1 **Identifying long non-coding RNAs and characterizing their functional roles in**

2 **swine mammary gland from colostrogenesis to lactogenesis**

3 Lijun Shi, Longchao Zhang, Ligang Wang, Xin Liu, Hongmei Gao, Xinhua Hou, Hua

4 Yan, Fuping Zhao, Wentao Cai[*] and Lixian Wang[*]

5 Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing

6 100193, China

7 [*] Corresponding author

8 Cai Wen-tao, Institute of Animal Science, Chinese Academy of Agricultural Sciences，

9 No. 2 Yuanmingyuan West Road, Haidian District, Beijing 100193, China. Email:

10 caiwentao@caas.cn.

11 Wang Li-xian, Institute of Animal Science, Chinese Academy of Agricultural Sciences，

12 No. 2 Yuanmingyuan West Road, Haidian District, Beijing 100193, China. E-mail:

13 iaswlx@263.net.

14

15

16   **Abstract**

17   Long non-coding RNAs (lncRNAs) play important regulatory roles in milk biological

18   processes. While, little is known about the identification and function of swine

19   lncRNAs in response to mammary gland development. Here, we identified 286

20   differentially expressed (DE) lncRNAs in mammary gland at different stages from 14

21   days prior to (-) parturition to day 1 after (+) parturition using the published RNA-seq

22   data. The expression pattern of these DE lncRNAs was examined, and most of lncRNAs

23   expressions were strongly changed from Day -2 to Day +1. Functional annotation

24   revealed that the DE lncRNAs were mainly involved in the mammary gland developing,

25   lactation developing, milk composition metabolism and colostrum function. By

26   performing the weighted gene co-expression network analysis (WGCNA), we

27   identified 7 out of 12 lncRNA-mRNA modules, including 35 lncRNAs and 319

28   mRNAs, which were highly associated with the mammary gland at Day -14, Day -2

29   and Day +1. Integrated analysis of the DE lncRNAs expression pattern examination,

30   targets prediction, function annotation and WGCNA suggested that 18 lncRNAs (such

31   as, *XLOC_020627*, *ENSSSCG00000051193*, *XLOC_025150*, *ENSSSCG00000042618*,

32   *XLOC_963181*, *ENSSSCG00000051701*, *XLOC_018030*, and *XLOC_025146*) and

33   their 20 target genes (such as, *ACTN4*, *ADCY1*, *CSN3*, *SMO*, *CSN2*, *PRKAG2*, *FIBCD1*,

34   and *GALNT7*), were considered as the promising candidates for swine parturition and

35   colostrum occurrence processes. Our research provided the insights into lncRNA

36   profiles and their regulating mechanisms from colostrogenesis to lactogenesis in swine.

37   **Keywords:** Swine, Mammary gland, Colostrum, LncRNA, WGCNA

38

Introduction

Although the litter size of sow has made a significant improvement during last two decades, the problem of growth performance and mortality in piglets become increasingly prominent. During gestation and lactation, the growth performance and health status of piglets have been critical factors impacting reproductive performance of modern sows (Kim 1999). Sow milk performance as a major limiting factor contributes to suboptimal growth and survival of piglets (Boyd & Kensinger 1998). The quality and quantity of milk in sows are highly variable, for example, the colostrum yield is reported to range from <1kg to 8.5kg (Quesnel et al. 2012; Vadmand et al. 2015). To enhance the sow lactation ability, it is necessary to understand the mammary gland development, and screen their regulatory factors (Hurley 2019). In 2016, Balzani *et al.* reported the heritability of udder morphology in crossbreed sows (Large White $\times$ Meishan), namely, teat length with $h^2 = 0.46$, teat diameter with $h^2 = 0.56$, interteat distance within the same row with $h^2 = 0.37$, abdominal midline with $h^2 = 0.22$, udder development score $h^2 = 0.25$, proportion of nonfunctional teats with $h^2 = 0.3$, and the proportion of teats oriented perpendicular to the udder with $h^2 = 0.1$ (Balzani et al. 2016). Generally, from colostrogenesis to lactogenesis, mammary glands might undergo significant functional differentiation for swine parturition and colostrum occurrence processes. In 2018, Palombo *et al.* used the RNA sequencing (RNA-seq) to detect the candidate genes impacting swine parturition and lactation in mammary gland of crossbred sows (Danish Landrace $\times$ Yorkshire), such as *CSN1S2*, *LALBA*, *WAP*, *SAA2*, and *BTN1A1* (Palombo et al. 2018).

4

61   Long non-coding RNA (LncRNA) is a newfound class of non-protein coding

62   transcripts in eukaryotes with a minimum length of 200 nt (Jin et al. 2020), and it plays

63   important regulatory role in biological processes (Liang et al. 2018). Many lncRNAs

64   have been reported along with the depth and quality of RNA-seq (St Laurent et al. 2015).

65   In swine, lots of potential regulatory lncRNAs have been identified from various tissues,

66   such as intramuscular adipose (Huang et al. 2018; Miao et al. 2018), longissimus dorsi

67   muscle (Wang et al. 2019a), preadipocytes (Li et al. 2019), and porcine endometrium

68   (Wang et al. 2017). In 2018, Liang *et al.* built a systematic S.scrofa lncRNA database

69   named lncRNAnet that contained 53,468 S.scrofa lncRNAs with their sequence

70   characteristics, genomic locations, conservation, overlapping single nucleotide

71   polymorphism and quantitative trait loci, and transcript abundance across nine tissues

72   (fat, heart, kidney, liver, lung, muscle, ovarium, spleen, and testis) (Liang *et al.* 2018).

73   However, no study has been performed to detect the lncRNAs regulating milk traits

74   from sow mammary gland.

75   In the present study, we identified lncRNAs in sow mammary gland at different

76   stages from 14 days prior to parturition to day 1 after parturition according to the

77   published RNA-seq data, and detected the differentially expressed (DE) lncRNAs.

78   Further, functional annotation and WGCNA were conducted to predict the regulatory

79   functions of DE lncRNAs. To the best of our knowledge, this is the first study to identify

80   lncRNAs from mammary gland in swine, and our results paved the way for a better

81   understanding of lncRNA functions in swine parturition and colostrum occurrence

82   processes.

5

Material and methods

RNA-seq dataset

The RNA-seq data sources (GSE101983) were involved in 15 mammary gland tissues, which were collected from 3 crossbred sows (Danish Landrace $\times$ Yorkshire) at Days 14, 10, 6 and 2 before (-) parturition and Day 1 after (+) parturition (Table S1)(Palombo *et al.* 2018).

Identification of lncRNAs

According to the raw data, clean reads were obtained by removing containing adapter molecules, reads including poly-N, and low-quality reads using Trimmomatic with defaults parameters (Bolger et al. 2014). All follow-up analyses were performed with the high-quality reads. Through STAR software, we aligned the clean reads to the pig reference genome (Sscrofa11.1).

