# Transcriptome Analysis Suggests the Roles of Long Intergenic Non-coding RNAs in the Growth Performance of Weaned Piglets

Lin Chen[1], Gaoli Shi[1], Guoting Chen[1], Jingxuan Li[1], Mengxun Li[1], Cheng Zou[1], Chengchi Fang[1] and Changchun Li[1,2]*

[1] Key Lab of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education, College of Animal Sciences and Technology, Huazhong Agricultural University, Wuhan, China, [2] The Cooperative Innovation Center for Sustainable Pig Production, Wuhan, China

Long intergenic non-coding RNAs (lincRNAs) have been considered to play a key regulatory role in various biological processes. An increasing number of studies have utilized transcriptome analysis to obtain lincRNAs with functions related to cancer, but lincRNAs affecting growth rates in weaned piglets are rarely described. Although lincRNAs have been systematically identified in various mouse tissues and cell lines, studies of lincRNA in pigs remain rare. Therefore, identifying and characterizing novel lincRNAs affecting the growth performance of weaned piglets is of great importance. Here, we reconstructed 101,988 lincRNA transcripts and identified 1,078 lincRNAs in two groups of longissimus dorsi muscle (LDM) and subcutaneous fat (SF) based on published RNA-seq datasets. These lincRNAs exhibit typical characteristics, such as shorter lengths and lower expression relative to protein-encoding genes. Gene ontology analysis revealed that some lincRNAs could be involved in weaned piglet related processes, such as insulin resistance and the AMPK signaling pathway. We also compared the positional relationship between differentially expressed lincRNAs (DELs) and quantitative trait loci (QTL) and found that some of DELs may play an important role in piglet growth and development. Our work details part of the lincRNAs that may affect the growth performance of weaned piglets and promotes future studies of lincRNAs for molecular-assisted development in weaned piglets.

Keywords: lincRNA, RNA-seq, muscle and fat development, growth performance, weaned piglets

## INTRODUCTION

Over the past decade, given developments in second-generation sequencing technologies, an increasing number of transcriptome studies have revealed a large number of non-coding transcripts in the genome of mammal (Pauli et al., 2012). The expression levels of these genes are somewhat lower than those of protein-coding genes and are often considered to be transcriptional noises

(Li et al., 2013). However, as science and technology have advanced, researchers have gradually realized that non-coding transcripts may possess their own unique regulatory mechanisms; these transcripts are called lincRNAs because they are located between genes and have a structure longer than 200 nt (Mizutani et al., 2012). New evidences reveal that, although they have low coding potential, lincRNAs present certain functions in many cellular processes, including gene regulation, apoptosis and embryonic development (Leung et al., 2013). Four functional types of lincRNA, including signal, decoy, guidance and support, have been roughly identified, and most of these transcripts have been studied in mice (Wang and Chang, 2011; Hong et al., 2018). Research involving pig lincRNA is minimal and thus deserves further exploration (Ran et al., 2016).

Weaned piglets, as a very important class of economic animal, are a typical biological model (Fisar and Hroudova, 2016). Newborn weaned piglets have low feed intake, slow growth and high mortality; thus, the piglets cannot gain weight easily (Lv et al., 2018). During the weaning transition, even if the environment and nutrient feed are adjusted, guaranteeing the growth rate of these animals is difficult (Jones et al., 2012). Previous studies have shown that inadequate energy intake and reduced lipid deposition rates are key factors of post-weaning dysplasia in piglets (Snoeck et al., 2004; Jones et al., 2012). We speculate that the development of muscle and fat (the key tissues for piglet development) during the weaning transition may reflect differences in the growth performance of weaned piglets (Yang H. et al., 2016; Albuquerque et al., 2017). In a previous study, pigs showed very obvious obesity, which was not found in other animals (Madeira et al., 2017). The pig model has been used to explore important diseases caused by excess fat, such as obesity, diabetes, coronary heart disease and arteriosclerosis (Robich et al., 2010). Studying the muscle growth of pigs can also enable the breeding of lean pigs to improve pork utilization (Zhang et al., 2015). However, compared with those of human and mouse, research on pig lincRNAs is scarce (De Quattro et al., 2018). The molecular mechanism of the development of muscle and fat in weaned piglets is yet to be explored.

In this study, we first performed a transcriptome assembly containing muscle and fat using data downloaded from GEO accession number GSE65983. We identified a total of 1,078 putative lincRNAs and characterized these lincRNAs. We then analyzed the expression of these lincRNAs in all samples and identified some DELs. We utilized differential expression analysis to explore DELs that may contribute to decreased growth performance in weaned piglets using the RNA-Seq transcriptome, based on the LDM and SF, to analyze the molecular basis of weaned piglet skeletal muscle and fat development. GO analysis was performed on the adjacent protein-coding genes of DELs to predict their function. Real-time quantitative PCR was used to verify the reliability of data analysis. This study not only enriches the lincRNA database for weaned piglet skeletal muscle and fat development, but also provides new insights into the molecular mechanisms affecting the growth performance of weaned piglets.

## MATERIALS AND METHODS

### Ethics Statement and Data Sources

All the experiments in our study were carried out according to the Key Laboratory of Agricultural Animal Genetics Breeding and Reproduction, Animal Protection and Use Committee, Wuhan 430070. In this study, the piglets used for sampling were all from the same batch and were raised under the same environmental and nutritional conditions (Pilcher et al., 2015). All samples were taken from the same part of subcutaneous back fat between the eighth and last ribs and LDM (Pilcher et al., 2015). 36 RNA-seq datasets from two different kinds of tissues, including LDM and subcutaneous back fat (18 replicates), were obtained from a previously published study and downloaded from the NCBI's GEO under GEO data set ID GSE65983 (Pilcher et al., 2015). The pig gene annotations were downloaded from ftp://ftp.ensembl.org/pub/release-93/gtf/sus_scrofa. Moreover, the NR reference sequence (RefSeq)NR database was downloaded from ftp://ftp.ncbi.nih.gov/blast/db/.
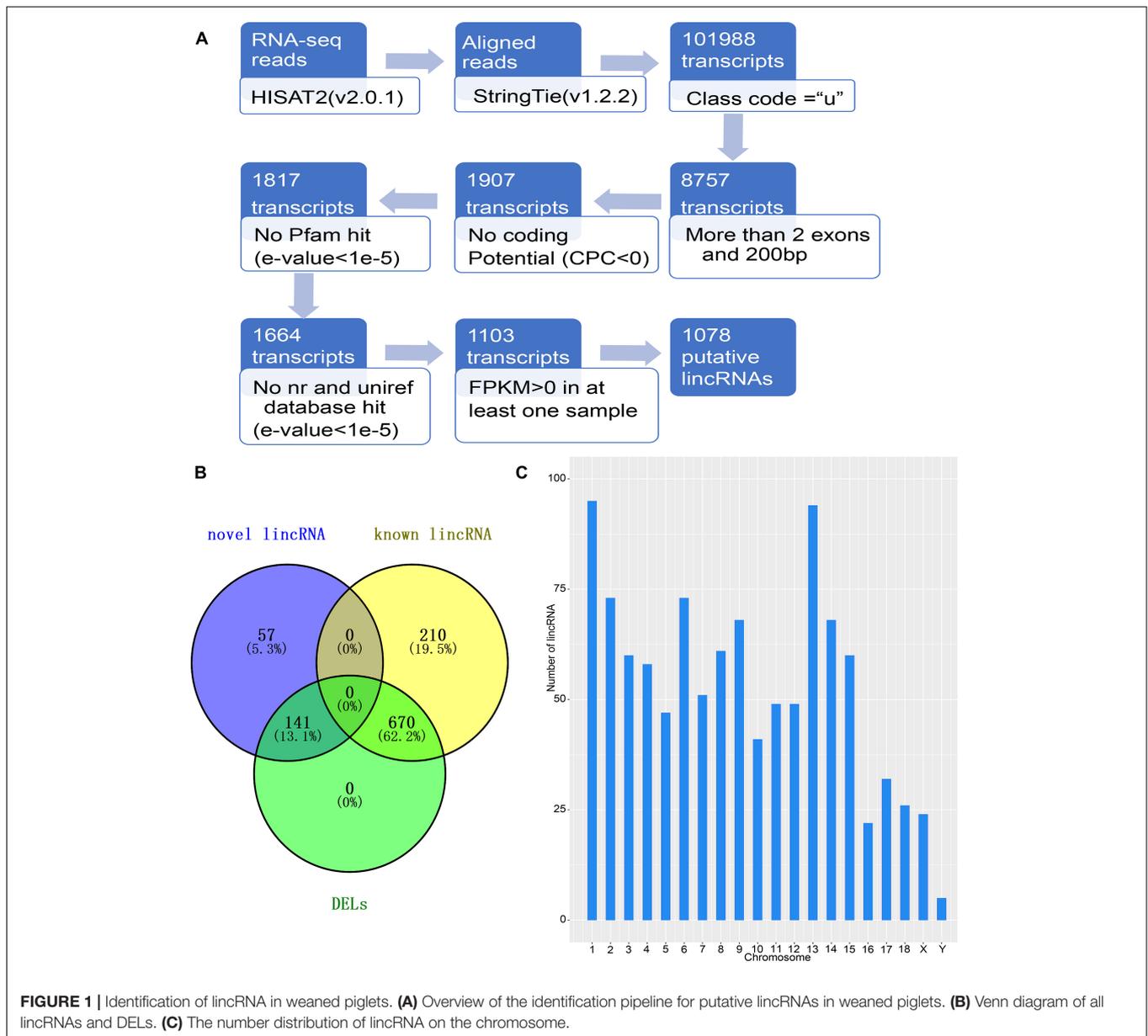
### RNA-Seq Reads Mapping and Transcriptome Assembly

