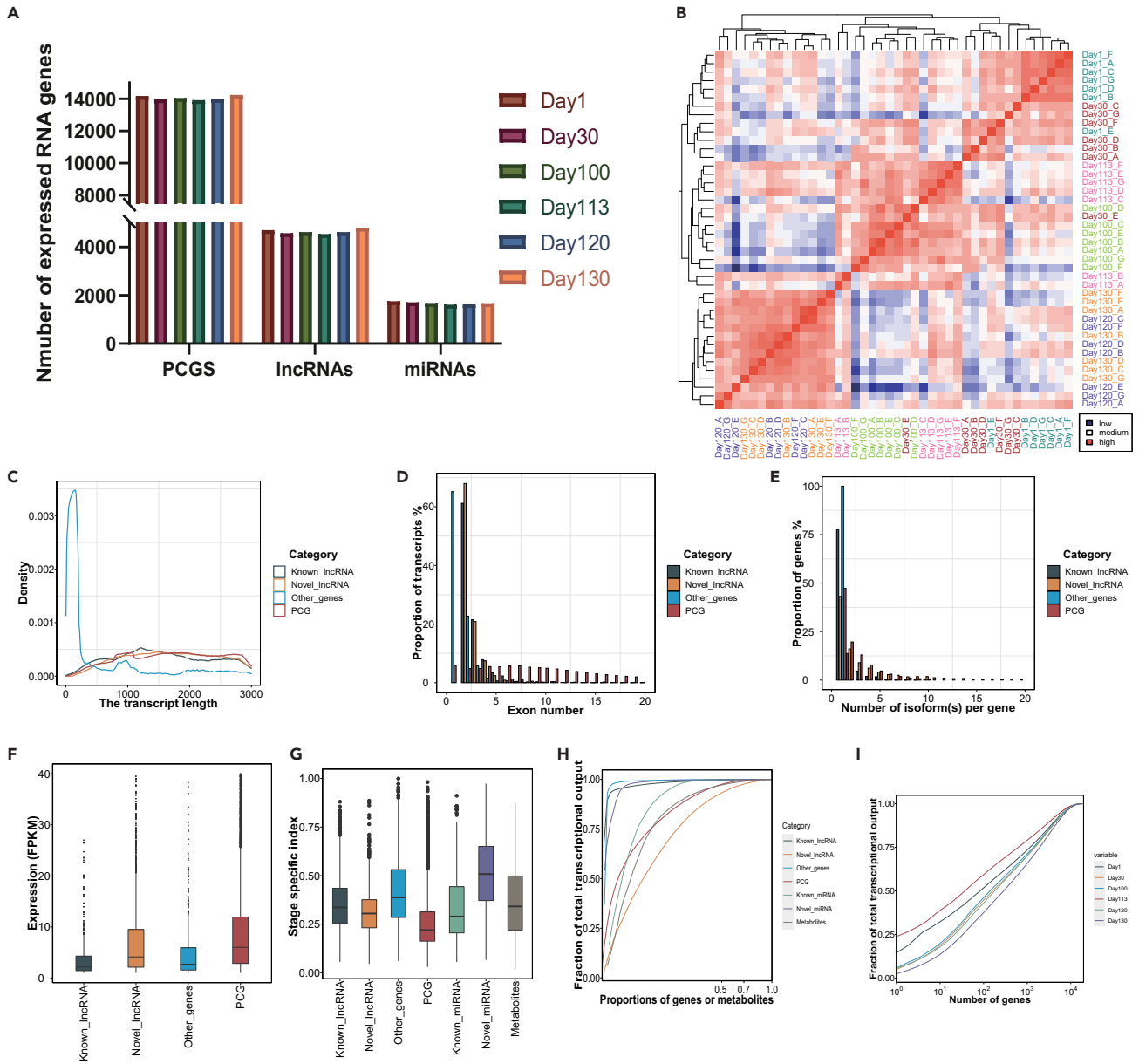


Figure 2. The landscape of sow blood transcriptome and metabolome

(A) Number of expressed PCGs, lncRNAs, and miRNAs in each stage.

(B) Hierarchical clustering analysis (HCA) of 629 metabolite abundance of 42 sow blood samples. Distance between samples was defined as distance = 1 - Spearman correlation.

(C–F) Feature comparisons of transcript length, number of exons, number of isoforms per gene, and gene expression between lncRNA and PCG.

(G) Stage specificity of transcripts shown through tissue specificity index.

(H) Cumulative distribution of the average percentage of entire transcription or metabolites abundance contribute. The mean values for samples from the same category were shown by lines. The proportion of transcripts arranged from the highest to lowest abundance was shown by the x axis. The accumulated fraction of transcripts relative to all transcripts was indicated by the y axis.

(I) Cumulative distribution of the average percentage for all transcription that genes contributed.

stages (Figures S2–S5) observed that some of them could be separated from each other, such as day 1 and day 113, day 1 and day 130, day 30 and day 113, and day 30 and day 120. We did the variance partition analysis and calculated the variance of each PCG, lncRNA, miRNA, and metabolite explained by stage, individual, and residuals. We found that the variance explained by stage was as much as that by the individual for

PCG, lncRNA, and miRNA, which may explain the mixed pattern of HCA analysis and PCA analysis that occurred in PCG, lncRNA, and miRNA (Figure S1D). While for metabolite, the variance explained by stage was far less than that explained by the individual; then, we could clearly distinguish samples by stage in HCA and PCA analyses (Figure S1D).

By investigating the initial sequence and exonic structural features, we found that lncRNAs had shorter lengths, fewer exons, lower transcript numbers, and lower expression compared to PCGs (Figures 2C–2F). Through calculating the stage specificity of genes and metabolites, we detected that the novel miRNA had a higher stage-specific index than other molecules, and the PCG had a lower stage-specific index (Figure 2G). For most molecule groups, a small number of hundred genes or metabolites were responsible for approximately 50% of the transcription abundance (Figure 2H). Furthermore, the stage day 113 had the highest fraction of total transcriptional output by genes (Figure 2I).

The development-dependent molecules expression and accumulation patterns

We used the DEseq2 package in R to identify transcriptome molecules that were DE (differentially expressed) RNAs between any two stages. In total, 3,652 DE PCGs, 694 lncRNAs, and 184 miRNAs were identified, and the overall DE transcriptome molecules were shown in (Table S5 and Figures 3A–3C). The number of DE RNA molecules was significantly different in any pair of stages compared. In detail, the adjacent stages day 1 vs day 30, day 100 vs day 113, day 113 vs day 120, and day 120 vs day 130 had fewer DE RNAs than the other compared groups. Particularly, there were no DE lncRNAs and miRNAs on day 1 vs day 30. Most DE PCGs and lncRNAs were detected on day 30 vs day 130 (1,389 DE mRNAs and 259 DE lncRNAs) and day 100 vs day 130 (2,165 mRNAs and 430 DE lncRNAs). By the pairwise comparison of metabolites between any two stages, there were 515 significantly different metabolites (SDMs) identified (Table S5 and Figure 3D), and the comparisons between other stages and mid-stage of milk production had more SDMs, such as day 1 vs day 130, day 30 vs day 130, and day 100 vs day 130.

For the DE lncRNAs, 4,040 strong lncRNA-mRNA correlation pairs (correlation >0.9 or < -0.9, and $p < 0.05$) were predicted, which contained 235 lncRNAs and 1,015 PCGs (Table S6). By screening the PCGs within 100 kb upstream and downstream of the DE lncRNAs, we detected 2,421 genes *cis*-target for 666 DE lncRNAs (Table S6). In total, we confirmed 3,312 target genes of 674 DE lncRNAs. For the miRNAs, we predicted 2,387 genes targeted for 106 DE miRNAs (Table S6).

To assess the development-dependent transcriptomic and metabolomic activities throughout the development of sow pregnancy, we conducted a stage-course differential PCG, lncRNA, and miRNA expression and metabolite accumulation analysis by comparing any two adjacent development stages. We showed the overall development-dependent patterns in Figures 3E–3H and Table S7. In total, we observed 15, 8, 4, and 21 expression or content patterns during day 1-day 30-day 100-day 113-120-day 130 for PCGs, lncRNAs, miRNAs, and metabolites, respectively, in which, more PCGs were involved in the DMMMM and MDMMM patterns, and more miRNAs and metabolites were involved in MMMDM pattern. The vast majority of PCGs, lncRNA, miRNAs, and metabolites stayed unchanged throughout the stages (over 73%), and no molecules' expressions continuously changed during the stage, either increasing (UUUUU) or decreasing (DDDDD). We performed the functional annotation of molecules in the same pattern and found that they were distinct across different patterns (Table S7). For instance, the PCGs at MUMDM and MMUMM patterns might respectively prepare for embryo implantation and lactation, and they were individually involved in hemopoiesis and structural constituent of the cytoskeleton, and protein and lipid metabolism process (Table S7).

Additionally, we conducted the time-series analysis to comprehensively reveal the development-dependent transcriptomic and metabolomic molecules, and we respectively found nine, 11, eight, and eight profiles with significant time-series feature in PCGs, lncRNA, miRNA, and metabolite (Figures 4 and S6, and Table S8). The molecules containing 107 PCGs, 98 lncRNAs, and 61 metabolites in profile 39 kept rising from day 1 to day 130, and 150 PCGs, 64 miRNAs, and 26 metabolites continued to decline throughout the period. In these significant profiles, 1,586 PCGs, 1,431 lncRNAs, 419 miRNAs, and 211 metabolites were included. Among them, 575 PCGs, 282 lncRNAs, and 70 miRNAs were DE molecules, and 174 metabolites were SDMs.