We used StringTie and Scripture to assemble the novel transcripts. All the transcripts with the same start, end position and exon–intron boundary, which were supported by at least two assembly programs or occurred in at least two samples were obtained (Guttman et al. 2010; Pertea et al. 2015). Further, the putative novel lncRNAs were detected via a customized multi-step pipeline: (1) The transcripts were removed, which were likely to be assembly artifacts or PCR run-on fragments based on class code annotated by gffcompare. For the different class, we retained those only annotated by 'u' and 'i', which indicated novel intergenic and intronic transcripts, respectively.

6

103   Transcripts with class code "=" were considered as known genes. (2) Transcripts with

104   length $\geqslant$ 200 nt and exon $\geqslant$ 2 were retained to avoid incomplete assemble and too

105   many splicing events. (3) Transcripts with low expression were removed using FPKM

106     $\geqslant$ 0.3 as cut off. The expression levels of transcripts were calculated in fragments per

107   kilo-base of exon per 10^6 mapped fragments (FPKM) by StringTie -e -B. (4) Maximum

108   ORF (open reading frame) lengths of less than 120 amino acids (360 nt) were obtained.

109   (5) Transcripts with predicted protein-coding potential were removed, that the protein-

110   coding potential criteria were CPC score > 0, PLEK score > 0, and CNCI score > 0.

111   Finally, the candidate sets of lncRNAs were made up by those without coding potentials.

112   **Differential expression analysis of lncRNAs**

113   To detect the DE lncRNAs and coding genes, the expression levels of these

114   lncRNAs and coding genes identified in this study from the mammary gland at Days -

115   10, -6, -2 and +1 were compared to their respective expression levels at Day -14 using

116   the EdegR R package. RNA-Seq read counts were modeled by a generalized linear

117   model that considered the experimental design with two factors (individuals and stages

118   of lactation). Therefore, the formula in EdegR showed as follow: Design= ~Individuals

119   + Stages. LncRNAs or genes with adjusted P-value < 0.05 were assigned as the

120   differences. The expression patterns of DE lncRNAs across five stages were performed

121   by k-mean method (Ahmad & Dey 2007). Through gap statistics, we determined that k

122   = 14 was the optimal choice for distinguishing these lncRNAs.

123   **Target gene prediction of lncRNAs and functional analysis**

124   LncRNAs can cis-regulate neighboring target genes and trans-regulate distant

125   target genes (Derrien et al. 2012; Guil & Esteller 2012). To predict cis-regulated target

126   genes of lncRNAs, the coding genes located within 100 kb upstream/downstream of

127   lncRNAs were checked. Further, we computed the expressed correlation coefficients

128   between lncRNAs and their neighboring genes by Spearman method, in which, the

129   significant lncRNA-mRNA correlation pairs were assigned when the P-value < 0.05.

130   To investigate the functions of DE lncRNAs, we performed gene ontology (GO)

131   and kyoto encyclopedia of genes and genomes (KEGG) enrichments using KOBAS

132   (http://kobas.cbi.pku.edu.cn/kobas3/genelist/) (Xie et al. 2011). The significant GO

133   terms and pathways were assigned with the P-value < 0.05.

134   **Module construction base on WGCNA**

135   WGCNA is a systematic in silico method for analysis of complex gene regulatory

136   networks, which can be used for finding modules of significantly correlated genes, and

137   for relating modules to one another and to external sample traits (Langfelder & Horvath

138   2008; Le et al. 2019). Here, we conducted WGCNA to detect the significant modules

139   by WGCNA R package. In the significant modules, we screened the lncRNAs and

140   genes, which might be as candidates to impact the swine lactation developing, milk fat

141   and protein metabolism and colostrum function.

142   **Results**

143   **Identification of lncRNAs**

8

144   In this study, RNA-seq data of 15 samples from the mammary gland of three

145   crossbred sows with five time points were used (Table S1) (Palombo *et al.* 2018). After

146   quality control, reads mapping and lncRNA identification, a total of 1,346 lncRNA

147   transcripts located in 1,084 lncRNA loci were detected (Table S2), in which, 483 (613

148   lncRNA transcripts) were novel lncRNAs, and 601 (733 lncRNA transcripts) were

149   known lncRNAs.

150   **Differential expression analysis**

151   A total of 286 lncRNAs (Table S3) were differentially expressed (adjusted P-value

152   < 0.05), including 9 (7 up-regulated and 2 down-regulated), 71 (41 up-regulated and 30

153   down-regulated), 169 (101 up-regulated and 68 down-regulated) and 206 (128 up-

154   regulated and 78 down-regulated) for -10vs-14, -6vs-14, -2vs-14 and 1vs-14

155   comparison groups, respectively (Figure 1A-D). Notably, six lncRNAs were

156   differentially expressed in all four comparison groups (Figure 2). Further, we examined

157   the expression pattern of DE lncRNAs across five lactation stage using k-means

158   clustering, and these lncRNAs could be divided into 14 distinct clusters. In cluster 2, 3,

159   7, 12, 13 and 14, most of DE lncRNAs expressions were rapidly decreased at Day -2,

160   and strongly increased at Day +1, which indicated these lncRNAs were dynamic

161   changed from colostrogenesis to lactogenesis (Figuer 3).

162   **Prediction of target genes**

9

163   To investigate the probable roles of the lncRNAs in swine mammary gland, we

164   predicted the target genes of lncRNAs. First, a search for protein-coding genes within

165   100 kb upstream/downstream of lncRNAs, and 6,589 genes were identified. Further,

166   the expressions correlation between lncRNAs and protein-coding genes were examined,

167   and 804 significant lncRNA-mRNA correlation pairs ($P < 0.05$) were detected,

168   including 685 protein coding genes and 443 lncRNAs. Among the 443 lncRNAs, 149

169   were differentially expressed, including 5, 41, 94 and 109 DE lncRNAs respectively

170   corresponded to -10vs-14, -6vs-14, -2vs-14 and +1vs-14 (Table S4).

171   In addition, we found that 256 protein-coding genes within 100 kb

172   upstream/downstream of the 149 DE lncRNAs were also confirmed to be strongly

173   correlated with them. To assess the function of DE lncRNAs, GO and KEGG

174   enrichment analyses were performed using the 256 target genes with KOBAS. There

175   were 311 significant enrichments were presented ($P < 0.05$; Table S5), including 298

176   GO terms and 13 KEGG pathways. These GO and KEGG enrichments were involved

177   in 87 genes, and most of them were directly related to the labor and colostrum. For

178   example, *CSN3*, *SMO*, and *CSN2* were involved in mammary associated activities.

179   **Module construction**

180   To explore the specific lncRNAs or genes that were highly associated with

181   lactation, we performed WGCNA using DE RNAs and their target genes. A total of 12

182   modules associated with the specific expression profiles of different samples were

183   identified (Figure 4A). Then, we calculated associations of each module with five

184   lactation stages, and found seven modules including Greenyellow, Green, Black,

185   Yellow, Brown, Blue, and Turquoise, were significantly associated with Day -14, Day

186   -2 and Day +1 ($P < 0.05$; Figure 4B). In detail, the Greenyellow module, including 12

187   lncRNAs and 6 target genes were highly associated with Day -2 ($P = 0.02$). The Green

188   (including 5 lncRNAs and 30 target genes), Black (including 1 lncRNAs and 23 target

189   genes), Yellow (including 5 lncRNAs and 49 target genes), and Brown (including 1

190   lncRNAs and 53 target genes) modules were significantly associated with the

191   mammary gland samples at Day +1 ($P =3e-04 \sim 0.009$). The Turquoise (including 9

192   lncRNAs and 91 target genes) and Blue (including 2 lncRNAs and 67 target genes)

193   modules were strongly associated with at Day -14 and Day +1 ($P = 0.004 \sim 0.04$). The

194   lncRNAs and target genes involved in the significant modules were shown in Table S6.