To ensure that high-quality data is provided to the initial spliced alignment step, the quality of raw reads was evaluated using FastQc (Xiao et al., 2015). Raw reads were filtered and trimmed using Trimmomatic (version0.3.2) with the default parameters (Xiao et al., 2015). The adapter and low-quality reads were removed for the number of mismatches in a read. The remaining high-quality clean reads were then mapped to the pig reference genome (SusScrofa11.1[1]) with the default parameters by the spliced reads aligner HISAT2(version 2.0.1) (Pertea et al., 2016). According to different database construction methods, we selected the corresponding parameters "known-splicesite-infile" and got the sam file containing all mapping readings for each sample. Using samtools software, the sam files were processed and converted into sorted bam files with parameters "view -H" and "sort -o" (Pertea et al., 2016). Finally, to facilitate transcript assembly and quantification, all bam files were assembled into one complete GTF files with the default parameters and merge function of StringTie (version1.2.2) software (Pertea et al., 2016). In addition, by using the assembled GTF file, StringTie software is able to estimate the expression levels of genes and transcripts in all samples for subsequent studies with the parameters "-e" and "-B" (Pertea et al., 2016).

### Pipeline for lincRNA Identification

The detection of NR transcriptomes for the detection of lincRNAs is basically based on previous laboratory studies (Zou et al., 2017a,b). The principal pipeline employed for lincRNA identification was as follows (**Figure 1A**): (1) filter transcripts that without 'u' characters to retain intergenic transcripts with the parameters "-r" of gffcompare in StringTie (version1.2.2) software (Pertea et al., 2016); (2) remove transcripts with single exons and less than 200 bp in length according to the merged

---

[1]ftp://ftp.ensembl.org/pub/release-93/fasta/sus_scrofa/dna/

**FIGURE 1 |** Identification of lincRNA in weaned piglets. **(A)** Overview of the identification pipeline for putative lincRNAs in weaned piglets. **(B)** Venn diagram of all lincRNAs and DELs. **(C)** The number distribution of lincRNA on the chromosome.

GTF file; (3) collect the positional information of the exons of the remaining transcripts, extract the genomic sequence of the corresponding exon, and fuse into a complete transcript sequence; (4) Use the CPC tool to assess the protein coding potential of complete transcript sequences, retaining transcripts that cannot encode proteins based on protein coding potential (Kong et al., 2007); (5) eliminate transcripts containing any known protein coding domain. Using Transeq[2] to translate transcript sequences into six possible protein domains and if the corresponding transcripts had a significant hit in the Pfam database[3], they were discarded by HMMER ($E$-value $<$ 1e-5) (Finn et al., 2015); (6) Using the BLASTX program to filter out any transcript which is similar to a known protein with an $E$-value $<$ 1e-5 in the NCBI, NR, and the UniRef90 database (Pirooznia et al., 2008); (7) Retained transcripts with FPKM values greater than 0 in at least one sample as candidate lincRNAs.

## Comparisons Between lincRNAs and Protein-Coding Transcripts

We selected the protein-encoded character as a marker in the transcript of the Ensemble gene annotation of the pig, resulting in a total of 45,788 protein-coding transcripts. We then identified the protein-coding transcripts and all lincRNA transcripts in terms of length, number and length of exons, and expression levels. Finally, we compared the different characteristics in transcript length, exon length, exon number and FPKM between

---

[2]https://www.ebi.ac.uk/Tools/st/emboss_transeq/

[3]http://pfam.xfam.org/search

lincRNA transcripts and protein-encoding gene transcripts and visualized them using R language, with *p*-value less than 0.05 was considered to be significant.

## Analysis of Adjacent Genes

Previous studies have shown that mammalian lncRNAs are preferentially located next to genes with developmental functions (Cabili et al., 2011). Therefore, for each DEL locus the nearest protein-coding neighbor within <100 kb was identified by BEDTools (version 2.17.0), and these neighboring genes are expressed in at least one sample (Quinlan and Hall, 2010). This resulted in a list of lncRNA loci/protein-coding loci pairs. Pearson correlation was used to explore the expression-based relationship between these pairs (Cabili et al., 2011). Ensembl was used to convert the protein-coding gene genes id into human genes id for annotation, visualization and integrated discovery (DAVID) (Huang et al., 2009). Then, we performed GO analysis and KEGG pathway analysis on all expressed protein-coding gene of DELs (Xing et al., 2016). Terms with a *p*-value < 0.05 were considered statistically significant (Benjamini and Cohen, 2017).

## Differentially Expressed lincRNAs Analysis

To perform differential expression analysis on these lincRNAs, we divided 36 samples into two groups, using DEseq2 to generate differential expression of all transcripts (Love et al., 2014). We screened the differentially expressed transcripts by using the fold change with an absolute value greater than 1 and an adjusted *p*-value of less than 0.05, and then crossed the entire lincRNA to obtain DELs. Finally, the results were processed by the DEseq2 package in the R computing environment to produce global gene expression figures and identify differentially expressed protein-coding and lincRNA transcripts based on $p_{adj} < 0.05$ (Benjamini et al., 2001).

## Functional Prediction of DELs by QTL

We further predicted the function of lincRNA by comparing the position of QTL and lincRNA on the whole genome to determine their overlapping regions with BEDTools2.17.0. The QTL database was downloaded from Animal QTL[4]. A total of 26,348 QTLs can be divided into five categories, including (meat and carcass, production, reproduction, health, and exterior) (Zhao et al., 2018).

## Validation of Gene Correlation by qPCR

To verify the association between the protein-coding gene and the lincRNA gene, 10 protein-coding genes and ten lincRNAs were selected for qPCR validation, the corresponding gene sequences were extracted and the gene-specific primers were designed using the Primer5 program. All RNA was extracted from the muscle and fat of three pigs using Trizol reagent (Invitrogen, Life Technologies, Carlsbad, CA, United States)