Stage-specific PCGs, lncRNAs, miRNAs, and metabolites

The PCGs, lncRNAs, miRNAs, and metabolites with more than four times higher expression in a given stage over any other stages were considered as the stage-specific molecules. In total, we identified 123, 23, 188,

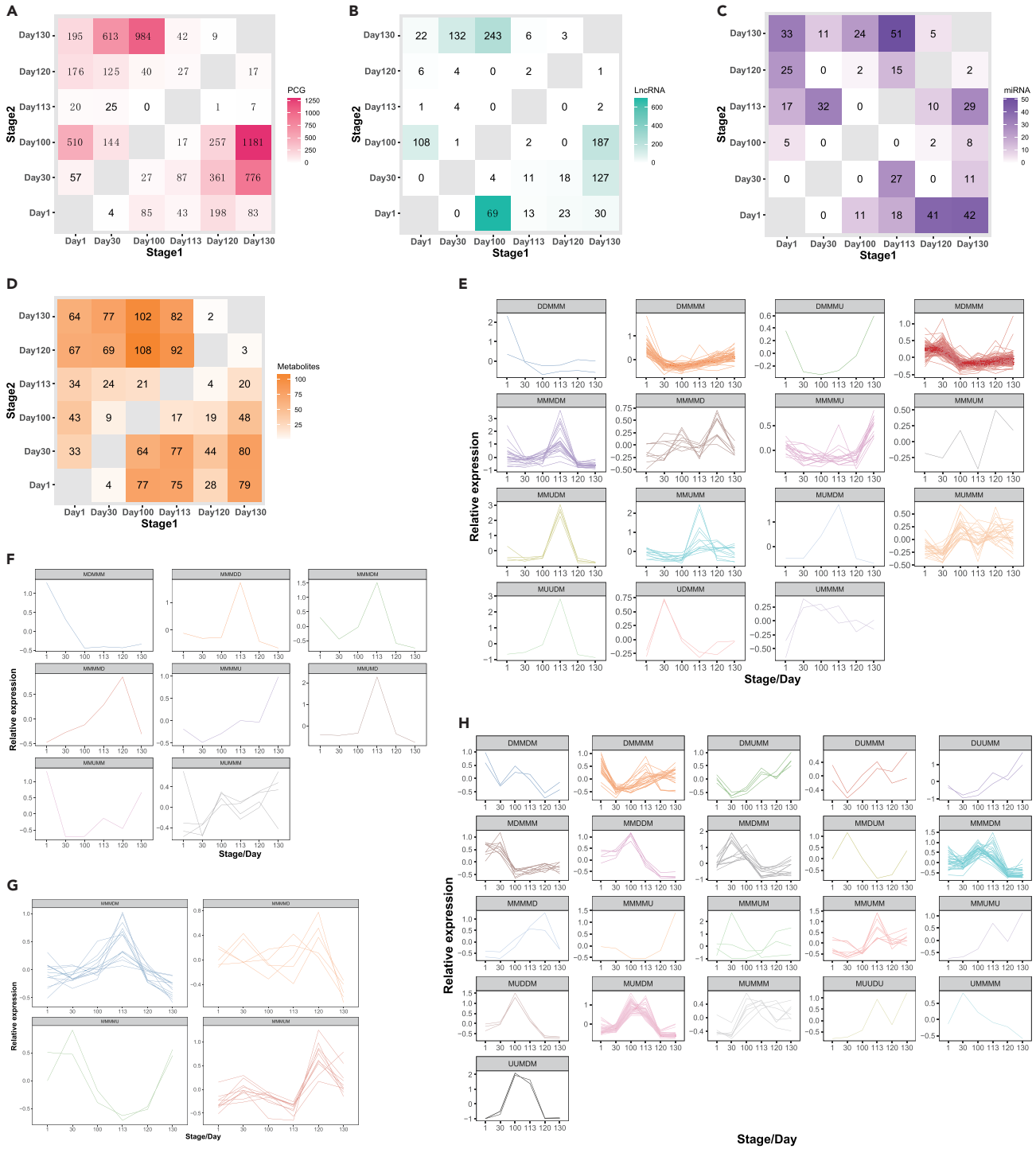


Figure 3. Results of Differentially expressed (DE) analyses

(A–C) respectively indicated the number of DE PCGs, lncRNAs and miRNAs between any two stages, and the digital in box was the number.

(D) Number of significantly different metabolites (SDMs) between any two stages, and the digital in box was the number.

(E–G) respectively indicated 15, eight, and four expression patterns from day 1 to day 130 for DE PCGs, lncRNAs, and miRNAs.

(H) 21 SDMs content change patterns from day 1 to day 130.

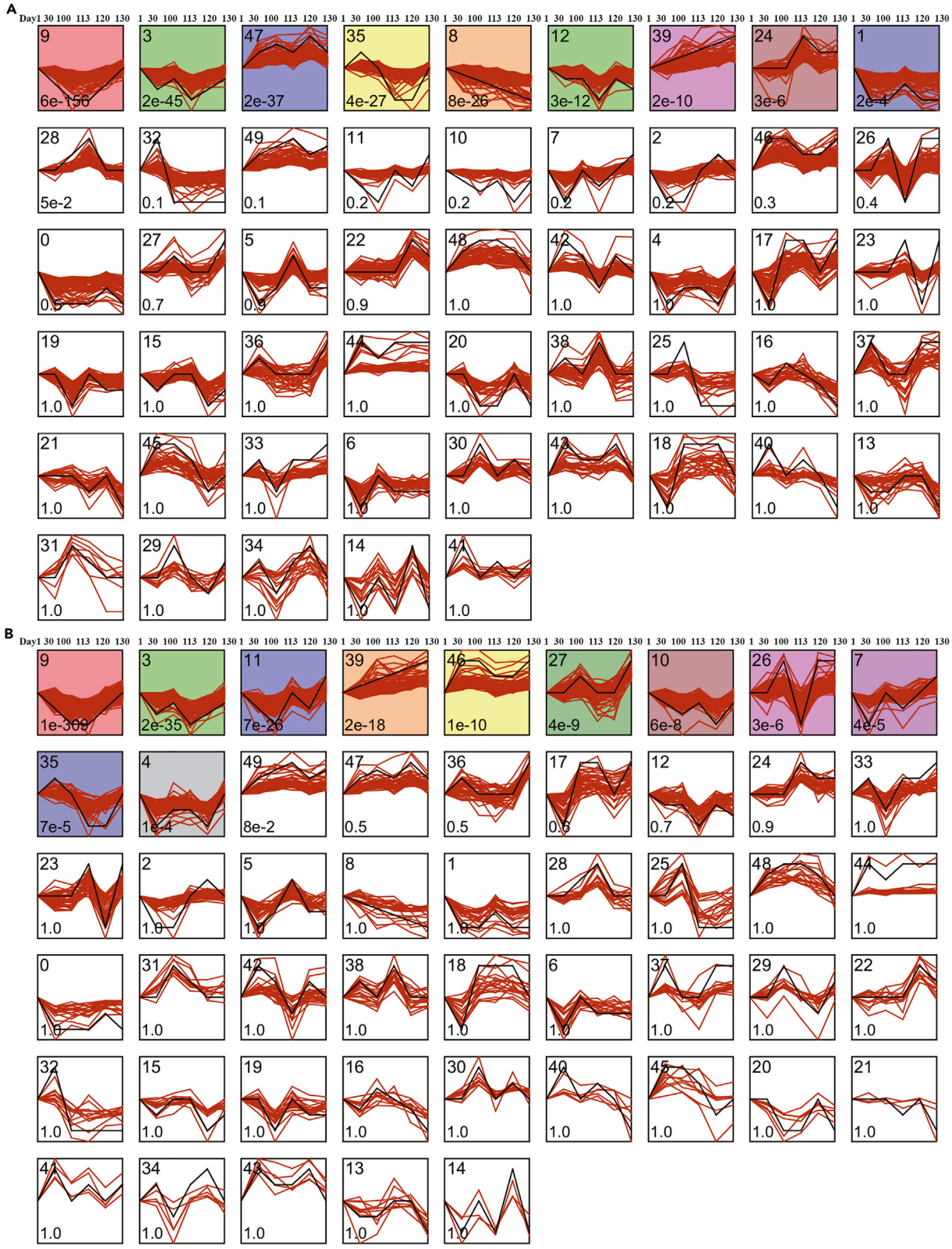

A
Day1 30 100 113 120 130
9 6e-156
3 2e-45
47 2e-37
35 4e-27
8 8e-26
12 3e-12
39 2e-10
24 3e-6
1 2e-4
28 5e-2
32 0.1
49 0.1
11 0.2
10 0.2
7 0.2
2 0.2
46 0.3
26 0.4
0 0.5
27 0.7
5 0.9
22 0.9
48 1.0
42 1.0
4 1.0
17 1.0
23 1.0
19 1.0
15 1.0
36 1.0
44 1.0
20 1.0
38 1.0
25 1.0
16 1.0
37 1.0
21 1.0
45 1.0
33 1.0
6 1.0
30 1.0
43 1.0
18 1.0
40 1.0
13 1.0
31 1.0
29 1.0
34 1.0
14 1.0
41 1.0
B
Day1 30 100 113 120 130
9 1e-309
3 2e-35
11 7e-26
39 2e-18
46 1e-10
27 4e-9
10 6e-8
26 3e-6
7 4e-5
35 7e-5
4 1e-4
49 8e-2
47 0.5
36 0.5
17 0.6
12 0.7
24 0.9
33 1.0
23 1.0
2 1.0
5 1.0
8 1.0
1 1.0
28 1.0
25 1.0
48 1.0
44 1.0
0 1.0
31 1.0
42 1.0
38 1.0
18 1.0
6 1.0
37 1.0
29 1.0
22 1.0
32 1.0
15 1.0
19 1.0
16 1.0
30 1.0
40 1.0
45 1.0
20 1.0
21 1.0
41 1.0
34 1.0
43 1.0
13 1.0
14 1.0

Figure 4. 50 patterns encompassing the PCGs and lncRNAs during different stages using STEM

(A) Result of PCGs.

(B) Result of lncRNAs. Profiles ordered by significance. The number in the top left-hand corner of a profile box was the profile ID number, and the number in the bottom left-hand corner of a profile box is the p-value. The colored profiles had a statistically significant number of genes assigned. Non-white profiles of the same color represent profiles grouped into a single cluster.

and 12 stage-specific PCGs, lncRNAs, miRNAs, and metabolites, respectively (Figures 5A–5D, Table S9). The stage of day 113 had the most specific PCGs (65) and lncRNAs (11), and day 30 and day 120 had the most specific miRNAs (Figures 5A and 5B, Table S9).