195   **Comprehensive analysis**

196   Among the 256 target genes, 121 were differentially expressed across five

197   lactation stages, and function annotation revealed that 25 of them targeted by 23

198   lncRNAs were involved in lactation developing, milk fat metabolism and colostrum

199   function (Table 1). Interestingly, 20 target genes (*ACTN4*, *ADCY1*, *CSN3*, *SMO*, *PTK7*,

200   *MPDZ*, *NPR1*, *CSN2*, *ATP2C2*, *PRKAG2*, *CD36*, *NPR1*, *ACSL3*, *GALNT15*,

201   *C6H1orf210*, *CDC20*, *SNRPD1*, *FIBCD1*, *ACTB* and *GALNT7*) and 3 lncRNAs

202   (*XLOC_025150*, *ENSSSCG00000011196*, and *XLOC_010589*) were also belong to

203   significant modules by WGCNA. These 20 target genes were correspond to 18

204   lncRNAs, which also included the 3 lncRNAs (*XLOC_025150*,

11

205   *ENSSSCG00000011196*, and *XLOC_010589*) involved in significant modules. Hence,

206   we suggested that these 18 lncRNAs targeting 20 genes were the candidates involved

207   in lactation of sow, which were considered as the potentially functional candidates

208   impacting the mammary gland developments from colostrogenesis to lactogenesis. In

209   addition, we conducted the network of promising candidate lncRNAs, genes and

210   pathways (Figure 5), in which, the functions of lncRNAs were clearly shown. For

211   example, XLOC_025150 might regulate *CSN3* gene, were involved in the mammary

212   gland development and lactation.

Discussion

214   In this study, we systematically analyzed the RNA-seq data of 15 swine mammary

215   gland samples collected at days 14, 10, 6 and 2 before (-) parturition to day 1 after (+)

216   parturition, and identified 286 DE lncRNAs targeting 256 genes. Further, we found that

217   these lncRNAs were involved in delivery and lactation developing, milk lipid

218   metabolism, and immune function of colostrum, and 18 lncRNAs targeting 20 genes

219   were proposed to be the candidates involved in lactation of sow.

220   At the parturition, the preparation for copious milk synthesis and secretion have

221   begun (Kensinger et al. 1982; Palombo *et al.* 2018), and the mammary gland can reach

222   the greatest degree of structural development. In this study, we examined the expression

223   pattern of DE lncRNAs, most of the lncRNAs expressions were strongly changed from

224   Day -2 to Day +1. Further, 18 promising functional lncRNAs targeting 20 genes were

225   mainly identified in -2vs-14 and 1vs-14 groups, which reflected a strong activation of

12

226   many metabolic processes before and after parturition. In the WGCNA, 20 promising

227   functional DE target genes were involved in the significant modules, which were highly

228   associated with the mammary gland samples at Day -14, Day -2 and Day +1. Hence,

229   our results were consistent with the concept: before and after farrowing with the

230   formation of colostrum, mammary glands undergo strong functional differentiation

231   immediately (Hurley 2019).

232   In general, lncRNAs exert regulatory functions at different levels of gene

233   expression to influence the tumor growth, cell-cycle, and apoptosis (Le *et al.* 2019;

234   Wang et al. 2019b). For example, lncRNA H19 is involved in the regulation of high

235   glucose-induced apoptosis through targeting VDAC1 (Li et al. 2016). Here, 11

236   lncRNA-mRNA target pairs, *XLOC_020627-ACTN4*, *ENSSSCG00000051193-ADCY1*,

237   *XLOC_025150-CSN3*,   *ENSSSCG00000042618-SMO*,   *XLOC_963181-PTK7*,

238   *ENSSSCG00000051701-MPDZ*,   *XLOC_018030-NPR1*,   *XLOC_025146-CSN2*,

239   *ENSSSCG00000041015-ATP2C2*,   *ENSSSCG00000046607-PRKAG2*,   and

240   *ENSSSCG00000049698-ACTB,* were involved in delivery and lactation metabolism,

241   such as tight junctions, oxytocin, development of the mammary gland and lactation.

242   Tight junctions of mammary gland from the pregnant animal are leaky, undergoing

243   closure around parturition to become the impermeable tight junctions of the lactating

244   animal (Nguyen & Neville 1998). In dairy cattle, after parturition, the start of copious

245   milk production requires the closure of tight junctions to form the blood-milk barrier

246   and prevent paracellular transfer of blood constituents into milk (such as lactate

247   dehydrogenase and serum albumin) and vice versa (such as appearance of α-

13

248   lactalbumin in blood) (Kessler et al. 2019). Oxytocin is a nonapeptide hormone that has

249   a central role in the regulation of parturition and lactation (Arrowsmith & Wray 2014),

250   and its best-known and most well-established roles are stimulation of uterine

251   contractions during parturition and milk release during lactation (Gimpl & Fahrenholz

252   2001). Development of the mammary gland occurs in defined stages, which are

253   connected to sexual development and reproduction (embryonic, prepubertal, pubertal,

254   pregnancy, lactation, and involution) (Hennighausen & Robinson 2001). Lactation is a

255   highly demanding lipid synthesis and transport process that is crucial for the

256   development of newborn mammals (Wan et al. 2007). Embryo development ending in

257   birth or egg hatching term was defined as the process whose specific outcome is the

258   progression of an embryo over time, from zygote formation until the end of the

259   embryonic life stage, and for mammals it is usually considered to be birth. The above

260   reports suggested that these GO and KEGG enrichments were related with the

261   parturition and lactation. Hence, we proposed that these lncRNAs in 11 lncRNA-

262   mRNA target pairs might be regulated their target genes to impact the delivery and

263   lactation processes.

264   Additionally, we found eight lncRNA-mRNA target pairs, *XLOC_020627-ACTN4*,

265   *ENSSSCG00000051193-ADCY1*,   *ENSSSCG00000042618-SMO*,   *XLOC_963181-*

266   *PTK7*, *XLOC_026156-CD36*, *XLOC_018030-NPR1*, *XLOC_010589-ACSL3*, and

267   *ENSSSCG00000046607-PRKAG2*, were enriched in PPARs, regulation of lipolysis in

268   adipocytes, cellular response to lipid, and adipocytokine signaling pathways, that might

269   be related with the milk components metabolism.