---

[4]https://www.animalgenome.org/cgi-bin/QTLdb/SS/index

from PSC, and the sample concentration was measured using the Agilent 2100 system. According to the manufacturer's instructions, SYBR Green (CWBIO, Beijing, China, CW0957) was used as a fluorescent dye, and 18S rRNA was used as an internal reference to detect the lncRNA gene and protein-coding gene in the Roche LightCyler 480 system (Roche, Mannheim, Germany). The delta delta *C*t ([Delta][Delta]Ct) method was used to analyze qPCR data and quantify gene expression (Rao et al., 2013).
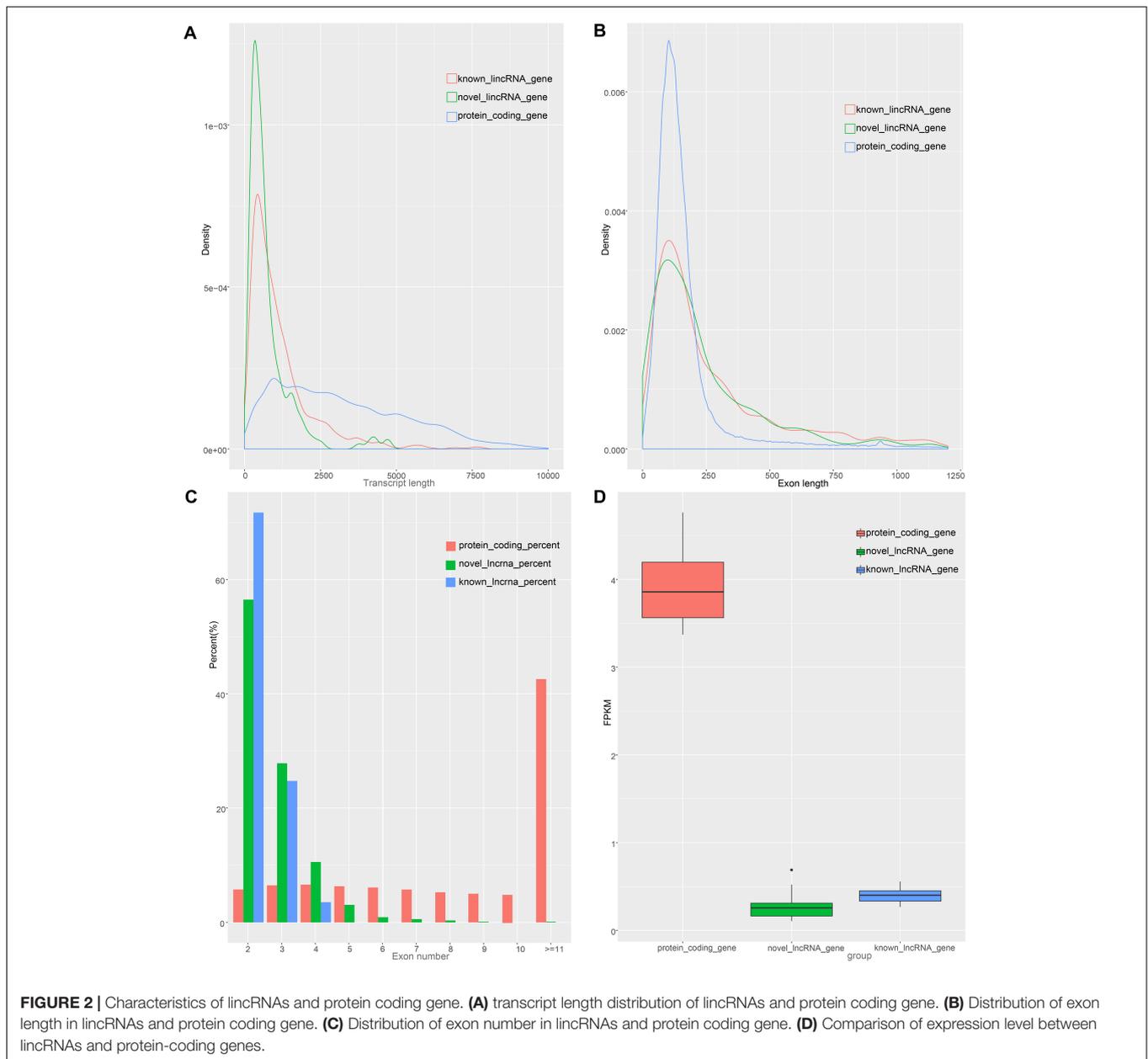
# RESULTS

## Genome-Wide Identification of lincRNAs Based on RNA-Seq Datasets

To identify the lincRNAs involved in the muscle and adipose tissue of weaned piglets, we used published data for RNA-seq (**Supplementary Table S1**) involving two groups of 36 samples (Pilcher et al., 2015). The RNA-seq datasets were sequenced on the Illumina Hiseq2000 platform (Pilcher et al., 2015). And the reads in the datasets are paired, with read lengths exceeding 200 bp (Liu et al., 2018). Based on the pipeline, which is explained in detail in the Section "Materials and Methods," we obtained new lincRNAs, as shown in **Figure 1A**. Briefly, beginning from a total of 115.6 million reads, we performed short-read gapped alignment using Hisat2 and recovered 108.5 million mapped reads onto the pig genome version 11.1 (Hong et al., 2018). Then, we used the publicly available assemble software StringTie to *de novo* reconstruct the transcriptome for each tissue based on the read-mapping results (Pauli et al., 2012). The reconstructed transcripts were merged into a single set of transcripts using the merge utility provided by StringTie. We first compared 101,988 transcripts with known genes and obtained a set of 8,757 potential intergenic transcripts. Then, we filtered by the exon number and transcript length and acquired 1,907 potential long intergenic transcripts. Next, we analyzed the coding potential of unannotated transcripts with the CPC tools, which retained 1,817 transcripts. We also scanned each transcript against a six-class protein framework to exclude any transcript that has or may have protein encoding ability. This filtering retained 1,664 potential lincRNA transcripts. We performed a sequence homology search with the BLAST tool and obtained 1,103 potential lincRNA transcripts. Using FPKM expression level of more than 0 at least in one tissue sample as a basis, we finally obtained 1,078 reliably putative lincRNAs, including 198 novel lincRNAs and 880 known lincRNAs (**Supplementary Table S3** and **Figure 1B**). We analyzed the position of each lincRNA and found that most lincRNAs are concentrated on chromosomes 1 and 13, which may be caused by the long length of these two chromosomes (**Figure 1C**).

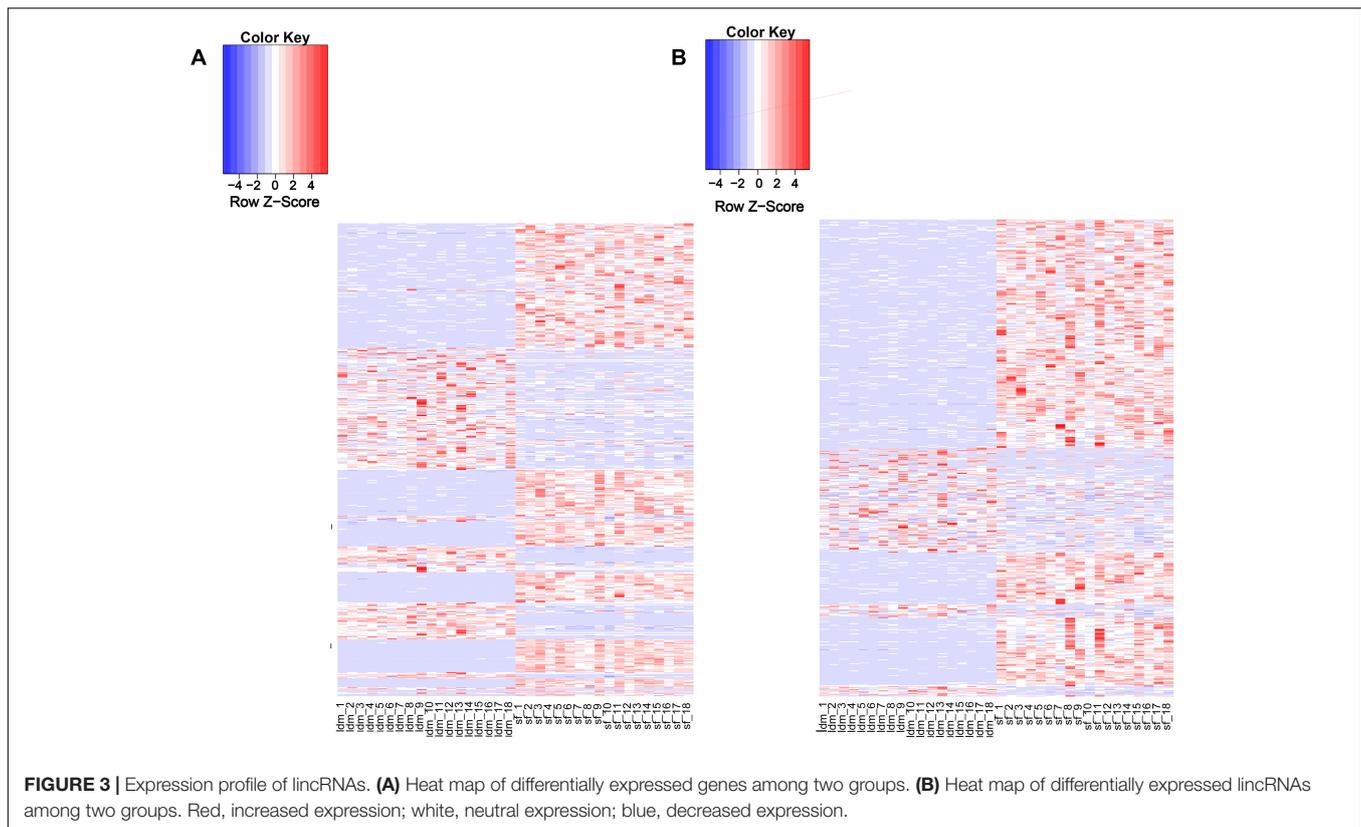## Characteristics of Known and Novel lincRNAs