Furthermore, the Phi coefficient was calculated to find the exclusive PCGs, lncRNAs, and miRNAs in each stage. Mean square contingency coefficient (Phi coefficient) ≥ 0.6 indicates that the molecule was expressed in (nearly) all samples from the stage (it was exclusive of the stage), and not almost expressed in other stages. A molecule with Phi coefficient ≤ -0.6 indicated that it was (nearly) not expressed in samples from the stage and it was almost expressed in other stages. Thousands of DE molecules were identified, while the stage-preferential expression molecules were only 11 (Figure S7). We also considered the 11 molecules as the stage-specific gene types (Table S9), in which, six were exclusive to stage day 130, and four of them were lncRNAs.

We performed the functional annotation for the stage-specific molecules and found that they had a significant relation with immune response and oxygen carrier activity (Table S9). In addition, the stage-specific molecules had different enrichments across development times (Table S9). For instance, the molecules at stage day 1 were strongly involved in the lipid and glycogen metabolic process, and the molecules at stages day 113 and day130 were significantly associated with protein and amino acid metabolism (Table S9).

Molecule co-expression network analysis

To find co-expression modules associated with sow gestation and lactation, we used PCGs, lncRNAs, miRNAs, and metabolites to conduct an unsigned weighted gene network analysis (WGCNA). A total of 26 co-expression modules including PCGs, lncRNAs, miRNAs, and metabolites were discovered (Figure 6A). Figure 6B displayed the number of molecules in each module. The expressions of six modules were noticeably (Bonferroni p-*value* < 0.05, Figure 6C) stage specific, namely, two, one, one, one, and two co-expression modules for days 1, 30, 100, 113, and 130, respectively (Figure 6A). According to the functional annotation, the stage-specific modules were strongly associated with signal transduction, metabolic process, immune responses, and cellular processes (Figure 6C and Table S10). The Darkgreen was highly associated with stages days 1 and 30, in which, 19 LysoPC metabolites were included. The purple module had 97 TG metabolites significantly associated with day 100, which might be needed for labor and lactation. 31 PC metabolites in the Lightgreen module were significantly related to milk production (day 130). Numerous PCGs, lncRNAs, miRNAs, and metabolites with unclear function were co-expressed with molecules annotated in the same modules, serving as helpful resources for pigs' molecular annotation.

Construction of regulatory networks of PCGs, lncRNAs, and miRNAs

We studied whether the stage-specific PCGs were strongly involved in targets of certain lncRNAs and miRNAs based on target genes for DE lncRNAs and miRNAs between the comparisons in any two stages. Table S11 observed the detailed summary statistics. These enriched lncRNAs and miRNAs might strongly influence sow pregnancy development. For instance, the targets of lncRNAs ENSSSCG00000041596 and XLOC_522486 were specifically expressed on day 113 and day 1, respectively (Figure 7A). The targets of lncRNA ENSSSCG00000041596 were significantly engaged in chaperonin-containing T-complex and complement activation (Table S11). The targets of miRNA ssc-miR-7142-3p specifically expressed at day 113 (Figure 7A) were significantly enriched in protein-glutamine gamma-glutamyltransferase activity, positive regulation of ATPase activity, and peptide crosslinking (Table S11).

We also conducted a competing endogenous RNA investigation for identifying the regulatory influence of lncRNAs on the expression of PCGs by mediating miRNAs during the stages. Table S12 provided a detailed description of the significantly ($p < 0.05$) negative correlation of miRNA-target pairs, containing 632 DE PCGs, 349 DE lncRNAs, and 73 DE miRNAs. Out of the 73 DE miRNAs, 11 miRNAs were found in the significant module Lightcyan connected with day 113, and one miRNA in the Darkgrey module was strongly associated with day 130. The Figure 7B displayed the lncRNA-miRNA-mRNA interactions for these 12 miRNAs.

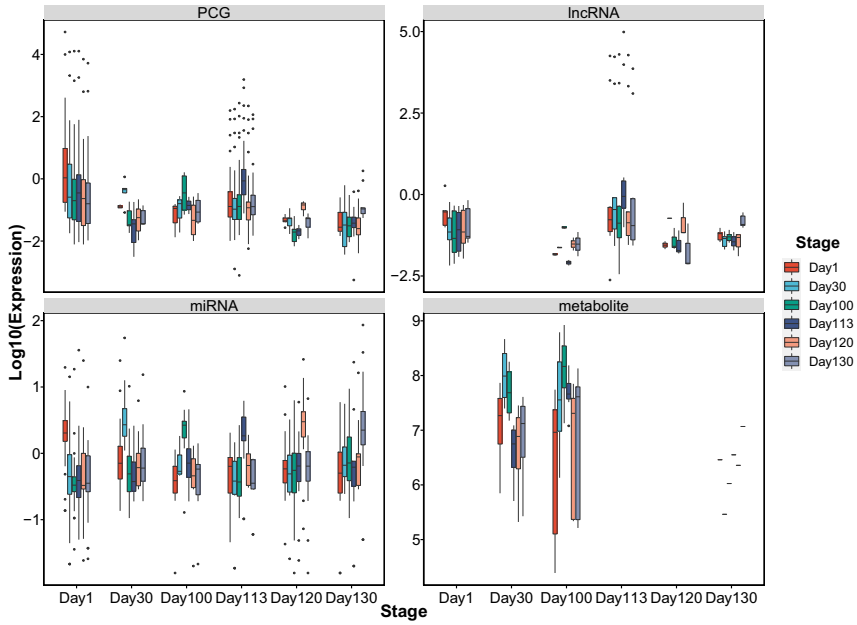


Figure 5. Expressions of stage-specific PCGs, lncRNAs, and miRNAs, and accumulation of stage-specific metabolites

The box shows the 25th–75th percentile (that is, the interquartile range; IQR); the line shows the median; the whiskers show 1.5 × IQR. These PCGs, lncRNAs, miRNAs, and metabolites with more than four times higher expression in a given stage over any other stages were considered stage-specific molecules. The Phi coefficient ≥0.6 indicated that the molecule was expressed in (nearly) all samples from the stage (it was exclusive of the stage), and was also considered a stage-specific molecule.

Prediction of TFs

For the DE and stage-specific PCGs, and the PCGs in significant modules of WGCNA, we predicted their bound TFs. In total, 3,762 PCGs were enriched with 33 significant motifs (*E-value* < 0.05, Figure S8), which were annotated 297 TF. Among them, 40, four, and two TFs were DE PCGs, stage-specific PCGs, and molecules in the significant modules of WGCNA, respectively (Table S13).

There were four TFs (EGR1, EGR3, KLE1, and MSGN1) involved in the development of sow pregnancy that showed stage-specific expression. For example, EGR1 and EGR3 were the stage-specific genes on day 1 and respectively participated in the locomotor rhythm and neuromuscular synaptic transmission (Table S9). MSGN1 as a basic helix-loop-helix transcription factor involved in the protein dimerization activity, showed stage-specific expression at day 120 (Table S9), and activated the genes in the Notch pathway, which acted on mammary stem cells to promote self-renewal.[18,19] For the PCGs in the significant modules of WGCNA, HIC1 and EGR2 were identified as the TFs in the significant Darkgreen module of day 1 and day 30 (Figure 6C, Table S10), and involved in transcription regulatory and motor neuron axon guidance.

Integrated analysis of metabolome and transcriptome in development stages

To analyze the regulatory network of the SDMs, a correlation analysis of DE PCGs/lncRNAs/miRNAs and SDMs was performed in each group. Correlation analysis showed that a total of 2,936 DE PCGs, 517 lncRNAs, and 147 miRNAs were significantly correlated with 300, 284, and 258 SDMs, respectively (correlation value ≥0.7 or ≤ −0.7 and p < 0.05, Tables S14 and S15). One stage-specific miRNA was strongly correlated with four metabolites at day 30, and one PCG, two lncRNAs, and two miRNAs as the specific molecules at day 100 were significantly correlated with seven metabolites (Tables S14 and S15). In the significant modules of WGCNA, 99 transcriptome molecules were significantly correlated with 45 metabolites. Numerous metabolites were positively or negatively regulated by multiple RNAs and a single transcriptome molecule could regulate multiple metabolites (Tables S14 and S15). For example, 11,14,17-Eicosatrienoic acid(FFA(20:3n3) was significantly correlated with 13 DE PCGs in comparison group day 1 vs day 30 (two negatively correlated and 11 positively correlated, Figure 8A). PCG ENSSSCG00000036983 of