14

270   For the piglets, colostrum plays an important role in ensuring their survival,

271   growth and health by providing energy, nutrients, immunoglobulins, growth factors and

272   many other bioactive components and cells (Quesnel & Farmer 2019). In the present

273   study, we found that 12 lncRNA-mRNA target pairs, *XLOC_020627-ACTN4*,

274   *ENSSSCG00000042618-SMO*, *XLOC_963181-PTK7*, *ENSSSCG00000011196-*

275   *GALNT15*, *ENSSSCG00000041987-C6H1orf210*, *ENSSSCG00000041987-CDC20*,

276   *ENSSSCG00000041015-ATP2C2*, *ENSSSCG00000046607-GALNTL5*,

277   *ENSSSCG00000042000-SNRPD1*, *XLOC_002435-FIBCD1*, *ENSSSCG00000049698-*

278   *ACTB* and *ENSSSCG00000048264-GALNT7*, were involved in leukocyte

279   transendothelial migration, response to retinoic acid, drug binding, mucin type O-

280   glycan biosynthesis, regulation of defense response to virus, and positive regulation of

281   ubiquitin-protein transferase activity enrichments, which might impact the immune

282   function of colostrum.

283   Based on these advantages, we proposed these 18 lncRNAs (*XLOC_020627*,

284   *ENSSSCG00000051193*, *XLOC_025150*, *ENSSSCG00000042618*, *XLOC_963181*,

285   *ENSSSCG00000051701*, *XLOC_018030*, *XLOC_025146*, *ENSSSCG00000041015*,

286   *ENSSSCG00000046607*, *ENSSSCG00000049698*, *XLOC_026156*, *XLOC_010589*,

287   *ENSSSCG00000011196*, *ENSSSCG00000041987*, *ENSSSCG00000042000*,

288   *XLOC_002435*, and *ENSSSCG00000048264*) as the promising candidates for swine

289   parturition and lactation. Our network results of 18 lncRNAs, their 20 target genes and

290   corresponding pathways further clearly showed the functions of the candidate lcnRNAs

291   and their target genes. There is a high probability that these lncRNAs and genes interact

15

292   closely and thus act as key regulators in mammary gland, thus, future work should focus

293   on the potential roles of them in sow lactation.

294   **Conclusions**

295   Sow milk production is the major factor which can limit growth and survival of

296   piglets. To enhance milk performance, it is necessary to understand the process of

297   mammary gland morphogenesis and to identify the regulatory factors of mammary

298   development. In this study, we identified 1,084 lncRNAs in swine mammary gland

299   using RNA-seq data, and 286 were DE lncRNAs in five lactation stages. Integrated

300   analysis of the DE lncRNAs expression pattern examination, targets prediction,

301   function annotation and WGCNA, we proposed that 18 lncRNAs (such as

302   *XLOC_020627*, *ENSSSCG00000051193*, *XLOC_025150*, *ENSSSCG00000042618*,

303   *XLOC_963181*, *ENSSSCG00000051701*, *XLOC_018030* and *XLOC_025146*) targeting

304   20 genes (such as *ACTN4*, *ADCY1*, *CSN3*, *SMO*, *CSN2*, *PRKAG2*, *FIBCD1* and

305   *GALNT7*) as the promising candidates were involved in swine parturition and colostrum

306   occurrence processes. These results provided a new insight for exploring critically

307   regulatory factors involved in reproductive performance of sow.

308   **Supplementary information**

309   **Table S1** The information of RNA sequencing data for 15 samples of sow mammary

310   gland tissue

311   **Table S2** The results of the identified lncRNAs in this study

312 **Table S3** The detailed information of DE lncRNAs in this study

313 **Table S4** The predicted target genes of DE lncRNAs

314 **Table S5** Significant GO and KEGG pathways of the target genes of DE lncRNAs

315 **Table S6** The lncRNAs and genes involved in significant modules

316 **Acknowledgements**

320 **Data availability and materials**

321 The data (BioProject ID: GSE101983) was in the Gene Expression Omnibus database.

322 **Conflicts of interest**

323 The authors declare that no competing interests exist.

324

17

References

Ahmad A. & Dey L. (2007) A k-mean clustering algorithm for mixed numeric and categorical data. Data & Knowledge Engineering 63, 503-27.

Arrowsmith S. & Wray S. (2014) Oxytocin: Its Mechanism of Action and Receptor Signalling in the Myometrium. Journal of Neuroendocrinology 26, 356-69.

Balzani A., Cordell H.J., Sutcliffe E. & Edwards S.A. (2016) Heritability of udder morphology and colostrum quality traits in swine. J Anim Sci 94, 3636-44.

Bolger A.M., Lohse M. & Usadel B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-20.

Boyd R.D. & Kensinger R.S. (1998) *Metabolic precursors for milk synthesis*.

Derrien T., Johnson R., Bussotti G., Tanzer A., Djebali S., Tilgner H., Guernec G., Martin D., Merkel A., Knowles D.G., Lagarde J., Veeravalli L., Ruan X., Ruan Y., Lassmann T., Carninci P., Brown J.B., Lipovich L., Gonzalez J.M., Thomas M., Davis C.A., Shiekhattar R., Gingeras T.R., Hubbard T.J., Notredame C., Harrow J. & Guigo R. (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 22, 1775-89.

Gimpl G. & Fahrenholz F. (2001) The Oxytocin Receptor System: Structure, function, and regulation. Physiological Reviews 81, 629-83.

Guil S. & Esteller M. (2012) Cis-acting noncoding RNAs: friends and foes. Nat Struct Mol Biol 19, 1068-75.

Guttman M., Garber M., Levin J.Z., Donaghey J., Robinson J., Adiconis X., Fan L., Koziol M.J., Gnirke A., Nusbaum C., Rinn J.L., Lander E.S. & Regev A. (2010) Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs (vol 28, pg 503, 2010). Nature Biotechnology 28, 756-.

Hennighausen L. & Robinson G.W. (2001) Signaling pathways in mammary gland development. Dev Cell 1, 467-75.

Huang W., Zhang X., Li A., Xie L. & Miao X. (2018) Genome-Wide Analysis of mRNAs and lncRNAs of Intramuscular Fat Related to Lipid Metabolism in Two Pig Breeds. Cell Physiol Biochem 50, 2406-22.

Hurley W.L. (2019) Review: Mammary gland development in swine: embryo to early lactation. Animal 13, S11-S9.

Jin Y., Zhang K., Huang W., Tang W., Li H., Dong W., Gu J. & Zhou J. (2020) Identification of functional lncRNAs in pseudorabies virus type II infected cells. Vet Microbiol 242, 108564.

Kensinger R.S., Collier R.J., Bazer F.W., Ducsay C.A. & Becker H.N. (1982) Nucleic acid, metabolic and histological changes in gilt mammary tissue during pregnancy and lactogenesis. J Anim Sci 54, 1297-308.

Kessler E.C., Wall S.K., Hernandez L.L., Gross J.J. & Bruckmaier R.M. (2019) Short communication: Mammary gland tight junction permeability after parturition is greater in dairy cows with elevated circulating serotonin concentrations. Journal

367   of Dairy Science 102, 1768-74.

368 Kim S. (1999) Mammary Gland Growth and Nutrient Mobilization in Lactating Sows:
369   A Dynamic Model to Describe Nutrient Flow.