To identify the characteristics of the lincRNA transcripts, we divided the lincRNA into two major classes, namely known and novel, based on the reconstructed transcriptome and compared

**FIGURE 2 |** Characteristics of lincRNAs and protein coding gene. **(A)** transcript length distribution of lincRNAs and protein coding gene. **(B)** Distribution of exon length in lincRNAs and protein coding gene. **(C)** Distribution of exon number in lincRNAs and protein coding gene. **(D)** Comparison of expression level between lincRNAs and protein-coding genes.

them to the obtained protein-coding transcript. According to the Ensembl database, we acquired 45,788 protein-coding transcripts corresponding to the 22,342 protein-coding genes annotated in pigs (**Supplementary Table S2**), and obtained 12,103 known lincRNA transcripts corresponding to 7,381 lincRNA genes in ALDB. In our study, the average transcript length of the new lincRNA gene was 776 bp, while the average transcript length of the known lincRNA genes was 1,159 bp; these lengths are both shorter than that of protein-coding gene (3,284 bp) (**Figure 2A**). In term of average exon length, the novel lincRNA gene is 334 bp in length, shorter than the known lincRNA gene (431 bp) but longer than the protein-coding genes (283 bp) (**Figure 2B**). In addition, we found that the average number of exons in the protein-coding gene is 11.7, which is significantly

higher than that of the new lincRNA genes (2.3) and the known lincRNA gene (2.7) (**Figure 2C**). Because lincRNAs lack protein coding ability, we compared the expression levels of 1,078 lincRNAs with the expression levels of protein-coding genes (**Supplementary Table S4**). The results showed that the FPKMs of the lincRNA genes and the unknown lincRNA genes are 0.39 and 0.25 (**Supplementary Table S5**), respectively, which are both lower than the average expression level of the protein-coding gene (FPKM: 3.9) (**Supplementary Table S7** and **Figure 2D**). In general, lincRNA genes exhibit shorter transcription lengths, longer exon lengths, fewer exons, and lower expression levels than protein-encoding genes. All of which, as previously reported, are typical features of lincRNA (Cabili et al., 2011; Djebali et al., 2012).

**FIGURE 3 |** Expression profile of lincRNAs. **(A)** Heat map of differentially expressed genes among two groups. **(B)** Heat map of differentially expressed lincRNAs among two groups. Red, increased expression; white, neutral expression; blue, decreased expression.
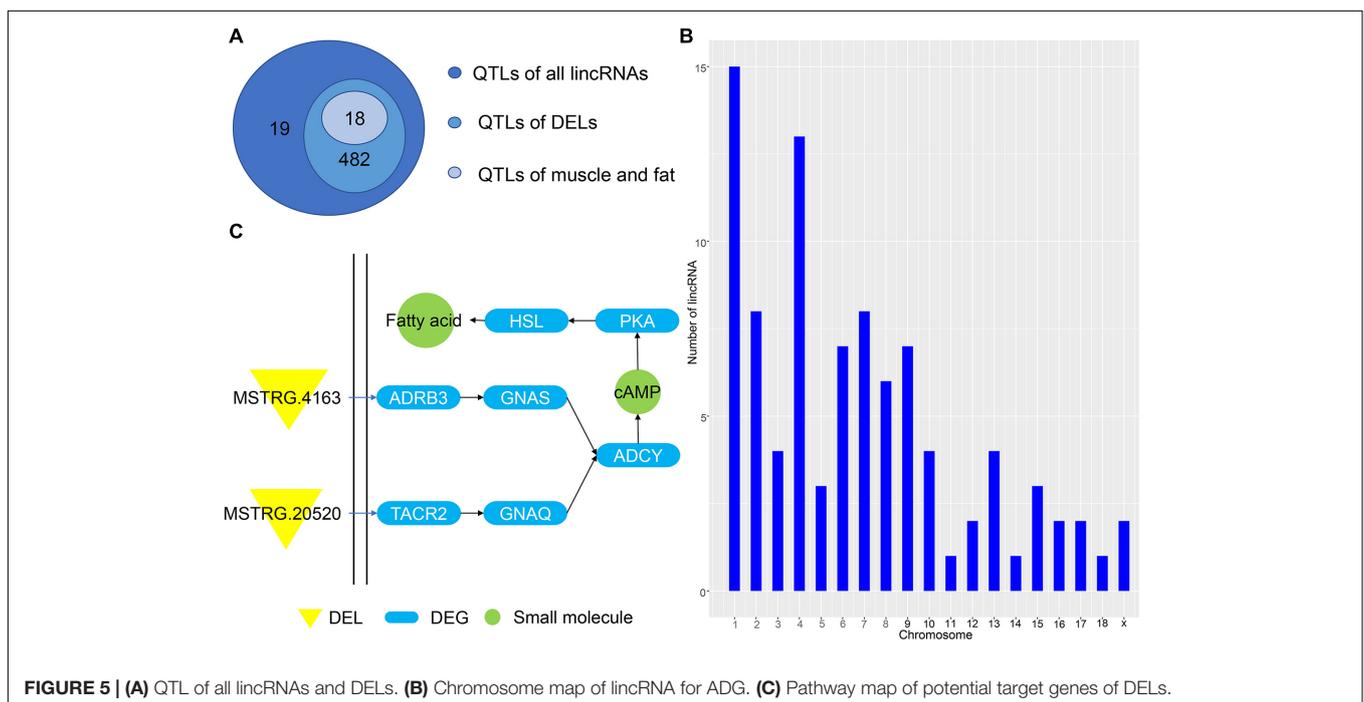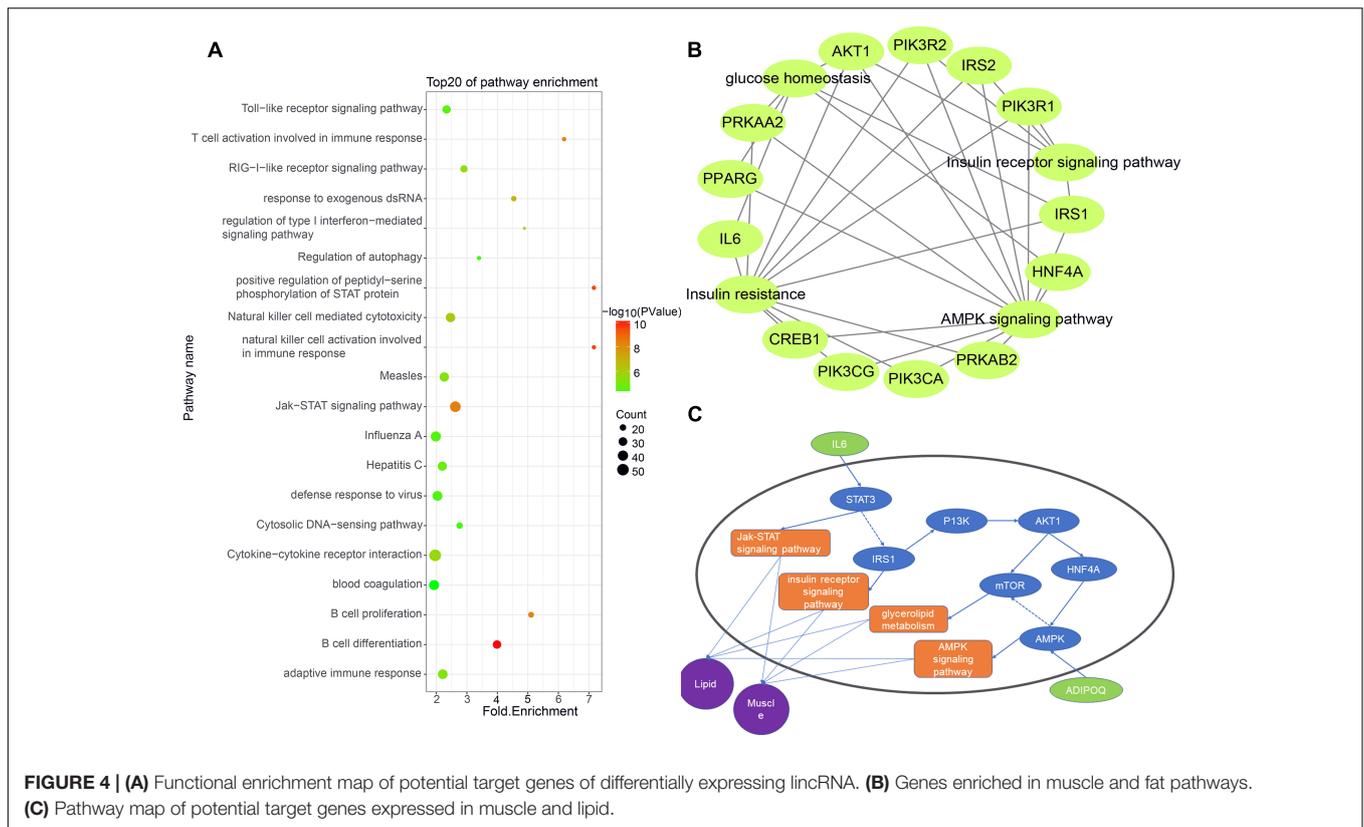
## Tissue Expression Profile of lincRNAs