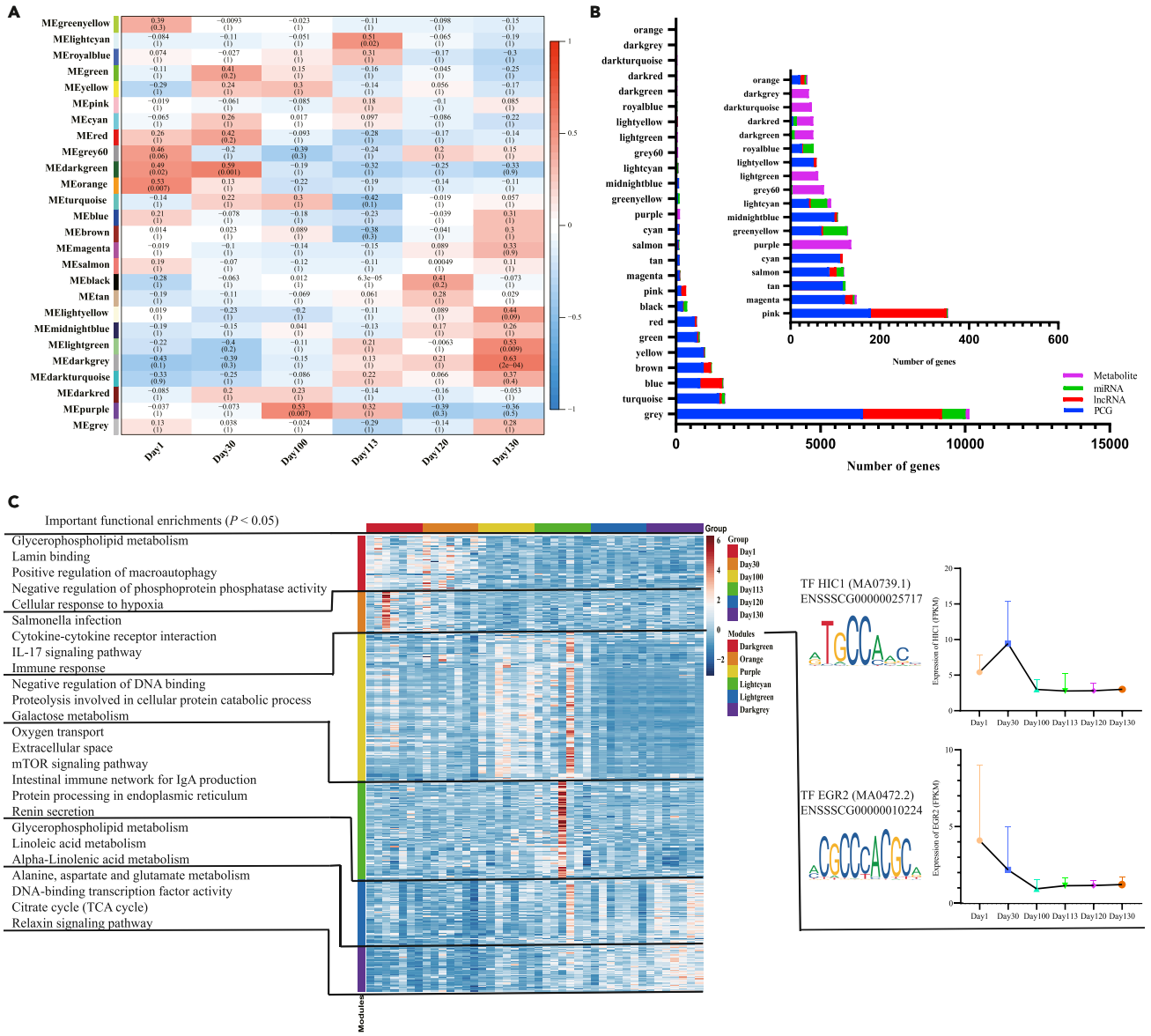


Figure 6. The weighted gene co-expression network analysis (WGCNA)

(A) Associations of modules with stages. The Bonferroni method was used for the multiple corrected statistical significance. Each cell contains the correlation and corrected p-value.

(B) Bar chart indicated numbers of transcriptome and metabolome molecules detected in individual modules.

(C) Heatmap showed the expressions of RNAs and metabolite accumulation in the six significant modules. The important functional enrichments, expressions of TFs, and motif of representative sequence were shown nearly to the corresponding module.

significant module-related day 1 was annotated to be positively correlated with ChE(20:4)*ChE(20:4) and PC(38:4)*PC(18:0_20:4) (Figure 8B). Stage-specific miRNA 1_737 of day 100 was positively correlated with seven metabolites (Chenodeoxycholic acid, hyocholic acid, chenodeoxycholic acid glycine conjugate, glycohyocholic acid, pregnenolone sulfate, pregnanolone sulfate, and X-MZ397RT475, Figure 8C). According to these results, a complex regulatory mechanism played between the variation in metabolite accumulation and transcriptome molecule expression during the stages of pig from estrus to lactation.

Combined analysis with the important pregnancy-related metabolites of human

We searched the reported studies of metabolites related to human pregnancy and discovered 50 metabolites commonly identified in humans and pigs. These 50 metabolites (Table S16) identified in humans

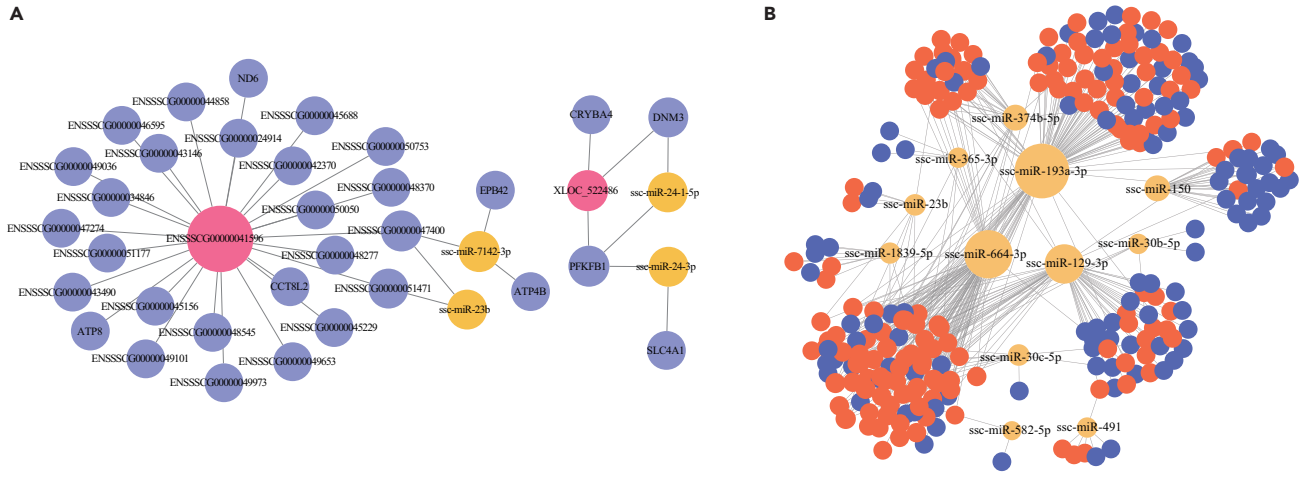


Figure 7. Regulatory networks of PCGs, lncRNAs, and miRNAs

The red, blue, and yellow indicated lncRNAs, PCGs, and miRNAs, respectively.

(A) The regulatory networks of stage-specific molecules for two lncRNAs and four miRNAs with their target genes.

(B) The lncRNA-miRNA-mRNA interactions including 12 miRNAs identified in the Lightcyan module associated with day 113 and Darkgrey module strongly associated with day 130.

mainly were from four studies,[4,20–22] in which, eight metabolites were important markers for human pregnancy. These eight metabolites were also the markers for pig pregnancy and were significantly detected by the DE analysis, stage-specific analysis, and WGCNA. Furthermore, four (4-Vinylphenol sulfate, choline, pregnanolone sulfate, and pregnenolone sulfate) of the eight metabolites were involved in benzoate metabolism, glycine, serine and threonine metabolism, progesterone metabolism, and steroid hormone biosynthesis, respectively, and they had strong correlations with 214 PCGs, 34 lncRNAs, and 42 miRNAs. We showed the regulatory networks of these four key metabolic indicators in Figures 9 and S9–S11.

DISCUSSION

In the present study, we built a sow RNA-Seq transcriptomic and metabolomic BloodMap including six stages from estrus to lactation. Thousands of molecules were commonly expressed across all six stages, despite some PCGs, lncRNAs, miRNAs, and metabolites displaying a stage-specific differential expression. Interestingly, day 113 had more specific molecules than other stages, which might provide for the body's needs for upcoming labor and lactation. The stage-specific molecules are well correlated with the biological functions of each stage. For example, day 100-specific lncRNAs were activated in embryonic development. These indicated that the sow might have various regulatory systems throughout the breeding development.

We observed an obvious clustering distinction of metabolites between different stages using HCA and PCA analyses, which indicated the similarity between biological replicates and differences between stages. For the transcriptome molecules, the clustering distinctions seemed weak, while some of stages could be separated from each other by the PCA results. Variance partition analysis confirmed that the individual effects were dominant factor affecting the expression of transcriptome molecules. To eliminate individual factors, we carefully conducted the differential expression analysis by considering the individual effect using the model "design = ~ Stage + Individual". By DE analysis, we found that the comparison groups day 30 vs day 130 and day 100 vs day 130 had thousands of DE PCGs, lncRNAs and miRNAs, and SDMs, which suggested a significant difference between pregnancy and lactation. The same expression pattern of molecules suggested their similar function. Four stage-specific TFs, EGR1, EGR3, KLF1, and MSGN1, were found to be associated with embryo implantation, fetal blood supply, and embryo development.[23–26] Two TFs, EGR2, and HIC1, in the significant modules associated with day 1 and day 30, were suggested to be involved in marking mammary morphogenesis and healthy.[27,28] Hence, these RNAs might be the key molecules regulating the development of sow pregnancy.

The identification of SDMs showed that the metabolite content of each stage was strongly different, while the stage-specific metabolite was relatively few, which was due to correlations among metabolites. By the WGCNA, we obtained 289 metabolites significantly associated with the development stages of the sow from estrus to