370 Langfelder P. & Horvath S. (2008) WGCNA: an R package for weighted correlation
371   network analysis. BMC Bioinformatics 9, 559.

372 Le K., Guo H., Zhang Q., Huang X., Xu M., Huang Z. & Yi P. (2019) Gene and lncRNA
373   co-expression network analysis reveals novel ceRNA network for triple-
374   negative breast cancer. Sci Rep 9, 15122.

375 Li S., Chen C.Z., Chai M.L., Wang J.W., Yuan B., Gao Y., Jiang H. & Zhang J.B. (2019)
376   Identification and Analysis of lncRNAs by Whole-Transcriptome Sequencing
377   in Porcine Preadipocytes Induced by BMP2. Cytogenetic and Genome Research
378   158, 133-44.

379 Li X., Wang H., Yao B., Xu W., Chen J. & Zhou X. (2016) lncRNA H19/miR-675 axis
380   regulates cardiomyocyte apoptosis by targeting VDAC1 in diabetic
381   cardiomyopathy. Sci Rep 6, 36340.

382 Liang G., Yang Y., Li H., Yu H., Li X., Tang Z. & Li K. (2018) LncRNAnet: a
383   comprehensive Sus scrofa lncRNA database. Animal Genetics 49, 632-5.

384 Miao Z., Wang S., Zhang J., Wei P., Guo L., Liu D., Wang Y. & Shi M. (2018)
385   Identification and comparison of long non-conding RNA in Jinhua and
386   Landrace pigs. Biochem Biophys Res Commun 506, 765-71.

387 Nguyen D.A. & Neville M.C. (1998) Tight junction regulation in the mammary gland.
388   J Mammary Gland Biol Neoplasia 3, 233-46.

389 Palombo V., Loor J.J., D'Andrea M., Vailati-Riboni M., Shahzad K., Krogh U. & Theil
390   P.K. (2018) Transcriptional profiling of swine mammary gland during the
391   transition from colostrogenesis to lactogenesis using RNA sequencing. Bmc
392   Genomics 19.

393 Pertea M., Pertea G.M., Antonescu C.M., Chang T.C., Mendell J.T. & Salzberg S.L.
394   (2015) StringTie enables improved reconstruction of a transcriptome from
395   RNA-seq reads. Nature Biotechnology 33, 290-5.

396 Quesnel H. & Farmer C. (2019) Review: nutritional and endocrine control of
397   colostrogenesis in swine. Animal 13, S26-S34.

398 Quesnel H., Farmer C. & Devillers N. (2012) Colostrum intake: Influence on piglet
399   performance and factors of variation. Livestock Science 146, 105-14.

400 St Laurent G., Wahlestedt C. & Kapranov P. (2015) The Landscape of long noncoding
401   RNA classification. Trends Genet 31, 239-51.

402 Vadmand C.N., Krogh U., Hansen C.F. & Theil P.K. (2015) Impact of sow and litter
403   characteristics on colostrum yield, time for onset of lactation, and milk yield of
404   sows. Journal of Animal Science 93, 2488-500.

405 Wan Y.H., Saghatelian A., Chong L.W., Zhang C.L., Cravatt B.F. & Evans R.M. (2007)
406   Maternal PPAR gamma protects nursing neonates by suppressing the production
407   of inflammatory milk. Genes & Development 21, 1895-908.

408 Wang J., Ren Q.L., Hua L.S., Chen J.F., Zhang J.Q., Bai H.J., Li H.L., Xu B., Shi Z.H.,
409   Cao H., Xing B.S. & Bai X.X. (2019a) Comprehensive Analysis of
410   Differentially Expressed mRNA, lncRNA and circRNA and Their ceRNA

19

411   Networks in the Longissimus Dorsi Muscle of Two Different Pig Breeds.
412   International Journal of Molecular Sciences 20.
413 Wang L., Cho K.B., Li Y., Tao G., Xie Z. & Guo B. (2019b) Long Noncoding RNA
414   (lncRNA)-Mediated Competing Endogenous RNA Networks Provide Novel
415   Potential Biomarkers and Therapeutic Targets for Colorectal Cancer.
416   International Journal of Molecular Sciences 20.
417 Wang Y., Hu T., Wu L., Liu X., Xue S. & Lei M. (2017) Identification of non-coding
418   and coding RNAs in porcine endometrium. Genomics 109, 43-50.
419 Xie C., Mao X., Huang J., Ding Y., Wu J., Dong S., Kong L., Gao G., Li C.Y. & Wei L.
420   (2011) KOBAS 2.0: a web server for annotation and identification of enriched
421   pathways and diseases. Nucleic Acids Res 39, W316-22.
422

Table 1 GO and KEGG enrichments of 23 potentially functional lncRNAs and their 25 target genes

LncRNA	Group	Target gene	ID	Term	*P* Value
			ssc04530	Tight junction	1.75E-03
XLOC_020627	-2vs-14	ACTN4*	GO:0035357	peroxisome proliferator activated receptor signaling pathway	4.03E-02
			ssc04670	Leukocyte transendothelial migration	1.92E-02
ENSSSCG00000051193	1vs-14	ADCY1*	ssc04921	Oxytocin signaling pathway	6.73E-03
			ssc04923	Regulation of lipolysis in adipocytes	3.09E-02
XLOC_025150*	-10vs-14, -6vs-14, -2vs-14, 1vs-14	CSN3*	GO:0030879	mammary gland development	1.49E-06
			GO:0007595	lactation	1.48E-03
			GO:0030879	mammary gland development	1.49E-06
			GO:0061180	mammary gland epithelium development	2.94E-03
ENSSSCG00000042618	-2vs-14	SMO*	GO:0060644	mammary gland epithelial cell differentiation	3.46E-02
			GO:0009792	embryo development ending in birth or egg hatching	4.80E-02
			GO:0071396	cellular response to lipid	2.42E-02
			GO:0008144	drug binding	3.27E-02
			GO:0009792	embryo development ending in birth or egg hatching	4.80E-02
XLOC_963181	-2vs-14, 1vs-14	PTK7*	GO:0071396	cellular response to lipid	2.42E-02
			GO:0008144	drug binding	3.27E-02
			GO:0032526	response to retinoic acid	4.35E-03
XLOC_026156	-10vs-14, -6vs-14, -2vs-14, 1vs-14	CD36*	ssc04920	Adipocytokine signaling pathway	6.51E-03
ENSSSCG00000011196*	-2vs-14, 1vs-14	GALNT15*	ssc00512	Mucin type O-glycan biosynthesis	6.33E-04
ENSSSCG00000036096	-6vs-14, -2vs-14, 1vs-14	CLDN6	ssc04530	Tight junction	1.75E-03
			ssc04670	Leukocyte transendothelial migration	1.92E-02
ENSSSCG00000042682	-6vs-14	RHBDF2	GO:0038127	ERBB signaling pathway	4.87E-03
			GO:0007173	epidermal growth factor receptor signaling pathway	3.85E-03
ENSSSCG00000041987	-6vs-14, -2vs-14, 1vs-14	C6H1orf210*	GO:0050688	regulation of defense response to virus	1.50E-02