Recent studies have shown that lincRNAs are expressed more tissue-specifically than protein-coding genes (Birney et al., 2007; Nakaya et al., 2007). We analyzed the expression patterns of all DELs and protein-encoding genes (**Supplementary Table S6**). About 97% (1078/1103) of the potential lincRNAs were expressed in more than one sample, and 54.8% (591/1078) lincRNA genes showed differential expression in both tissue types (**Supplementary Table S9**). By contrast, among all protein-coding genes, only 2.7% (601/22342) were differentially expressed in the two tissues (**Supplementary Table S8**). The proportion of DEL genes was significantly greater than differentially expressed protein-encoding genes. Therefore, we suggest that the lincRNA genes show a more tissue-specific expression landscape than protein-coding genes (**Figures 3A,B**). To explore the function of lincRNAs, after profiling 880 known lincRNAs and 198 novel expression in all groups, we conducted the differential expression analysis among the two groups using DEseq2 in the R package and detected 811 DELs. Using *p*-value and fold change as bases, we found 254 up-regulated and 337 down-regulated DELs in the LDM and SF groups.

## Nearest-Neighbor Analysis of lincRNAs

Some studies have shown that lincRNAs may regulate their neighboring genes in a *cis*-acting manner (Cabili et al., 2011). To explore the expression relationship between lincRNAs and their neighboring protein-coding genes, we performed GO analysis and KEGG analysis on expressed protein-coding genes

transcribed near lincRNA (<100 kb) (Zhao et al., 2015). Using Ensembl and DAVID, we obtained relevant pathway information for these protein-coding genes. The results showed that 955 of 2,251 protein-coding genes significantly participated in 84 biological processes and 37 pathways (**Supplementary Table S2**). Many protein-coding genes are involved in biological processes or pathways related to muscle growth and fat deposition, such as glucose homeostasis, lipid transport and insulin receptor signaling pathway. Interestingly, at the top 20 of these biological processes and pathways is the Jak-STAT signaling pathway, which corresponds to the biological processes of positive regulation of peptidyl-serine phosphorylation of STAT protein (**Figure 4A**).

We selected biological processes and pathways related to muscle development and fat deposition from all the results, including glucose homeostasis, glycogen metabolism, lipid transport, cortical actin cytoskeletal organization, the insulin receptor signaling pathways, insulin resistance and the AMPK signaling pathways (**Figures 4B,C**). Enrichment analysis of the protein-coding genes in these pathways revealed that 12 genes appeared multiple times, with the AKT and IRS1 genes appearing four times in all pathways. According to previous studies, AKT may act on MYOD to affect muscle differentiation (Li X. et al., 2018). Moreover, Li D. et al. (2018) found that lncHR1 can act on AKT to inhibit SREBP1C level and regulate fat metabolism. Zhu et al. (2018) found that miR-146b regulates glucose homeostasis in porcine primary adipocytes by targeting IRS1. These studies indicate that the target genes we found are indeed capable of regulating muscle development and fat metabolism.

**FIGURE 4 | (A)** Functional enrichment map of potential target genes of differentially expressing lincRNA. **(B)** Genes enriched in muscle and fat pathways. **(C)** Pathway map of potential target genes expressed in muscle and lipid.



**FIGURE 5 | (A)** QTL of all lincRNAs and DELs. **(B)** Chromosome map of lincRNA for ADG. **(C)** Pathway map of potential target genes of DELs.

## Function Prediction of lincRNA

Predicting the functions of lincRNAs remains challenging due to their lack of annotation and low expression levels (Han et al., 2016). Here, we analyzed the positional relationship between QTLs and lincRNAs to determine their overlapping regions to infer the function of lincRNA. We compared all lincRNAs and DELs to the porcine QTL database to speculate the function of the identified lincRNA. A total of 1078 lincRNAs overlapped with

501 QTL regions, of which 811 DELs corresponded to 482 QTLs (**Figure 5A**). Eighteen QTLs had multiple functions, including average daily gain, muscle fiber development and intramuscular fat; all of which are related to the muscle and fat development of weaned piglets. Then, we studied the average daily gain of QTLs and found that 93 lincRNAs fall into the QTL interval and are distributed on all chromosomes (**Figure 5B**). These lincRNAs are concentrated on chromosomes 1 and 4, consistent with previous results in **Figure 1**, thereby indicating that most of the identified lincRNAs are associated with average daily gain. In order to understand the specific functions of DEL, we enriched several pathways using GO analysis, the most important of which is the cAMP pathway. In this pathway, we found that the DEL MSTRG.4163 can regulate the differentially expressed potential target genes ADRB3 (**Figure 5C**). Previous studies have confirmed that activation of ADRB3 stimulates cAMP to act on PKA and promotes lipolysis through HSL phosphorylation (Shin et al., 2016). And cAMP can regulate the expression of AKT to change glucose metabolism (Xie et al., 2014; Arcaro et al., 2018).