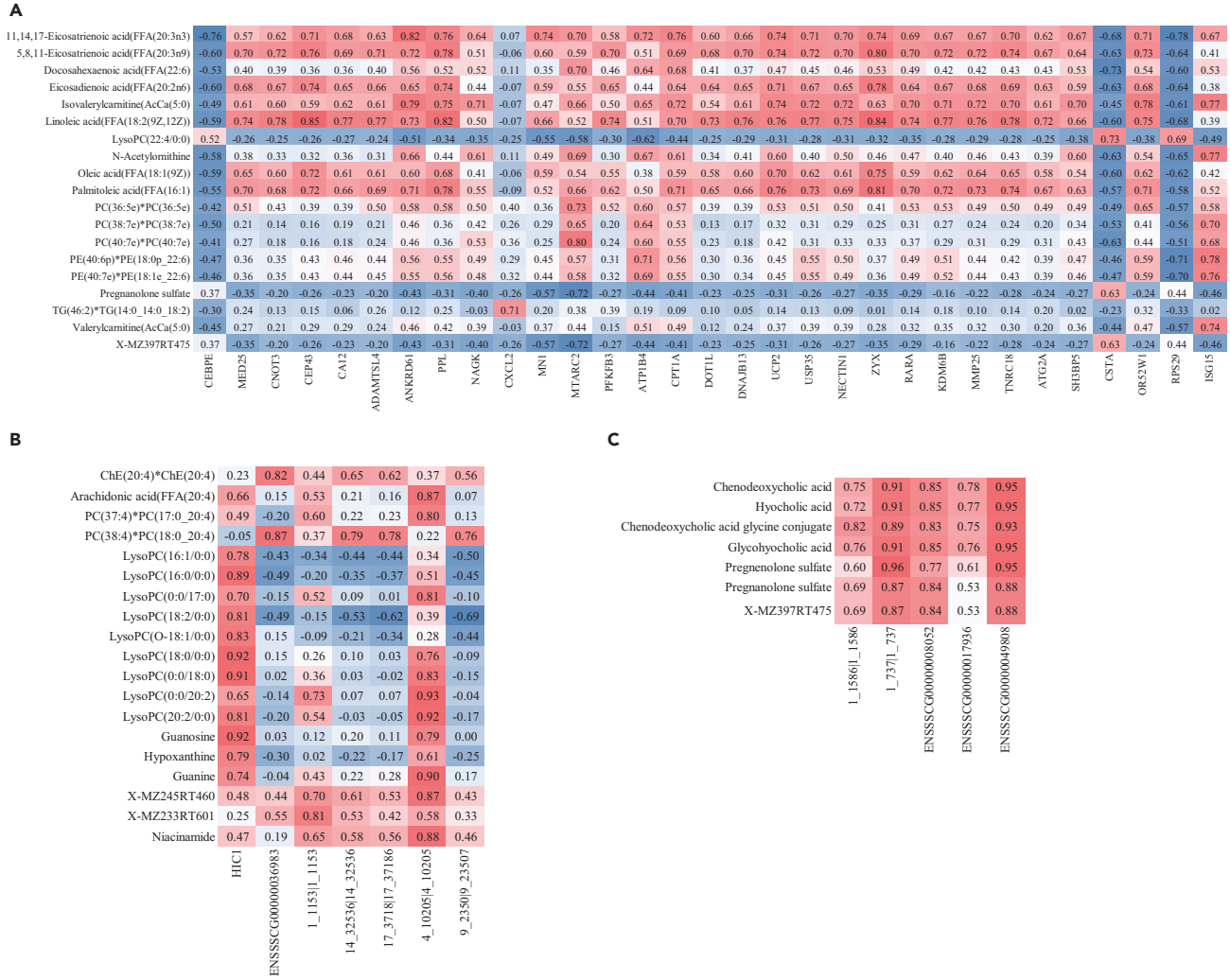


Figure 8. Heatplot of the associations between transcriptome and metabolome

PCGs, lncRNAs, and miRNAs in rows and metabolites in columns observed the interrelation between RNAs and metabolites. The positive and negative correlations were indicated by red and blue, respectively.

(A) Heatplot of the correlations between DE RNAs and SDMs on day 1 vs day 30.

(B) Heatplot of the correlations between stage-specific RNAs and metabolites of day 100.

(C) Heatplot of the correlations between RNAs and metabolites of significant modules associated with day 1.

lactation, and these metabolites might be the marker for the corresponding stage. In addition, these metabolites might be the important markers of human pregnancy to research-related metabolic diseases. For instance, the accumulation of pregnenolone sulfate increased in the stage of pregnancy in humans,[4] and was significantly high on day 100 of pigs' pregnancy, which indicated that the pregnenolone sulfate had a similar function in humans and pigs. Furthermore, we conducted the integrated analysis of metabolome and transcriptome to show that many transcriptome molecules and metabolites were involved in the same functional pathways of a specific stage. The expression of transcriptome RNAs exhibited a strong association with the metabolite, which indicated the metabolites had the opportunity to serve as the gene expression markers.

According to the DE and time-series analyses, stage-specific molecules identification, WGCNA, and integrated analysis of metabolome and transcriptome, many promising candidate molecules involved in the reproductive processes were detected. For instance, nine PCGs (*SLC45A1*, *MAL*, *EGR3*, *EGR1*, *HBB*, *SNCG*, *NR4A1*, *AHSP*, and *DNM3*), one lncRNA (ENSSSCG00000045096), and one metabolite (glycohyocholic acid) were jointly confirmed by DE and time-series analyses, and stage-specific molecules identification. *EGR3* and *EGR1* were also observed in the prediction of TFs. *SLC45A1* was associated with

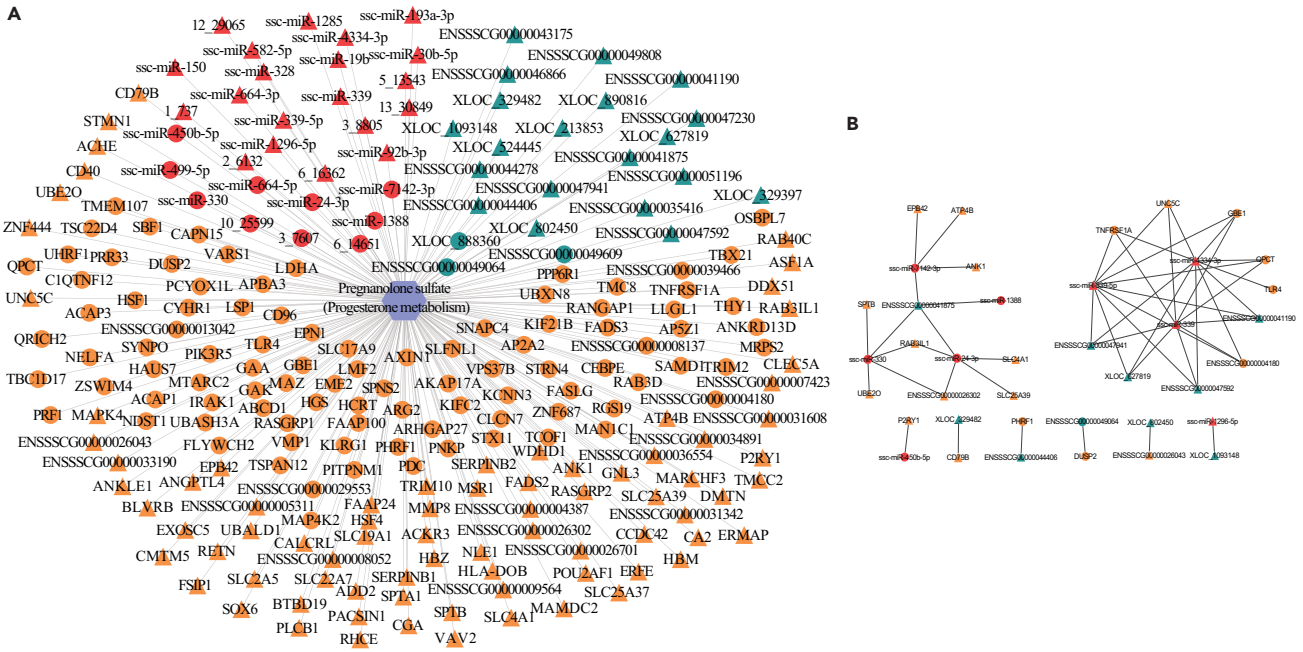


Figure 9. Regulatory networks of metabolite pregnanolone sulfate commonly identified in humans and pigs

(A) The transcriptome and metabolite correlation. The tan, blue-green, and red indicated PCGs, lncRNAs, and miRNAs, respectively. Pregnanolone sulfate was mainly involved in progesterone metabolism. The circles and triangles indicated negative and positive correlations, respectively.

(B) The ceRNA regulatory network of PCGs, lncRNAs, and miRNAs involved in the interrelated network.

development of infancy,[29] and *MAL* was highly expressed in early endosome membranes.[30] *HBB* was associated with a variety of physiological properties and its p.K65R SNP had strong impacts on reproductive performance of Awassi ewes.[31] *SNCG* was overexpressed in various types of human tumors of reproductive organs, such as breast, ovary, and cervix.[32–34] *NR4A1*, an immediate-early gene encoding an orphan nuclear receptor, had a key role in the ovulatory process, and was a potential molecular marker in selection for increasing litter size in pigs.[35] *AHSP* had protection against oxidative stressors, which were related to miscarriage.[36] *DNM3* was the member of a superfamily of large GTPase proteins dynamin, and dynamin polymers contributed to the membrane fission and scission of nascent vesicles from parent membranes.[37] The function research of lncRNA ENSSSCG00000045096 was limited, while its target PCG *CDS1* was identified to own significantly higher expressions on day 113, day 120, and day 130 than that on day 30. The previous study proposed that *CDS1* was an essential gene for cell growth,[38] and was linked to DNA replication.[39] Glycohyocholic acid was reported to be a potential marker of a pregnancy-specific liver disorder.[40] These implied that the key molecules might be involved in embryo implantation and development process of pigs, and be the important genetic markers.

The datasets were accessible by the NCBI database and supplemental information of this study, which consists of 42 samples from multiple development stages across the sow pregnancy. It may be used to detect novel genes, lncRNAs, and miRNAs, and to better annotate the pig transcriptome. The profile of a novel gene, lncRNA or miRNA expression, and metabolite accumulation throughout 42 samples of this study can be used as the fingerprint to infer its normal biological function in the development of sows' pregnancy. In addition, stage-specific expression analysis, expression patterns of key molecules, and integrated analysis of metabolome and transcriptome could be used to examine the pharmacological and toxicological effects of medications that may differ depending on the stage of human pregnancy. Moreover, the BloodMap reported here may be utilized as a foundation for cross-species comparison, improving the way that findings on preclinical animal safety are translated into health effects in humans.

Limitations of the study

The present study provided a relatively large dataset and suggested the important biomarkers for mammal pregnancy with their preliminary regulatory mechanism. Our research was the early bioinformatics analysis;

herein, the validation of these markers was lacking. Another limitation was the small sample size, resulting in a non-negligible individual effect, while we used the model "design = ~ Stage + Individual" to decrease the influence of individual effects on the difference analysis.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animal individuals and samples collection
- METHOD DETAILS
  - RNA and metabolites extraction
  - RNA library construction and sequencing
  - Metabolomic analysis and evaluation of metabolite contents
  - Identification of protein-coding genes (PCGs), lncRNAs, and miRNAs
  - Analysis of transcriptomic and metabolomic molecule expression profiles
  - Blood-enriched stage-specific PCGs, lncRNAs, miRNAs, and metabolites
  - Analysis of development-dependent molecules expression and content patterns
  - Transcription factor (TF) binding motif analysis
  - The prediction of target genes of DE lncRNAs and miRNAs
  - Construction of ceRNA regulatory networks and WGCNA
  - Integrated analysis of transcriptome and metabolome
  - Pathway analysis
  - Construction of pathways regulated the human and pig common metabolic marker
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105825.