Gene ID	Comparison	Gene	Term ID	Description	p-value
		CDC20*	GO:0051443	positive regulation of ubiquitin-protein transferase activity	4.03E-02
ENSSSCG00000051701	-2vs-14, 1vs-14	MPDZ*	ssc04530	Tight junction	1.75E-03
XLOC_018030	-2vs-14	NPR1*	ssc04921	Oxytocin signaling pathway	6.73E-03
			ssc04923	Regulation of lipolysis in adipocytes	3.09E-02
XLOC_025146	-10vs-14, -6vs-14, -2vs-14, 1vs-14	CSN2*	GO:0030879	mammary gland development	1.49E-06
			GO:0007595	lactation	1.48E-03
			GO:0030879	mammary gland development	1.49E-06
ENSSSCG00000041015	1vs-14	ATP2C2*	GO:0061180	mammary gland epithelium development	2.94E-03
			GO:0008144	drug binding	3.27E-02
XLOC_005878	-6vs-14, -2vs-14	MEOX1	GO:0009792	embryo development ending in birth or egg hatching	4.80E-02
XLOC_010589*	-2vs-14, 1vs-14	ACSL3*	ssc04920	Adipocytokine signaling pathway	6.51E-03
		GALNTL5*	ssc00512	Mucin type O-glycan biosynthesis	6.33E-04
ENSSSCG00000046607	-6vs-14, -2vs-14, 1vs-14	PRKAG2*	ssc04530	Tight junction	1.75E-03
			ssc04921	Oxytocin signaling pathway	6.73E-03
			ssc04920	Adipocytokine signaling pathway	6.51E-03
			GO:0030879	mammary gland development	1.49E-06
			GO:0071396	cellular response to lipid	2.42E-02
ENSSSCG00000050015	-6vs-14, -2vs-14, 1vs-14	SOX9	GO:0038127	ERBB signaling pathway	4.87E-03
			GO:0032526	response to retinoic acid	4.35E-03
			GO:0007173	epidermal growth factor receptor signaling pathway	3.85E-03
ENSSSCG00000042000	-2vs-14, 1vs-14	SNRPD1*	GO:0050688	regulation of defense response to virus	1.50E-02
XLOC_002435	-2vs-14	FIBCD1*	GO:0008144	drug binding	3.27E-02
			ssc04530	Tight junction	1.75E-03
ENSSSCG00000049698	-2vs-14	ACTB*	ssc04921	Oxytocin signaling pathway	6.73E-03
			ssc04670	Leukocyte transendothelial migration	1.92E-02
			GO:0008144	drug binding	3.27E-02

ENSSSCG00000048264	1vs-14	GALNT7*	ssc00512	Mucin type O-glycan biosynthesis	6.33E-04
XLOC_003621	1vs-14	MAP3K8	GO:0008144	drug binding	3.27E-02

424   Note: * indicates the lncRNAs and their target genes involved in the significant module by WGCNA.
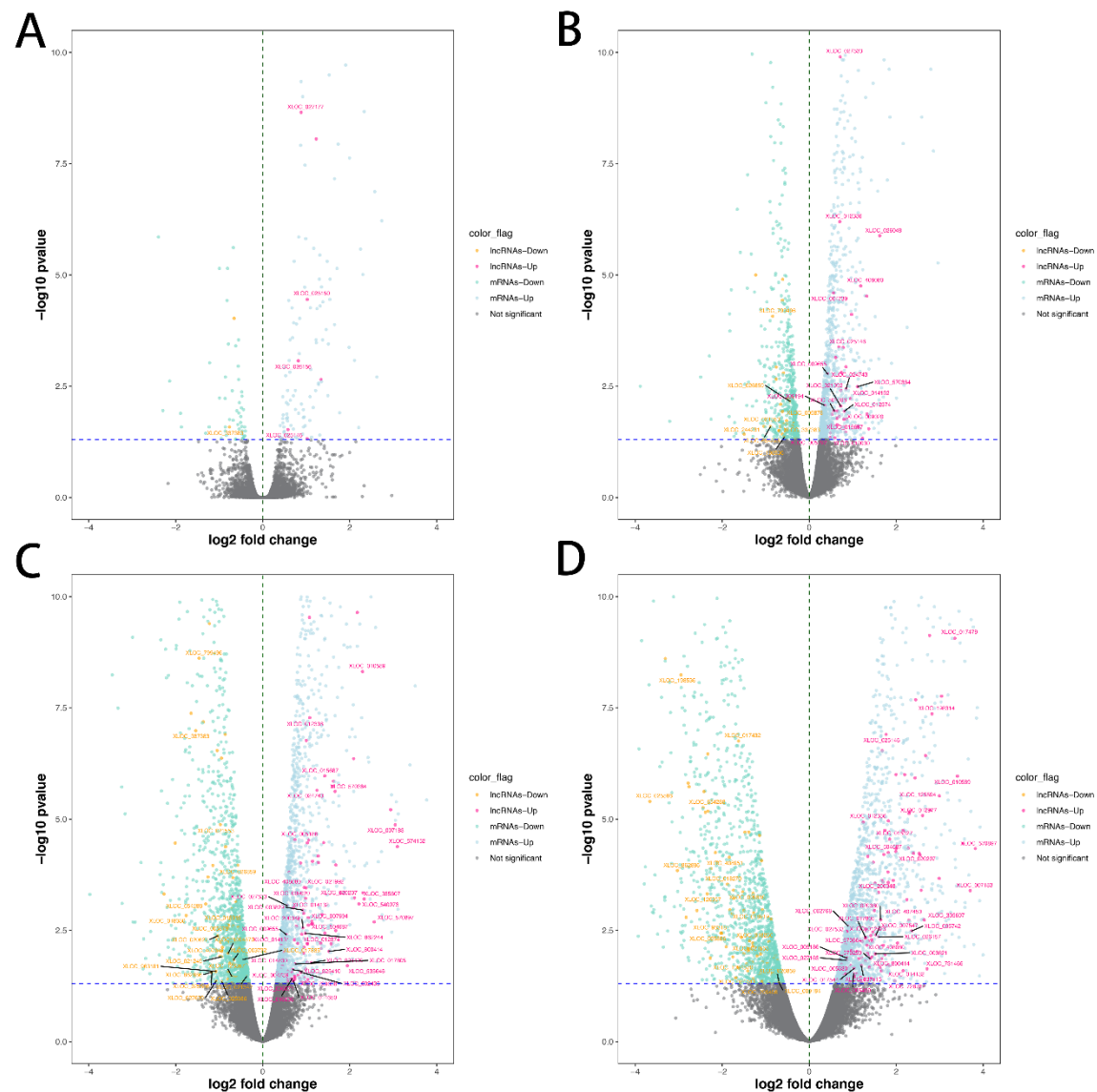
425