## Correlation Validation of lincRNA and mRNA by qPCR

During proximity genetic analysis, we predicted potential target genes that differentially express lincRNAs. To verify the reliability of the data analysis, we selected four pairs of lincRNA genes and protein-coding genes for real-time quantitative PCR. The lincRNAs were differentially expressed, while COL6A3, COL4A2, and C1S were highly expressed in SF, and four protein-coding genes were expressed in LDM. The experimental results showed that the $p$-values of the four pairs of genes were less than 0.05, the Pearson correlation coefficient of COL4A2 and MSTRG.5937 was 0.65, and the Pearson correlation coefficient of C1S and MSTRG.22831 was 0.733 (**Figure 6**). The correlation coefficient and $p$-values of the genes were calculated according to their expression level in the original data set, and their relationship was basically consistent with the basic trend of experimental verification.
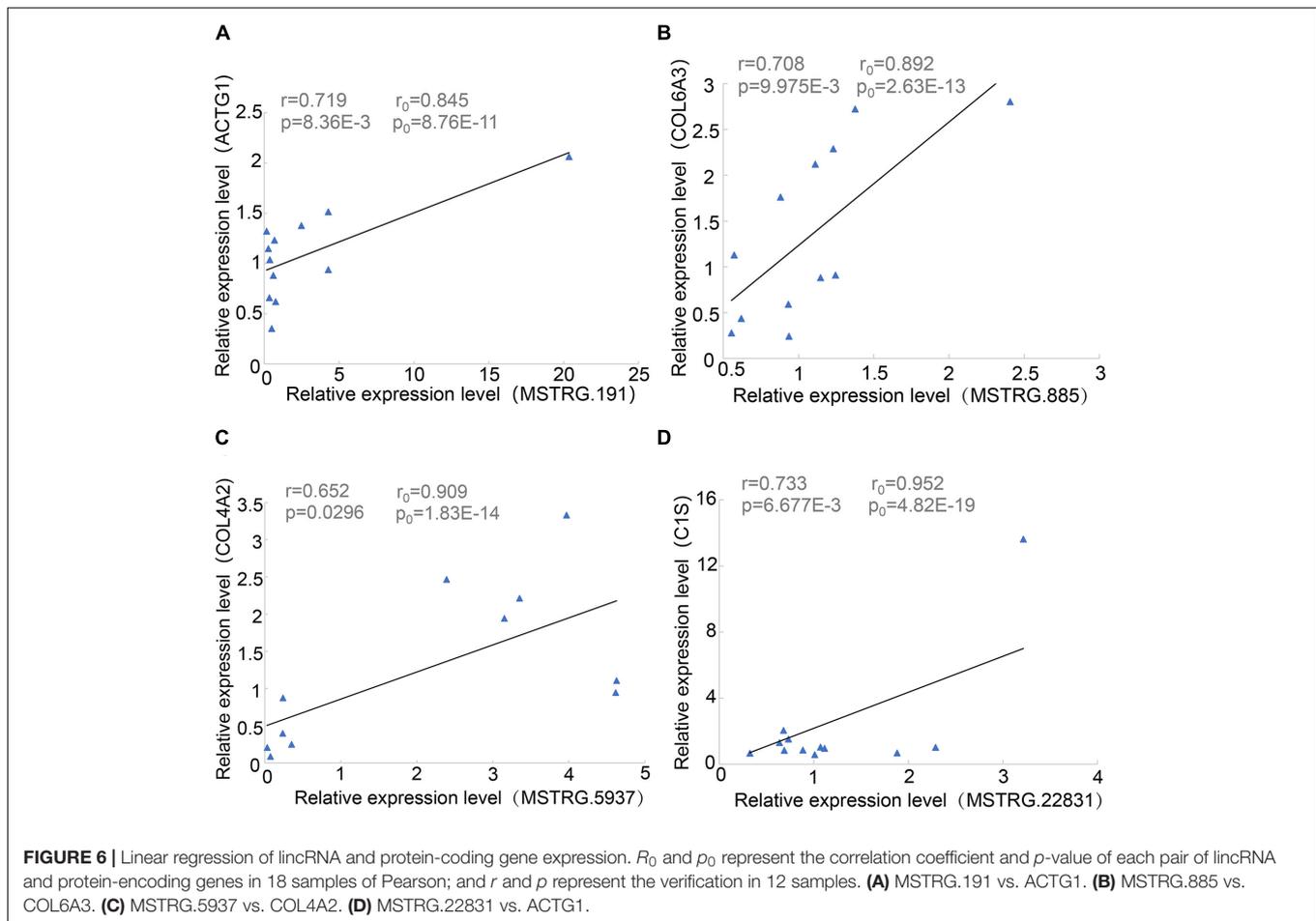
## DISCUSSION

In this study, we performed a comprehensive analysis of 36 RNA-seq libraries based on the designed pipeline and we identified 1078 lincRNAs in LDM and SF of weaned piglets. Compared with the known lincRNA database, we found 198 new putative lincRNAs. To detect the genomic characteristics of the identified lncRNA, we compared the known and novel lincRNAs with the porcine protein-coding gene and found that lincRNAs have shorter transcripts, fewer exons, longer exon lengths and lower expression levels (Koufariotis et al., 2015; Han et al., 2016; Yang M. et al., 2016). Previous studies have shown that the size of the pig genome is essentially the same as that of humans and mice; however, fewer lincRNAs have been identified in pigs than in the above two species, thereby indicating that a large number of lincRNAs in pigs have not been found (Kwok and Tay, 2017; Zhou et al., 2017). Therefore, the 198

new putative lincRNAs identified in this work not only enrich the lincRNA database but also extend the lincRNA annotation information of pigs.

Muscle and adipose, the two most important tissues in animals, are closely related to the growth and development of organisms (Kociucka et al., 2017; Liu et al., 2017). Muscle and fat development can affect the growth rate of livestock, meat production and meat quality, and also cause diseases such as obesity in human. Studying the developmental mechanisms of muscle and fat is important to obtain an in-depth understanding of the above issues. To study the relationship between the development of two tissues and the growth performance of piglets, we focused on differential expression analysis. By comparing DELs in weaned piglets between the two tissue types, we can effectively identify important lincRNA genes during the growth process. Analysis of the expression of these lincRNAs in muscle and fat promotes the understanding of their tissue specificity and potential function. We identified a total of 811 DELs, corresponding to 591 lincRNA genes between muscle and adipose tissue that may result in differences in growth performance. To predict the function of these lincRNAs, we examined their potential target genes in the vicinity of the 100 k region of DELs. However, the mechanism by which lincRNA regulates its potential target genes remains unclear and requires further investigation.

Enrichment analysis revealed that the processes and pathways enriched by potential target genes directly or indirectly affects muscle and fat development (Ayuso et al., 2015; Ibanez-Escriche et al., 2016). Several reports that lincRNA may have *cis*-regulatory relationships with these potential protein-coding genes have been published (Quattro et al., 2017). Some lincRNAs have be speculated to play an regulatory role in muscle development and fat deposition, which can change the expression pattern of target genes and promote the growth of weaned piglets. Three pathways are related to muscle and fat, namely glucose homeostasis, the AMPK signaling pathway and the insulin receptor signaling pathway. Previous studies have reported that glucose homeostasis can maintain the normal metabolism of cells, ensure the proliferation and differentiation of muscle cells and promote the formation of muscle fibers (Lauritsen et al., 2002). As a cellular energy receptor, AMPK responds to low levels of ATP and, when activated, positively regulates the signal transduction pathway that complements cellular ATP supply (Scheffler and Gerrard, 2016). The insulin receptor signaling pathway, once blocked, weakens the function of islet B cells and forms a large number of free fatty acids, leading to the dysregulation of fat and sugar metabolism in the body. Over time, the body forms fat deposits (Rajala and Anderson, 2010). These pathways and biological processes can affect the formation of muscle and adipose tissue to some extent. Although the mechanisms by which these lincRNAs regulate their potential target genes are unclear, these enriched biological processes suggest that some of the lincRNAs we identified may have functions that regulate muscle growth and fat metabolism. Our study provides new insights into the discovery and annotation of lincRNAs associated with growth

**FIGURE 6 |** Linear regression of lincRNA and protein-coding gene expression. $R_0$ and $p_0$ represent the correlation coefficient and $p$-value of each pair of lincRNA and protein-encoding genes in 18 samples of Pearson; and $r$ and $p$ represent the verification in 12 samples. **(A)** MSTRG.191 vs. ACTG1. **(B)** MSTRG.885 vs. COL6A3. **(C)** MSTRG.5937 vs. COL4A2. **(D)** MSTRG.22831 vs. ACTG1.

performance in weaned piglets, thereby providing an opportunity to further explore the molecular mechanisms underlying pig growth and development.

In this study, we also performed a simple QTL analysis to locate lincRNAs on QTLs. Since most QTLs are regulated by multiple genes, although studies have identified major genes by mapping QTL maps, accurate identification of target genes corresponding to a certain lincRNA remains challenging (Grisart et al., 2002). Moreover, most QTLs are fairly large, whereas lincRNAs are relatively short. A large number of lincRNAs may fall on the same QTL, causing difficulties in predicting the functions of specific lincRNAs. Based on the regulation of lincRNAs to their neighboring genes (<100 k), we may be able to predict the functions of lincRNAs more accurately. In this paper, we studied the average daily gain of QTLs and found that 93 lincRNAs fall into the QTL interval and are distributed on all chromosomes, but only 2 lincRNAs correspond to QTL intervals less than 100 kb. MSTRG.9938.1 is located on chromosome 15, almost coinciding with the QTL interval of 40 b, and is more likely to affect the average daily gain of piglets than MSTRG.4760.27, which corresponds to the 66 kb QTL interval on chromosome 12. Although our QTL study cannot directly determine the functions of lincRNAs, it can screen out a number of lincRNAs that may be associated with different production

traits. This result helps to enrich the genetic database associated with piglet growth performance and lays the foundation for improved piglet performance.