ACKNOWLEDGMENTS

We would like to thank pig breeders at the Institute of Animal Science, Chinese Academy of Agricultural Sciences for assistance in the collection of blood samples. This work was financially supported by the China Agriculture Research System of MOF and MARA, Agricultural Science and Technology Innovation Program (CAAS-ZDRW202006-04), and Agricultural Science and Technology Innovation Project (ASTIP-IAS02).

AUTHOR CONTRIBUTIONS

This research was conceived and designed by Lx.W. and L.S. The DNA was collected by L.S. with the help of H.Y., Z.S., J.L., and Lg.W. The data analysis was performed by L.S. with the help of H.L. and X.H. This manuscript was prepared by L.S. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 6, 2022
Revised: October 10, 2022
Accepted: December 12, 2022
Published: January 20, 2023

REFERENCES

1. Bassols, A., Costa, C., Eckersall, P.D., Osada, J., Sabrià, J., and Tibau, J. (2014). The pig as an animal model for human pathologies: a proteomics perspective. Proteomics. Clin. Appl. *8*, 715–731. https://doi.org/10.1002/prca.201300099.

2. Zhang, Z., Chen, Z., Ye, S., He, Y., Huang, S., Yuan, X., Chen, Z., Zhang, H., and Li, J. (2019). Genome-wide association study for reproductive traits in a duroc pig population. Animals *9*, 732. https://doi.org/10.3390/ani9100732.

3. Bazer, F.W., Song, G., Kim, J., Dunlap, K.A., Satterfield, M.C., Johnson, G.A., Burghardt, R.C., and Wu, G. (2012). Uterine biology in pigs and sheep. J. Anim. Sci. Biotechnol. *3*, 23. https://doi.org/10.1186/2049-1891-3-23.

4. Liang, L., Rasmussen, M.L.H., Piening, B., Shen, X., Chen, S., Röst, H., Snyder, J.K., Tibshirani, R., Skotte, L., Lee, N.C., et al. (2020). Metabolic dynamics and prediction of gestational age and time to delivery in pregnant women. Cell *181*, 1680–1692.e15. https://doi.org/10.1016/j.cell.2020.05.002.

5. Sinowatz, F., and Friess, A.E. (1983). Uterine glands of the pig during pregnancy. An ultrastructural and cytochemical study. Anat. Embryol. *166*, 121–134. https://doi.org/10.1007/BF00317948.

6. Ziecik, A.J., Waclawik, A., Kaczmarek, M.M., Blitek, A., Jalali, B.M., and Andronowska, A. (2011). Mechanisms for the establishment of pregnancy in the pig. Reprod. Domest. Anim. *46*, 31–41. https://doi.org/10.1111/j.1439-0531.2011.01843.x.

7. Zang, X., Gu, T., Hu, Q., Xu, Z., Xie, Y., Zhou, C., Zheng, E., Huang, S., Xu, Z., Meng, F., et al. (2021). Global transcriptomic analyses reveal genes involved in conceptus development during the implantation stages in pigs. Front. Genet. *12*, 584995. https://doi.org/10.3389/fgene.2021.584995.

8. Palombo, V., Loor, J.J., D'Andrea, M., Vailati-Riboni, M., Shahzad, K., Krogh, U., and Theil, P.K. (2018). Transcriptional profiling of swine mammary gland during the transition from colostrogenesis to lactogenesis using RNA sequencing. BMC Genom. *19*, 322. https://doi.org/10.1186/s12864-018-4719-5.

9. Armit, C., Venkataraman, S., Richardson, L., Stevenson, P., Moss, J., Graham, L., Ross, A., Yang, Y., Burton, N., Rao, J., et al. (2012). eMouseAtlas, EMAGE, and the spatial dimension of the transcriptome. Mamm. Genome *23*, 514–524. https://doi.org/10.1007/s00335-012-9407-1.

10. Henry, A.M., and Hohmann, J.G. (2012). High-resolution gene expression atlases for adult and developing mouse brain and spinal cord. Mamm. Genome *23*, 539–549. https://doi.org/10.1007/s00335-012-9406-2.

11. Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature *489*, 101–108. https://doi.org/10.1038/nature11233.

12. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74. https://doi.org/10.1038/nature11247.

13. Li, Q., Liang, R., Li, Y., Gao, Y., Li, Q., Sun, D., and Li, J. (2020). Identification of candidate genes for milk production traits by RNA sequencing on bovine liver at different lactation stages. BMC Genet. *21*, 72. https://doi.org/10.1186/s12863-020-00882-y.

14. Guo, H., Guo, H., Zhang, L., Tang, Z., Yu, X., Wu, J., and Zeng, F. (2019). Metabolome and transcriptome association analysis reveals dynamic regulation of purine metabolism and flavonoid synthesis in transdifferentiation during somatic embryogenesis in cotton. Int. J. Mol. Sci. *20*, 2070. https://doi.org/10.3390/ijms20092070.

15. Banerjee, P., Carmelo, V.A.O., and Kadarmideen, H.N. (2020). Integrative analysis of metabolomic and transcriptomic profiles uncovers biological pathways of feed efficiency in pigs. Metabolites *10*, 275. https://doi.org/10.3390/metabo10070275.

16. Yang, Y., Yuan, H., Liu, X., Wang, Z., Li, Y., Ren, Y., Gao, C., Jiao, T., Cai, Y., and Zhao, S. (2022). Transcriptome and metabolome integration provides new insights into the regulatory networks of Tibetan pig alveolar type II epithelial cells in response to hypoxia. Front. Genet. *13*, 812411. https://doi.org/10.3389/fgene.2022.812411.

17. Wan, J.J., Lin, C.H., Ren, E.D., Su, Y., and Zhu, W.Y. (2019). Effects of early intervention with maternal fecal bacteria and antibiotics on liver metabolome and transcription in neonatal pigs. Front. Physiol. *10*, 171. https://doi.org/10.3389/fphys.2019.00171.

18. Chalamalasetty, R.B., Dunty, W.C., Jr., Biris, K.K., Ajima, R., Iacovino, M., Beisaw, A., Feigenbaum, L., Chapman, D.L., Yoon, J.K., Kyba, M., and Yamaguchi, T.P. (2011). The Wnt3a/beta-catenin target gene Mesogenin1 controls the segmentation clock by activating a Notch signalling program. Nat. Commun. *2*, 390. https://doi.org/10.1038/ncomms1381.

19. Dontu, G., Jackson, K.W., McNicholas, E., Kawamura, M.J., Abdallah, W.M., and Wicha, M.S. (2004). Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. Breast Cancer Res. *6*, R605–R615. https://doi.org/10.1186/bcr920.

20. Stelzer, I.A., Ghaemi, M.S., Han, X., Ando, K., Hédou, J.J., Feyaerts, D., Peterson, L.S., Rumer, K.K., Tsai, E.S., Ganio, E.A., et al. (2021). Integrated trajectories of the maternal metabolome, proteome, and immunome predict labor onset. Sci. Transl. Med. *13*, eabd9898. https://doi.org/10.1126/scitranslmed.abd9898.

21. Orczyk-Pawilowicz, M., Jawien, E., Deja, S., Hirnle, L., Zabek, A., and Mlynarz, P. (2016). Metabolomics of human amniotic fluid and maternal plasma during normal pregnancy. PLoS One *11*, e0152740. https://doi.org/10.1371/journal.pone.0152740.

22. Handelman, S.K., Romero, R., Tarca, A.L., Pacora, P., Ingram, B., Maymon, E., Chaiworapongsa, T., Hassan, S.S., and Erez, O. (2019). The plasma metabolome of women in early pregnancy differs from that of non-pregnant women. PLoS One *14*, e0224682. https://doi.org/10.1371/journal.pone.0224682.

23. Liang, X.H., Deng, W.B., Li, M., Zhao, Z.A., Wang, T.S., Feng, X.H., Cao, Y.J., Duan, E.K., and Yang, Z.M. (2014). Egr1 protein acts downstream of estrogen-leukemia inhibitory factor (LIF)-STAT3 pathway and plays a role during implantation through targeting Wnt4. J. Biol. Chem. *289*, 23534–23545. https://doi.org/10.1074/jbc.M114.588897.

24. Shin, H., Kwon, S., Song, H., and Lim, H.J. (2014). The transcription factor Egr3 is a putative component of the microtubule organizing center in mouse oocytes. PLoS One *9*, e94708. https://doi.org/10.1371/journal.pone.0094708.

25. Ravindranath, Y., Johnson, R.M., Goyette, G., Buck, S., Gadgeel, M., and Gallagher, P.G. (2018). KLF1 E325K-associated congenital dyserythropoietic anemia type IV: insights into the variable clinical severity. J. Pediatr. Hematol. Oncol. *40*, E405–E409. https://doi.org/10.1097/Mph.0000000000001056.

26. Liu, D., Zhang, Q., Zhang, H., Tang, L., Li, W., Zhang, D., Wu, G., Ye, S., Ban, Q., and He, K. (2016). A comprehensive transcriptomic analysis of differentiating embryonic stem cells in response to the overexpression of Mesogenin 1. Aging *8*, 2324–2336. https://doi.org/10.18632/aging.101049.

27. Gutierrez, G., Sun, P., Han, Y., and Dai, X. (2022). Defining mammary basal cell transcriptional states using single-cell RNA-sequencing. Sci. Rep. *12*, 4893. https://doi.org/10.1038/s41598-022-08870-1.

28. Wang, Y., Weng, X., Wang, L., Hao, M., Li, Y., Hou, L., Liang, Y., Wu, T., Yao, M., Lin, G., et al. (2018). HIC1 deletion promotes breast cancer progression by activating tumor cell/fibroblast crosstalk. J. Clin. Invest. *128*, 5235–5250. https://doi.org/10.1172/JCI99974.