Figure 1 Identification of differentially expressed (DE) lncRNAs and mRNAs. (A) Volcano plot dispalys DE lncRNAs and mRNAs of Day -10 vs Day -14. **(B)** Volcano plot dispalys DE lncRNAs and mRNAs of Day -6 vs Day -14. **(C)** Volcano plot dispalys DE lncRNAs and mRNAs of Day -2 vs Day -14. **(D)** Volcano plot dispalys DE lncRNAs and mRNAs of Day +1 vs Day -14. The novel DE lncRNAs was marked.
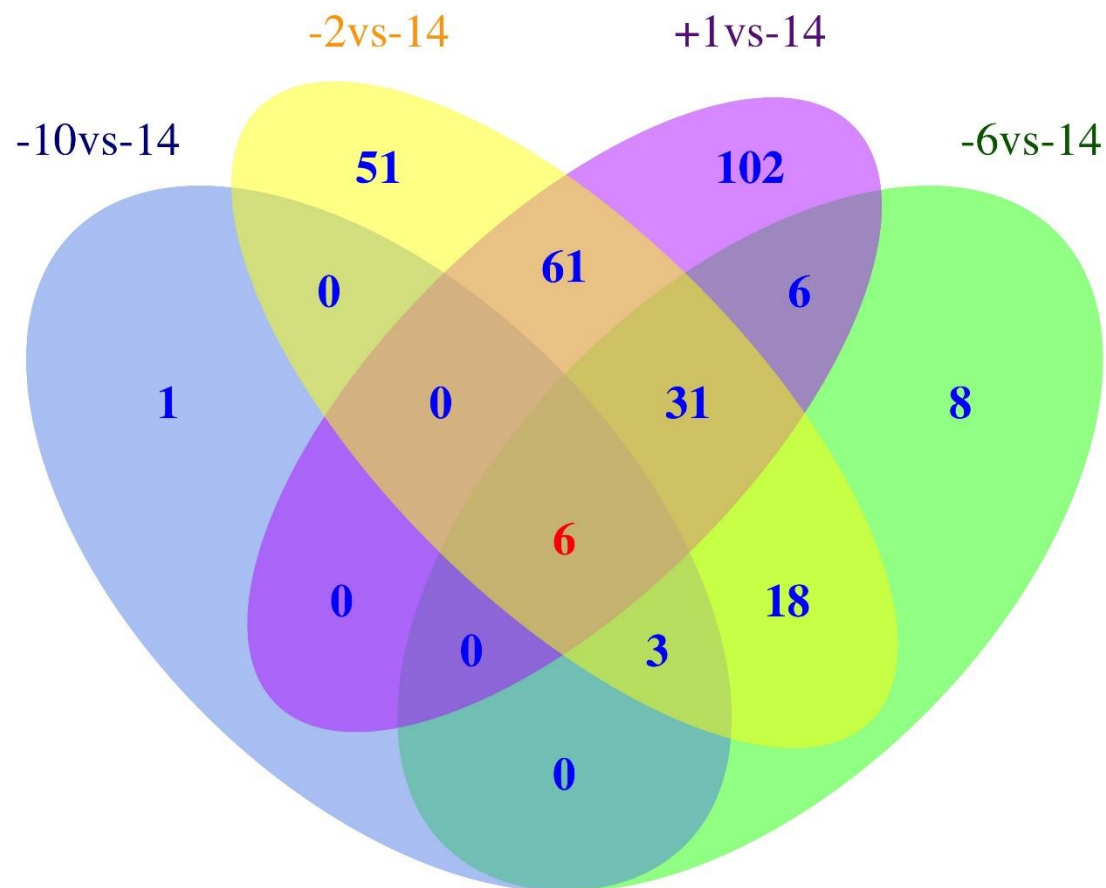
Figure 2 Venn diagram shows the number of differentially expreesed (DE) lncRNAs overlapping in different swine mammary gland development stages or unique to each developmental stage. -14: at days 14 before parturition. -10: at days 10 before parturition. -6: at days 6 before parturition. -2: at days 2 before parturition. +1: at days 1 after parturition.
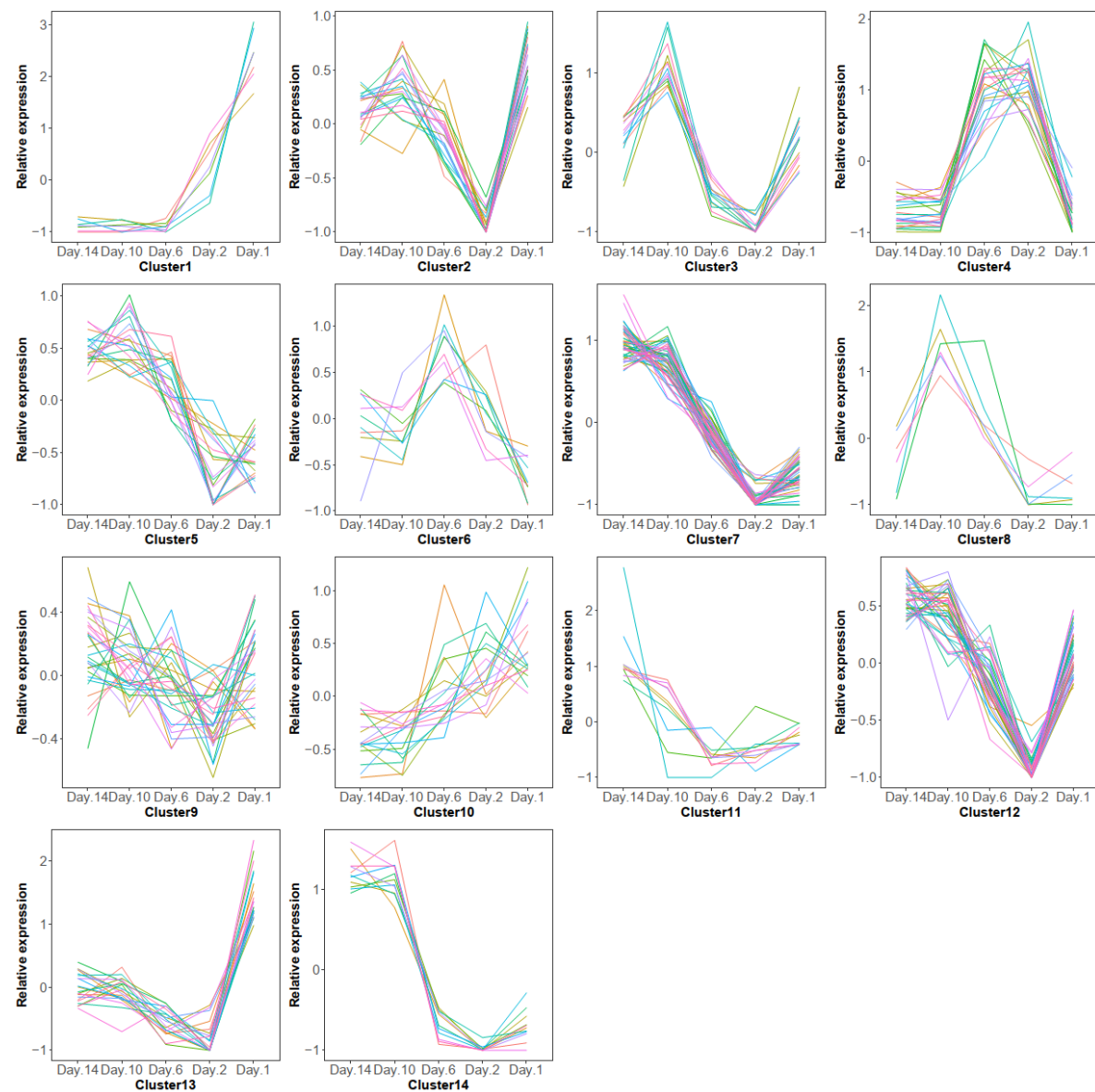