In summary, we identified and characterized several lincRNAs involved in muscle and fat development in weaned piglets. Tissue-specific analysis identified partially DELs, and adjacent gene analysis and functional enrichment revealed that many lincRNAs may act on potential target genes to participate in muscle growth and lipid deposition-related processes. QTL positional alignment also indicated that these genes can act on muscle or fat to further regulate average daily gains. Given that the functions and molecular mechanisms of lincRNAs are unclear and that lincRNAs annotated on pigs are not complete, this work provides a valuable resource for enriching the pig lincRNA database and helps improve the understanding on the potential functions of lincRNAs during the weaning transition phase for muscle and fat development.

## DATA AVAILABILITY

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA275632.

## AUTHOR CONTRIBUTIONS

CL conceived and designed the experiments and explained the data. LC analyzed the main content of the data with the help of GS, GC, CZ, and CF. JL performed the experiments with the help of ML. LC wrote the manuscript with the help of CL.

## FUNDING

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00196/full#supplementary-material

**TABLE S1 |** Summary of data from RNA-seq.

**TABLE S2 |** GTF of protein coding gene.

**TABLE S3 |** GTF of lincRNA gene.

**TABLE S4 |** FPKM of all transcripts.

**TABLE S5 |** FPKM of all lincRNA transcripts.

**TABLE S6 |** Count value of all genes.

**TABLE S7 |** FPKM of protein coding gene.

**TABLE S8 |** Count value of all differentially expressed genes.

**TABLE S9 |** Count value of all differentially expressed lincRNA genes.

## REFERENCES

Albuquerque, A., Neves, J. A., Redondeiro, M., Laranjo, M., Felix, M. R., Freitas, A., et al. (2017). Long term betaine supplementation regulates genes involved in lipid and cholesterol metabolism of two muscles from an obese pig breed. *Meat Sci.* 124, 25–33. doi: 10.1016/j.meatsci.2016.10.012

Arcaro, C. A., Assis, R. P., Zanon, N. M., Paula-Gomes, S., Navegantes, L., Kettelhut, I. C., et al. (2018). Involvement of cAMP/EPAC/Akt signaling in the antiproteolytic effects of pentoxifylline on skeletal muscles of diabetic rats. *J. Appl. Physiol.* 124, 704–716. doi: 10.1152/japplphysiol.00499.2017

Ayuso, M., Fernandez, A., Nunez, Y., Benitez, R., Isabel, B., Barragan, C., et al. (2015). Comparative analysis of muscle transcriptome between pig genotypes identifies genes and regulatory mechanisms associated to growth, fatness and metabolism. *PLoS One* 10:e0145162. doi: 10.1371/journal.pone.0145162

Benjamini, Y., and Cohen, R. (2017). Weighted false discovery rate controlling procedures for clinical trials. *Biostatistics* 18, 91–104. doi: 10.1093/biostatistics/kxw030

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* 125, 279–284. doi: 10.1016/S0166-4328(01)00297-2

Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799–816. doi: 10.1038/nature05874

Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927. doi: 10.1101/gad.17446611

De Quattro, C., Mica, E., Pe, M. E., and Bertolini, E. (2018). Brachypodium distachyon long noncoding RNAs: genome-wide identification and expression analysis. *Methods Mol. Biol.* 1667, 31–42. doi: 10.1007/978-1-4939-7278-4_3

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108. doi: 10.1038/nature11233

Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., et al. (2015). HMMER web server: 2015 update. *Nucleic Acids Res.* 43, W30–W38. doi: 10.1093/nar/gkv397

Fisar, Z., and Hroudova, J. (2016). Pig brain mitochondria as a biological model for study of mitochondrial respiration. *Folia Biol.* 62, 15–25.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., et al. (2002). Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Res.* 12, 222–231. doi: 10.1101/gr.224202

Han, S., Liang, Y., Li, Y., and Du, W. (2016). Long noncoding RNA identification: comparing machine learning based tools for long noncoding transcripts discrimination. *Biomed. Res. Int.* 2016:8496165. doi: 10.1155/2016/8496165

Hong, S. H., Kwon, J. T., Kim, J., Jeong, J., Kim, J., Lee, S., et al. (2018). Profiling of testis-specific long noncoding RNAs in mice. *BMC Genomics* 19:539. doi: 10.1186/s12864-018-4931-3

Huang, D. W., Sherman, B. T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., et al. (2009). Extracting biological meaning from large gene lists with DAVID. *Curr. Protoc. Bioinformatics* 27, 13.11.1–13.11.13. doi: 10.1002/0471250953.bi1311s27

Ibanez-Escriche, N., Magallon, E., Gonzalez, E., Tejeda, J. F., and Noguera, J. L. (2016). Genetic parameters and crossbreeding effects of fat deposition and fatty acid profiles in Iberian pig lines. *J. Anim. Sci.* 94, 28–37. doi: 10.2527/jas.2015-9433

Jones, C. K., Gabler, N. K., Main, R. G., and Patience, J. F. (2012). Characterizing growth and carcass composition differences in pigs with varying weaning weights and postweaning performance. *J. Anim. Sci.* 90, 4072–4080. doi: 10.2527/jas.2011-4793

Kociucka, B., Stachecka, J., Szydlowski, M., and Szczerbal, I. (2017). Rapid communication: the correlation between histone modifications and expression of key genes involved in accumulation of adipose tissue in the pig. *J. Anim. Sci.* 95, 4514–4519. doi: 10.2527/jas2017.2010

Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., Wei, L., et al. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 35, W345–W349. doi: 10.1093/nar/gkm391

Koufariotis, L. T., Chen, Y. P., Chamberlain, A., Vander, J. C., and Hayes, B. J. (2015). A catalogue of novel bovine long noncoding RNA across 18 tissues. *PLoS One* 10:e0141225. doi: 10.1371/journal.pone.0141225

Kwok, Z. H., and Tay, Y. (2017). Long noncoding RNAs: lincs between human health and disease. *Biochem. Soc. Trans.* 45, 805–812. doi: 10.1042/BST20160376

Lauritsen, T. L., Grunnet, N., Rasmussen, A., Secher, N. H., and Quistorff, B. (2002). The effect of hepatectomy on glucose homeostasis in pig and in man. *J. Hepatol.* 36, 99–104. doi: 10.1016/S0168-8278(01)00241-0

Leung, A., Trac, C., Jin, W., Lanting, L., Akbany, A., Saetrom, P., et al. (2013). Novel long noncoding RNAs are regulated by angiotensin II in vascular smooth muscle cells. *Circ. Res.* 113, 266–278. doi: 10.1161/CIRCRESAHA.112.300849

Li, D., Guo, L., Deng, B., Li, M., Yang, T., Yang, F., et al. (2018). Long noncoding RNA HR1 participates in the expression of SREBP1c through phosphorylation of the PDK1/AKT/FoxO1 pathway. *Mol. Med. Rep.* 18, 2850–2856. doi: 10.3892/mmr.2018.9278

Li, X., Wu, Z., Fu, X., and Han, W. (2013). Long noncoding RNAs: insights from biological features and functions to diseases. *Med. Res. Rev.* 33, 517–553. doi: 10.1002/med.21254

Li, X., Zhu, Y., Zhang, H., Ma, G., Wu, G., Xiang, A., et al. (2018). MicroRNA-106a-5p inhibited C2C12 myogenesis via targeting PIK3R1 and modulating the PI3K/AKT signaling. *Genes* 9:E333. doi: 10.3390/genes9070333

Liu, W., Cheng, C., Lin, Y., XuHan, X., and Lai, Z. (2018). Genome-wide identification and characterization of mRNAs and lncRNAs involved in cold stress in the wild banana (*Musa itinerans*). *PLoS One* 13:e0200002. doi: 10.1371/journal.pone.0200002