29. Srour, M., Shimokawa, N., Hamdan, F.F., Nassif, C., Poulin, C., Al Gazali, L., Rosenfeld, J.A., Koibuchi, N., Rouleau, G.A., Al Shamsi, A., and Michaud, J.L. (2017). Dysfunction of the cerebral glucose transporter SLC45A1 in individuals with intellectual disability and epilepsy. Am. J. Hum. Genet. *100*, 824–830. https://doi.org/10.1016/j.ajhg.2017.03.009.

30. Puertollano, R., and Alonso, M.A. (1999). MAL, an integral element of the apical sorting machinery, is an itinerant protein that cycles between the trans-Golgi network and the plasma membrane. Mol. Biol. Cell *10*, 3435–3447. https://doi.org/10.1091/mbc.10.10.3435.

31. Al-Nafie, A.T., Al-Thuwaini, T.M., and Al-Shuhaib, M.B.S. (2022). A novel association between hemoglobin subunit beta gene and reproductive performance in Awassi ewes. J. Saudi Soc. Agric. Sci. *21*, 1–7. https://doi.org/10.1016/j.jssas.2021.06.018.

32. Bruening, W., Giasson, B.I., Klein-Szanto, A.J., Lee, V.M., Trojanowski, J.Q., and Godwin, A.K. (2000). Synucleins are expressed in the majority of breast and ovarian carcinomas and in preneoplastic lesions of the ovary. Cancer *88*, 2154–2163.

33. Guo, J., Shou, C., Meng, L., Jiang, B., Dong, B., Yao, L., Xie, Y., Zhang, J., Chen, Y., Budman, D.R., and Shi, Y.E. (2007). Neuronal protein synuclein gamma predicts poor clinical outcome in breast cancer. Int. J. Cancer *121*, 1296–1305. https://doi.org/10.1002/ijc.22763.

34. Liu, H., Liu, W., Wu, Y., Zhou, Y., Xue, R., Luo, C., Wang, L., Zhao, W., Jiang, J.D., and Liu, J. (2005). Loss of epigenetic control of synuclein-gamma gene as a molecular indicator of metastasis in a wide range of human cancers. Cancer Res. *65*, 7635–7643. https://doi.org/10.1158/0008-5472.Can-05-1089.

35. Liu, L.Q., Li, F.E., Deng, C.Y., and Xiong, Y.Z. (2011). Molecular cloning, tissue expression and association of porcine NR4A1 gene with reproductive traits. Mol. Biol. Rep. *38*, 103–114.

36. Emanuelli, M., Cecati, M., Sartini, D., Stortoni, P., Corradetti, A., Giannubilo, S.R., Turi, A., and Tranquilli, A.L. (2009). Placental alpha hemoglobin stabilizing protein (AHSP) and recurrent miscarriage. Cell Stress Chaperones *14*, 193–197. https://doi.org/10.1007/s12192-008-0072-y.

37. Romeu, A., and Arola, L. (2014). Classical dynamin DNM1 and DNM3 genes attain maximum expression in the normal human central nervous system. BMC Res. Notes *7*, 188. https://doi.org/10.1186/1756-0500-7-188.

38. Shen, H., Heacock, P.N., Clancey, C.J., and Dowhan, W. (1996). The CDS1 gene encoding CDP-diacylglycerol synthase in Saccharomyces cerevisiae is essential for cell growth (*). J. Biol. Chem. *271*, 789–795. https://doi.org/10.1074/jbc.271.2.789.

39. Lindsay, H.D., Griffiths, D.J., Edwards, R.J., Christensen, P.U., Murray, J.M., Osman, F., Walworth, N., and Carr, A.M. (1998). S-phase-specific activation of Cds1 kinase defines a subpathway of the checkpoint response in Schizosaccharomyces pombe. Genes Dev. *12*, 382–395. https://doi.org/10.1101/gad.12.3.382.

40. Cui, Y., Xu, B., Zhang, X., He, Y., Shao, Y., and Ding, M. (2018). Diagnostic and therapeutic profiles of serum bile acids in women with intrahepatic cholestasis of pregnancy-a pseudo-targeted metabolomics study. Clin. Chim. Acta *483*, 135–141. https://doi.org/10.1016/j.cca.2018.04.035.

41. Parkhomchuk, D., Borodina, T., Amstislavskiy, V., Banaru, M., Hallen, L., Krobitsch, S., Lehrach, H., and Soldatov, A. (2009). Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res. *37*, e123. https://doi.org/10.1093/nar/gkp596.

42. Cai, W., Li, C., Liu, S., Zhou, C., Yin, H., Song, J., Zhang, Q., and Zhang, S. (2018). Genome wide identification of novel long non-coding RNAs and their potential associations with milk proteins in Chinese Holstein cows. Front. Genet. *9*, 281. https://doi.org/10.3389/fgene.2018.00281.

43. Cai, W., Li, C., Li, J., Song, J., and Zhang, S. (2021). Integrated small RNA sequencing, transcriptome and GWAS data reveal microRNA regulation in response to milk protein traits in Chinese Holstein cattle. Front. Genet. *12*, 726706.

44. Ernst, J., and Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. BMC Bioinf. *7*, 191. https://doi.org/10.1186/1471-2105-7-191.

45. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

46. John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. PLoS Biol. *2*, e363–e1879. https://doi.org/10.1371/journal.pbio.0020363.

47. Razani, B., Zarnegar, B., Ytterberg, A.J., Shiba, T., Dempsey, P.W., Ware, C.F., Loo, J.A., and Cheng, G. (2010). Negative feedback in noncanonical NF-kappa B signaling modulates NIK stability through IKK alpha-mediated phosphorylation. Sci. Signal. *3*, ra41. https://doi.org/10.1126/scisignal.2000778.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Health pig blood tissue	Changping Experimental Base of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (Beijing, China)	https://ias.caas.cn/
Deposited data		
Raw and analyzed data	This paper	PRJNA857151
Raw data of metabolite	This paper	Table S2
Experimental models: Organisms/strains		
Pig: Yorkshire pig	Changping Experimental Base of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (Beijing, China)	https://ias.caas.cn/
Software and algorithms		
UPLC-HRMS instrument	Thermo Scientific	N/A
AcquityTM HSS C18 column	Waters Co.	N/A
AcquityTM BEH amide column	Waters Co.	N/A
Accucore C30 core-shell column	Thermo Scientific	N/A
Ensembl database	https://asia.ensembl.org/Sus_scrofa/Info/Index?db=core	N/A
XCalibur software	Thermo Scientific	N/A
STAR	http://star.mit.edu/	http://star.mit.edu/
Cytoscape software	https://cytoscape.org/	https://cytoscape.org/
BEDTools	https://bedtools.readthedocs.io/en/latest/	https://bedtools.readthedocs.io/en/latest/
StringTie	https://bioinformaticshome.com/tools/rna-seq/descriptions/StringTie.html#gsc.tab=0	https://bioinformaticshome.com/tools/rna-seq/descriptions/StringTie.html#gsc.tab=0
MetaboAnalyst5.0	https://www.metaboanalyst.ca/faces/home.xhtml	https://www.metaboanalyst.ca/faces/home.xhtml
KOBAS	http://kobas.cbi.pku.edu.cn/kobas3/genelist/	http://kobas.cbi.pku.edu.cn/kobas3/genelist/
MEME Suite	http://meme-suite.org/index.html	http://meme-suite.org/index.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Lixian Wang (iaswlx@263.net).

Materials availability

The present study did not generate new unique reagents.

Data and code availability

The RNA-seq data used here had been submitted to the NCBI Sequence Read Archive (SRA) under accession PRJNA857151 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA857151). The Raw mass spectral data of metabolism was in Table S2.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this manuscript is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal individuals and samples collection

All the experimental animals were 550 days old Yorkshire pigs from Changping Experimental Base of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (Beijing, China). Seven Yorkshire pigs were from the same flock management and would prepare for the third pregnancy. The blood samples were collected from estrus to lactation, namely, Day 1 (estrus and mating period; on day1 of pregnancy), Day30 (implantation period; on day30 of pregnancy), Day100 (embryonic period; on day100 of pregnancy), Day113 (last few days before farrowing, at day113 of pregnancy), Day120 (early stage of milk production, at day6 after parturition), and Day 130 (mid-stage of milk production, at day 16 after parturition). A total of seven biological replicates were carried out for each stage.

All 42 blood samples of seven pigs were collected. We divided each blood sample into two parts: one was disposed of by red cell lysis buffer (RT122, Tiangen, Beijing, China) to obtain the white cell, and the remaining blood was used for isolating serum by centrifugation for 15 minutes (3,000 rpm). The white cell and serum samples were snap-frozen in liquid nitrogen.

METHOD DETAILS

RNA and metabolites extraction

42 blood white cell samples were used for extracting the total RNA with the TRIzol reagent (Life Technologies, CA, USA). The 1% agarose gels were utilized to track the RNA degradation and contamination. A NanoPhotometer® spectrophotometer (IMPLEN, CA, USA) was to check the RNA purity. The RNA concentration and integrity were respectively measured through the Qubit® RNA Assay Kit in Qubit® 2.0 Fluorometer (Life Technologies) and RNA Nano 6000 Assay Kit.

Each serum sample was split into two portions, and the polar and non-polar lipid metabolites were extracted for metabolomic profiling analysis. Polar metabolites extraction: precisely absorbing 100 μL serum to the EP tube, adding 400 μL methanol-acetyol (1/1, v/v), vortexing 120 s, centrifugaling 15 min with 15,000 g, transferring 200 μL of the liquid to two 1.5 mL EP tubes, respectively. Lipid metabolites extraction: adding 120 μL methanol (including lipid extract for internal standard Ceramide (D18: 1/26: 0-O-18: 1 (D9)) to 20 μL of serum, vortexing 180 s, adding 360 μL MTBE and 100 μL ultra-pure water in turn, after mixing the vortex, 10 min lay layers in the refrigerator at 4° C, 15,000 g × 15 min centrifuged layers, precise transfer of 300 μL of upper lipid extract objects accurately in the EP tube. Finally, the polar and lipid metabolites were dried in the centrifugal concentration (Thermo Scientific, USA) with a low-temperature vacuum reduction.