Figure 3 Expression pattern of differentially expreesed (DE) lncRNAs across five stages. The colored lines represent that the different lncRNA changes.
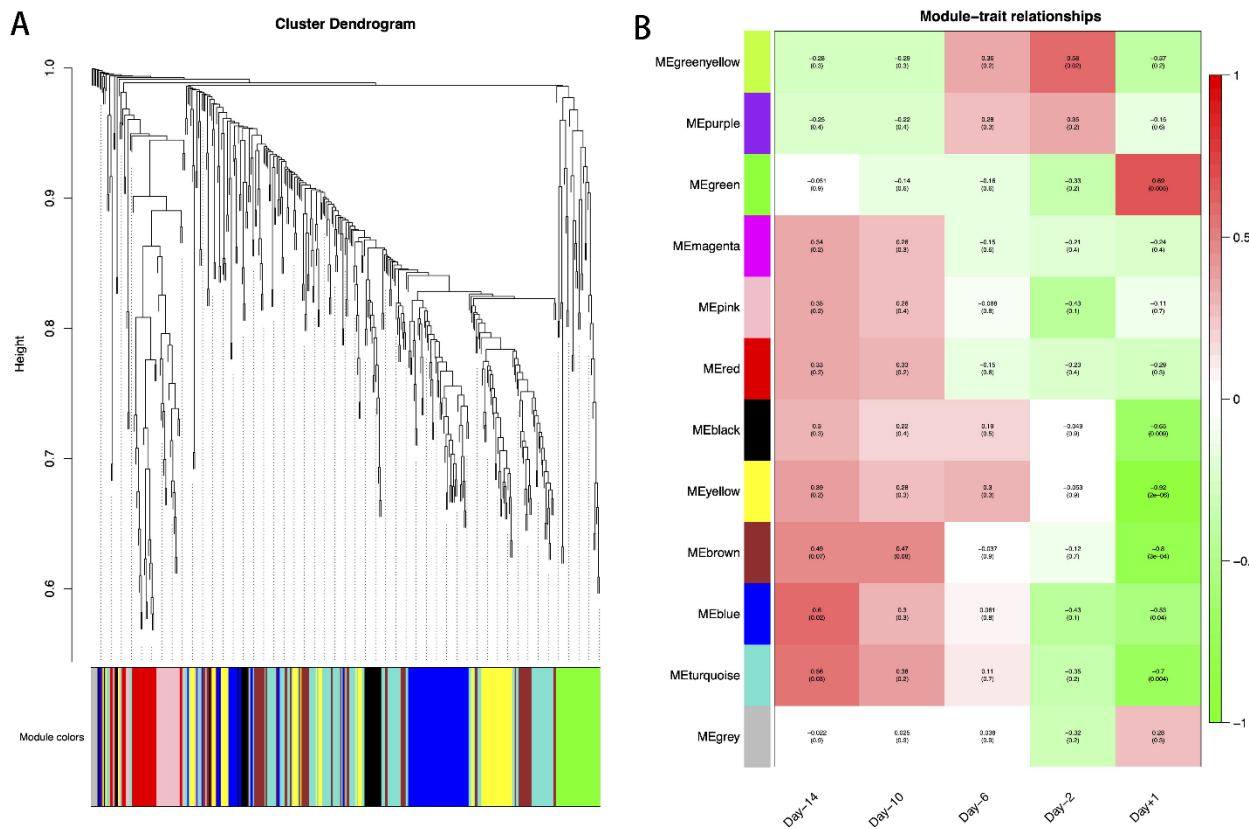
Figiure 4 Weighted gene co-expression network analysis (WGCNA) at each swine mammary gland development. (A) Hierarchical cluster tree showing co-expression modules identified by WGCNA. **(B)** Module-sample relationships. Each column indicates a sample, and each row represents a module. The correlation coefficient and *p*-value of each sample-module relationship are displayed. Red represents high correlation value, and green represents low correlation value.
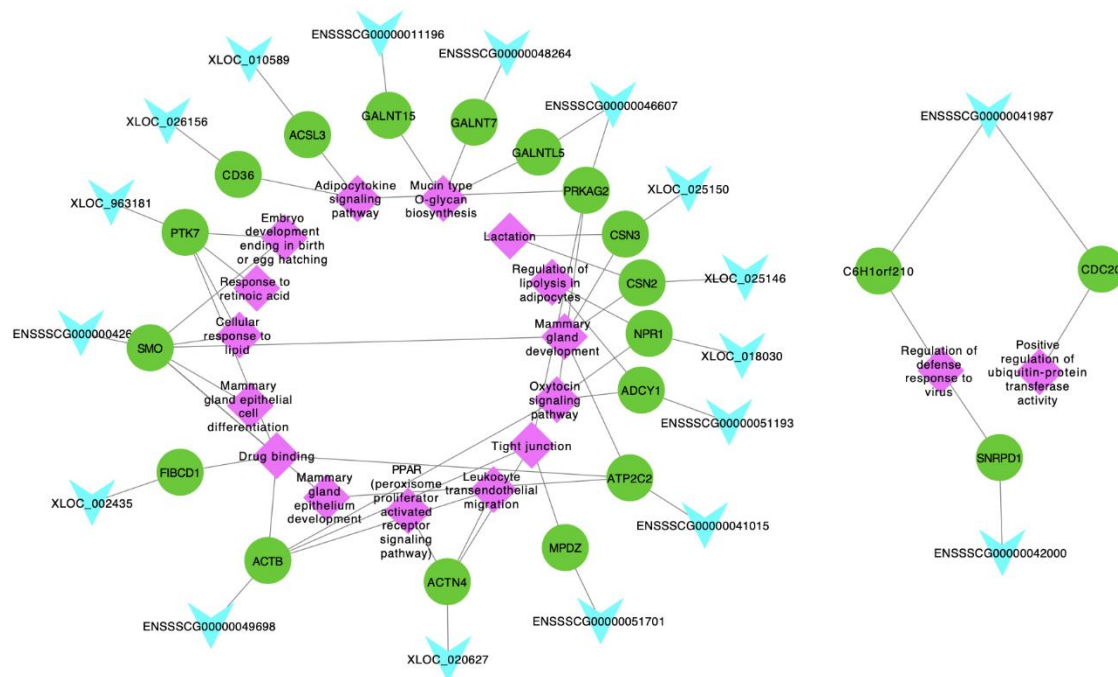
Figure 5 Network plot of the promising candidate lncRNAs, genes and pathways.
The light blue, green and purple indicate lncRNAs, genes and pathways, respectively.