Liu, Y., Yang, X., Jing, X., He, X., Wang, L., Liu, Y., et al. (2017). Transcriptomics analysis on excellent meat quality traits of skeletal muscles of the chinese indigenous min pig compared with the large white breed. *Int. J. Mol. Sci.* 19:E21. doi: 10.3390/ijms19010021

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Lv, D., Xiong, X., Yang, H., Wang, M., He, Y., Liu, Y., et al. (2018). Effect of dietary soy oil, glucose, and glutamine on growth performance, amino acid profile, blood profile, immunity, and antioxidant capacity in weaned piglets. *Sci. China Life Sci.* 61, 1233–1242. doi: 10.1007/s11427-018-9301-y

Madeira, M. S., Lopes, P. A., Costa, P., Coelho, D., Alfaia, C. M., and Prates, J. (2017). Reduced protein diets increase intramuscular fat of psoas major, a red muscle, in lean and fatty pig genotypes. *Animal* 11, 2094–2102. doi: 10.1017/S1751731117000921

Mizutani, R., Wakamatsu, A., Tanaka, N., Yoshida, H., Tochigi, N., Suzuki, Y., et al. (2012). Identification and characterization of novel genotoxic stress-inducible nuclear long noncoding RNAs in mammalian cells. *PLoS One* 7:e34949. doi: 10.1371/journal.pone.0034949

Nakaya, H. I., Amaral, P. P., Louro, R., Lopes, A., Fachel, A. A., Moreira, Y. B., et al. (2007). Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. *Genome Biol.* 8:R43. doi: 10.1186/gb-2007-8-3-r43

Pauli, A., Valen, E., Lin, M. F., Garber, M., Vastenhouw, N. L., Levin, J. Z., et al. (2012). Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res.* 22, 577–591. doi: 10.1101/gr.133009.111

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, stringtie and ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095

Pilcher, C. M., Jones, C. K., Schroyen, M., Severin, A. J., Patience, J. F., Tuggle, C. K., et al. (2015). Transcript profiles in longissimus dorsi muscle and subcutaneous adipose tissue: a comparison of pigs with different postweaning growth rates. *J. Anim. Sci.* 93, 2134–2143. doi: 10.2527/jas.2014-8593

Pirooznia, M., Perkins, E. J., and Deng, Y. (2008). Batch blast extractor: an automated blastx parser application. *BMC Genomics* 9(Suppl. 2):S10. doi: 10.1186/1471-2164-9-S2-S10

Quattro, C., Enrico, P. M., and Bertolini, E. (2017). Long noncoding RNAs in the model species *Brachypodium distachyon*. *Sci. Rep.* 7:11252. doi: 10.1038/s41598-017-11206-z

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

Rajala, R. V., and Anderson, R. E. (2010). Rhodopsin-regulated insulin receptor signaling pathway in rod photoreceptor neurons. *Mol. Neurobiol.* 42, 39–47. doi: 10.1007/s12035-010-8130-8

Ran, M., Chen, B., Li, Z., Wu, M., Liu, X., He, C., et al. (2016). Systematic identification of long noncoding RNAs in immature and mature porcine testes. *Biol. Reprod.* 94:77. doi: 10.1095/biolreprod.115.136911

Rao, X., Huang, X., Zhou, Z., and Lin, X. (2013). An improvement of the 2ˆ(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. *Biostat. Bioinform. Biomath.* 3, 71–85.

Robich, M. P., Osipov, R. M., Nezafat, R., Feng, J., Clements, R. T., Bianchi, C., et al. (2010). Resveratrol improves myocardial perfusion in a swine model of hypercholesterolemia and chronic myocardial ischemia. *Circulation* 122, S142–S149. doi: 10.1161/CIRCULATIONAHA.109.920132

Scheffler, T. L., and Gerrard, D. E. (2016). Growth and development symposium: adenosine monophosphate-activated protein kinase and mitochondria in

Rendement Napole pig growth. *J. Anim. Sci.* 94, 3601–3612. doi: 10.2527/jas.2015-0055

Shin, J. H., Lee, S. H., Kim, Y. N., Kim, I. Y., Kim, Y. J., Kyeong, D. S., et al. (2016). AHNAK deficiency promotes browning and lipolysis in mice via increased responsiveness to beta-adrenergic signalling. *Sci. Rep.* 6:23426. doi: 10.1038/srep23426

Snoeck, V., Huyghebaert, N., Cox, E., Vermeire, A., Saunders, J., Remon, J. P., et al. (2004). Gastrointestinal transit time of nondisintegrating radio-opaque pellets in suckling and recently weaned piglets. *J. Control. Release* 94, 143–153. doi: 10.1016/j.jconrel.2003.09.015

Wang, K. C., and Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. *Mol. Cell* 43, 904–14. doi: 10.1016/j.molcel.2011.08.018

Xiao, H., Yuan, Z., Guo, D., Hou, B., Yin, C., Zhang, W., et al. (2015). Genome-wide identification of long noncoding RNA genes and their potential association with fecundity and virulence in rice brown planthopper, Nilaparvata lugens. *BMC Genomics* 16:749. doi: 10.1186/s12864-015-1953-y

Xie, X., Xu, X. M., Li, N., Zhang, Y. H., Zhao, Y., Ma, C. Y., et al. (2014). DMH1 increases glucose metabolism through activating Akt in L6 rat skeletal muscle cells. *PLoS One* 9:e107776. doi: 10.1371/journal.pone.0107776

Xing, Z., Chu, C., Chen, L., and Kong, X. (2016). The use of gene ontology terms and KEGG pathways for analysis and prediction of oncogenes. *Biochim. Biophys. Acta* 1860, 2725–2734. doi: 10.1016/j.bbagen.2016.01.012

Yang, H., Xu, X. L., Ma, H. M., and Jiang, J. (2016). Integrative analysis of transcriptomics and proteomics of skeletal muscles of the Chinese indigenous Shaziling pig compared with the Yorkshire breed. *BMC Genet.* 17:80. doi: 10.1186/s12863-016-0389-y

Yang, M., Tian, J., Guo, X., Yang, Y., Guan, R., Qiu, M., et al. (2016). Long noncoding RNA are aberrantly expressed in human papillary thyroid carcinoma. *Oncol. Lett.* 12, 544–552. doi: 10.3892/ol.2016.4653

Zhang, C., Luo, J. Q., Zheng, P., Yu, B., Huang, Z. Q., Mao, X. B., et al. (2015). Differential expression of lipid metabolism-related genes and myosin heavy chain isoform genes in pig muscle tissue leading to different meat quality. *Animal* 9, 1073–1080. doi: 10.1017/S1751731115000324

Zhao, P., Zheng, X., Feng, W., Wang, H., Kang, H., Ning, C., et al. (2018). Profiling long noncoding RNA of multi-tissue transcriptome enhances porcine noncoding genome annotation. *Epigenomics* 10, 301–320. doi: 10.2217/epi-2017-0149

Zhao, W., Mu, Y., Ma, L., Wang, C., Tang, Z., Yang, S., et al. (2015). Systematic identification and characterization of long intergenic non-coding RNAs in fetal porcine skeletal muscle development. *Sci. Rep.* 5:8957. doi: 10.1038/srep08957

Zhou, J., Zhang, S., Wang, H., and Sun, H. (2017). LncFunNet: an integrated computational framework for identification of functional long noncoding RNAs in mouse skeletal muscle cells. *Nucleic Acids Res.* 45:e108. doi: 10.1093/nar/gkx232

Zhu, Y. L., Chen, T., Xiong, J. L., Wu, D., Xi, Q. Y., Luo, J. Y., et al. (2018). miR-146b inhibits glucose consumption by targeting IRS1 gene in porcine primary adipocytes. *Int. J. Mol. Sci.* 19:E783. doi: 10.3390/ijms19030783

Zou, C., Li, J., Luo, W., Li, L., Hu, A., Fu, Y., et al. (2017a). Transcriptome analysis reveals long intergenic non-coding RNAs involved in skeletal muscle growth and development in pig. *Sci. Rep.* 7:8704. doi: 10.1038/s41598-017-07998-9

Zou, C., Li, S., Deng, L., Guan, Y., Chen, D., Yuan, X., et al. (2017b). Transcriptome analysis reveals long intergenic noncoding RNAs contributed to growth and meat quality differences between yorkshire and wannanhua pig. *Genes* 8:E203. doi: 10.3390/genes8080203