RNA library construction and sequencing

For mRNA and lncRNA, the strand-specific library was created by removing ribosomal RNA to.[41] Through the EpicentreRibo-Zero™ rRNA Removal Kit (Epicentre, WI, USA), we respectively eliminated the rRNA from total RNA, and got the rRNA-free residue. Subsequently, the RNA was broken into short fragments of 250–300 bp, which were used as the template to synthesize the cDNA. The PCR amplification and products were separately conducted and purified (AMPure XP system) to construct a library. Finally, according to the effective concentration of the library and data output requirements, we sequenced the libraries by a Novaseq 6000 platform (Illumina, NBE, USA), and obtained 150 bp paired-end reads.

With a Small RNA Sample Pre Kit (Illumina, NBE, USA), we constructed a library of small RNA. Briefly, both ends of small RNA were directly added adapters and then were reversely transcribed to synthesize cDNA, in which, the total RNA samples were as the sating materials. The target DNA fragments were separated by PAGE gel electrophoresis following PCR amplification, and the cDNA library was created by reusing the sliced gel. Based on the effective concentration of the library and data output requirements, we sequenced the libraries through a Novaseq 6000 platform (Illumina, NBE, USA), and achieved 50 bp single-end reads.

Metabolomic analysis and evaluation of metabolite contents

Based on the manufacturer's instructions, we analyzed the metabolites by the UPLC-HRMS instrument (Thermo Scientific, San Jose, CA, USA). The reported research also described the detailed information,[14] in which, polar metabolites used AcquityTM HSS C18 column (Waters Co., USA, 2.1 × 100 mm) and AcquityTM BEH amide column (Waters Co., USA, 1.7 μm, 2.1 × 100 mm), and lipid metabolites utilized Accucore C30 core-shell column (Thermo Scientific, USA). The high-resolution mass spectrometry detection

conditions were performed according to three resolutions, namely 70,000, 17,500, and 120,000 full widths half maximum resolution through the XCalibur software (Thermo Scientific, USA) to obtain the MS/MS spectra data.

With the Compound Discoverer software, we carried out the full scan and handled the profile data to extract the comprehensive component as the instructions of manufacturer. Further, through the area under curve values extracted by with XCalibur Quan Browser, the polar metabolite quantitative information was achieved and exported. For the lipidomics data, we used the LipidSearch software to process peak picking, and identify lipids.

Identification of protein-coding genes (PCGs), lncRNAs, and miRNAs

Based on the clean reads, all follow-up analyses were conducted. For the mRNAs and lncRNAs, we aligned the clean reads to the pig reference genome (Scrofa11.1) by the STAR software. StringTie was used for counting the fragment of genes, and fragments per kilobase of exon model per million mapped fragments (FPKM) indicated the expression of PCGs and lncRNAs. StringTie and Scripture (beta2) were used to detect the novel lncRNA transcripts, and the detailed descriptions were in the previous study,[42] while the initial identification of lncRNAs was not filtered by the expression levels. For small RNAs, by Cutadapt 2.8, we trimmed the quality and removed the adaptor of the reads, and identified the miRNAs according to the description method in the report research.[43] The reference genome was Scrofa11.1, and the novel miRNAs were detected by combining unannotated sequences with the known miRNA annotation from *Ovis aries*, *Bos taurus*, and *Gallus gallus*.

Analysis of transcriptomic and metabolomic molecule expression profiles

In this study, a transcriptomic molecule was regarded as expressed if its expression value in FPKM for PCGs and lncRNAs, and counts per million (CPM) for miRNAs were equal to or greater than 0.1 in five (10%) samples. If the expression in FPKM or CPM of a transcriptomic molecule was more than 0.1 in two samples for each stage, the molecule was considered expressed in the stage. Further, the transcriptomic molecule was regarded as a commonly expressed RNA with its expressions occurring in all developmental stages.

We performed the HCA for the PCGs, lncRNAs, miRNAs, and metabolites according to a distance matrix of the Spearman correlation of these samples, and carried out a two-dimensional scatter plot to visualize the global relationships among samples with the PCA analyses. We conducted the initial sequence and exonic structural features analysis and examined the stage specificity of gene abundance shown by the tau score (τ) (0–1), in which, 0 and 1 indicated ubiquitously transcribed and highly tissue-specific genes, respectively. We analyzed the stage specificity of transcripts observed through the stage specificity index and conducted the cumulative distribution of the average percentage of all transcription or metabolites abundance supplied through genes or metabolites when arranged from most to least expressed in all groups.

Blood-enriched stage-specific PCGs, lncRNAs, miRNAs, and metabolites

DE PCGs, lncRNAs, and miRNAs between any two stages were identified by the DEseq2 package in R with the corrected p-*values* less than 0.05, in which, we carefully considered the individual effect using the model "design = ~ Stage + Individual". We conducted the t-test for the metabolites, of which, a p-*value* less than 0.05 and fold change (FC) $\geq$ 2 or $\leq$0.5 denoted statistical significance. The orthogonal PLS-DA was utilized to identify the differences between these samples and to eliminate the irrelevant differences. The variables were regarded to be different if the variable importance in projection (VIP) values were greater than 1. Finally, the SDMs were screened by p-*values* less than 0.05, FC $\geq$ 2 or $\leq$0.5, and VIP greater than 1. The stage-specific molecules were identified using different multiple thresholds of 4. A PCG, lncRNA, miRNA, or metabolite was considered to be a stage-specific molecule if its expression level or content level was more than four times higher in a given stage over any other stage.

We calculated the Phi coefficient to find the exclusive PCGs, lncRNAs, and miRNAs in each stage. The condition selected samples that were either expressed (FPKM/CPM $\geq$0.1) or were not expressed (FPKM/CPM <0.1). We used a threshold of 0.1 to conservatively select stage-specific PCGS, lncRNAs, and miRNAs that had substantial expression. The second condition selected samples coming either from the tested stage or from all other stages. We calculated the Phi coefficient using the function Phi coefficient of the R package psych. The Phi coefficient measured the association between two binary variables. We defined those PCGs, lncRNAs, and miRNAs with Phi coefficient $\geq$0.6 as the stage-exclusive genes.

Analysis of development-dependent molecules expression and content patterns

According to the DE molecules, we analyzed their expression patterns between any two adjacent stages with the younger stage as the denominator. A DE upregulated molecule between two adjacent stages was grouped into the "up" pattern. A DE downregulated molecule between two adjacent stages was grouped into 'decrease', and a molecule with no difference in expression between two adjacent stages was grouped into "maintain". Thus, a DE molecule was classified as one out of 125 patterns, ranging from up-up-up-up-up (UUUUU), maintain-maintain-maintain-maintain-maintain (MMMMM), to decrease-decrease-decrease-decrease-decrease (DDDDD).

In addition, we conducted the time-series analysis for all the PCGs, lncRNAs, miRNAs, and metabolites using STEM (Short Time-series expression Miner),[44] in which, the "maximum number of model profiles" and the "maximum unit change in model profiles between time points" were set at 50 and 2, respectively.

Transcription factor (TF) binding motif analysis

The enrichment of TFs bound in the promoters of DE PCGs was detected by the MEME Suite (http://meme-suite.org/index.html). We extracted 2,000 bp sequences upstream of DE PCGs by the BioMart program of Ensembl database (https://asia.ensembl.org/Sus_scrofa/Info/Index?db=core) to predict the motifs enriched 8–15 widths through the STREME tool (*E-value* < 0.05). The JASPAR CORE vertebrates database was used for detecting and annotating the motifs with the Tomtom tool. We screened the DE PCGs matched to TFs with the Find Individual Motif Occurrences (FIMO) tool, in which, the threshold was p-*value* < 1E−04.

The prediction of target genes of DE lncRNAs and miRNAs

By the BEDTools,[45] the *cis*-targets (PCGs within 100 kb up- and downstream regions of lncRNAs) of DE lncRNAs were identified. The Pearson correlation between the DE lncRNAs and the identified PCGs was calculated to detect the *trans*-targets with Pearson coefficient of correlations >0.9 and p-values < 0.05. In addition, by the miRanda[46] with energy < −10 and TargetScan[46] with context score percentile >50, the target genes of miRNAs were identified simultaneously. Then the Pearson correlation between the DE miRNAs and their target genes was calculated to confirm the final target genes with a negative correlation and p-*values* < 0.05.

Construction of ceRNA regulatory networks and WGCNA

As the method of target genes prediction of DE miRNAs, we predicted the target lncRNAs for the DE miRNAs, in which, pairs with the negative Pearson coefficient were selected. The ceRNA was created by all the DE miRNAs and their target PCGs and lncRNA. The ceRNA network was visualized by Cytoscape software.[47]

A gene co-expression network was created by the R package WGCNA, where 13,295 genes, 4,209 lncRNAs, and 1,242 miRNAs were identified in all six stages. The significant modules with Bonferroni p-values <0.05 were defined as the functional modules for the developmental stages of pigs from oestrus to lactation.

Integrated analysis of transcriptome and metabolome

The DE, stage-specific, and significant module molecules were used for the integrative analysis of different development stages. The integrations of metabolome and transcriptome were conducted through the Spearman method. Only the detected correlations with a Pearson's correlation coefficient value ≥ 0.7 or ≤ -0.7 and $p < 0.05$ were selected. Heat plots were used to show the connections between RNAs and metabolites.

Pathway analysis

In this study, we performed the functional annotation of RNAs and metabolites by the KOBAS (http://kobas.cbi.pku.edu.cn/kobas3/genelist/) and MetaboAnalyst5.0 (https://www.metaboanalyst.ca/faces/home.xhtml), respectively. The enrichments with p-*values* less than 0.05 were regarded as significant.

Construction of pathways regulated the human and pig common metabolic marker

We compared the results of the pig with that of humans previously reported to identify the common metabolic marker. Further, combining transcriptome results and functional analysis, we constructed the pathways that regulated the common metabolic marker.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using R. Details for specific analyses were described in the figure legends, results, and method details.