Identification of porcine imprinted genes involved in skeletal muscle development by high-throughput sequencing

Xinhua Hou

Institute of Animal Science, CAAS <https://orcid.org/0000-0001-9485-7766>

Zishuai Wang

Chinese Academy of Agricultural Sciences Agricultural Genomes Institute at Shenzhen

Ligang Wang

Chinese Academy of Agricultural Sciences Institute of Animal Science

Fuping Zhao

Chinese Academy of Agricultural Sciences Institute of Animal Science

Xin Liu

Chinese Academy of Agricultural Sciences Institute of Animal Science

Hongmei Gao

Chinese Academy of Agricultural Sciences Institute of Animal Science

Lijun Shi

Chinese Academy of Agricultural Sciences Institute of Animal Science

Hua Yan

Chinese Academy of Agricultural Sciences Institute of Animal Science

Longchao Zhang

Chinese Academy of Agricultural Sciences Institute of Animal Science

Lixian Wang (✉ iaswlx@263.net)

Chinese Academy of Agricultural Sciences <https://orcid.org/0000-0002-1852-7308>

Research article

Keywords: pig, imprinted gene, skeletal muscle, RNA-seq, DNA-seq

Posted Date: September 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-77481/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Imprinted genes—exhibiting parent-specific transcription—play essential roles in the process of mammalian development and growth. To further understand the imprinted genes involved in skeletal muscle development, DNA-seq and RNA-seq were used to explore the characteristics of imprinted genes involved in skeletal muscle development from porcine reciprocal crosses.

Results

A total of 211 paternally and 417 maternally imprinted genes were obtained in this study. Of the candidate imprinted genes, 50 paternally and 112 maternally imprinted genes possessed cytosine and guanine dinucleotide (CpG) islands in their promoters that may have regulated the imprinted gene expression. Imprint-related motifs were predicted and PBX1 motifs may mediate the expression of imprinted genes in the process of skeletal-muscle development. Functional analysis showed that a maternally imprinted gene of EPHB1 was involved in skeletal muscle cell proliferation. Imprinted genes involved in the biological processes of cell proliferation, differentiation and fusion were also analyzed. More maternally imprinted genes, including ADRA1D, E2F1, FBXO40, GHRH and WNT5A, were involved in skeletal muscle development than paternally imprinted genes.

Conclusions

Skeletal muscle development is crucial for meat production. This study identified 628 porcine imprinted genes in skeletal muscle and revealed their functional role in skeletal muscle development. Our findings should further assist with the potential use of imprinted genes in pig breeding.

Background

Genomic imprinting is a form of allelic imbalance in gene expression in a parent-of-origin-dependent manner [1]. The expression of imprinted genes is specific to tissue and development stage [2]. Many imprinted genes are expressed during embryonic periods and regulate embryonic growth and placental functions. Studies of mice have revealed that after high level of expressing during the fetal and newborn periods, many imprinted genes show decreased expression levels within weeks of birth [3, 4]. Skeletal muscle fibers are formed by the fusion and differentiation of myoblasts and mainly generated during the fetal and newborn periods [5, 6]. In a previous study on mice, seven imprinted genes—including Bex1, H19, p57Kip2, Igf2, Peg1/Mest, Peg3/Pw1 and Zac1—were found to be important to muscle differentiation [7], which suggests that imprinted genes may play vital roles in skeletal muscle development. Some imprinted genes have also been confirmed in porcine skeletal muscle. MAGEL2 is paternally expressed in skeletal muscle from 65-day-old pig fetuses, which may affect carcass traits and be related to porcine

development [8]. The DLX5 gene is maternally expressed in skeletal muscle, and the C-to-T mutation (polymorphism) in exon 1 has been shown to have a highly significant association with fat meat percentage (FMP), carcass length (CL), bone percentage (BP), average backfat thickness (ABT), thorax–waist fat thickness (TFT) and buttock fat thickness (BFT) [9]. Several porcine genes, including NAP1L5, PEG3, CDKN1C, NDN, MAGEL2, MEST, PLAGL1, PEG10, NECD, SNRPN, PEG3, NAP1L5, NNAT and DIRAS3, are paternally expressed in skeletal muscle [10–15]. However, these studies only concentrate on a limited number of porcine imprinted genes that have been reported in humans or mice. To obtain a comprehensive view of the imprinted genes related to porcine skeletal muscle development, imprinted genes were detected in the skeletal muscle of postnatal 0-day-old piglets from reciprocal crosses using both RNA-seq and DNA-seq in this study.

Methods

Animal and tissue collection

Yorkshire (Y) and Min (M) pigs were used to construct a pair of reciprocal crosses ($Y \times M$ and $M \times Y$). For each reciprocal F1 group, four newborn piglets (including two male and two female) were anesthetized with a propofol injection (10 mg/kg) and then euthanized by exsanguination. Longissimus dorsi muscles of piglets were collected and frozen in liquid nitrogen. Ear samples from four F0 pigs were collected and kept in 70% ethyl alcohol. All pigs used in this study were obtained from the experiment pig farm of Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS). All animals were treated following the guidelines established by the Ministry of Science and Technology. The experimental procedures were approved by the Science Research Department of the Institute of Animal Sciences, CAAS.

Genome sequencing

DNA from the ear samples of the F0 pigs was extracted using a DNeasy Blood & Tissue Kit (Qiagen) according to the manufacture's protocol. The quality of DNA was evaluated using an Agilent 2100 bioanalyzer (Agilent). After ultrasonication with a peak of 350 bp, the DNA was end-repaired using NEBNext Ultra II End Prep Enzyme Mix (New England Biolabs). The DNA fragments were ligated with NEBNext Adaptor (New England Biolabs) and amplified by PCR using NEBNext Ultra II Q5 Master Mix (New England Biolabs). The purified PCR fragments were sequenced on an Illumina HiSeq X-ten platform.

Transcriptome sequencing

Total RNA was extracted from the longissimus dorsi muscles of the F1 reciprocal pigs with TRIzol reagent (Invitrogen). The quality of RNA was evaluated using an Agilent 2100 bioanalyzer (Agilent). After reverse transcription of first-strand cDNA, second-strand cDNA was synthesized using NEBNext Second Strand Synthesis Enzyme Mix (New England Biolabs). The cDNA fragments were purified and ligated with NEBNext Adaptor (New England Biolabs). PCR was performed to amplify the cDNA fragments. After

assessment of the quality of the libraries by an Agilent 2100 Bioanalyzer, RNA-seq was performed on an Illumina HiSeq X-ten platform.

Reads mapping and SNVs identification

Adaptor sequences were filtered, and low-quality reads were removed using the FASTX-Toolkit (v 0.0.14). Genomes and annotation files of *Sus scrofa* (Sscrofa11.1) were downloaded from the ENSEMBL database. The genome sequencing reads from the F0 pigs were mapped to reference genomes using BWA-MEM (0.7.1) with default parameters; RNA-seq reads from the F1 pigs were mapped to reference genomes using HISAT2 (v2.0.1) with default parameters. The SAM files were sorted and converted to BAM files by SAMtools (v1.3.1). After reads alignment, the single-nucleotide variations (SNVs) were called by SAMtools (v1.3.1) for each individual. Then, each SNV was annotated using snpEFF (v4.2). Any SNV with a read-depth of less than 10 and a quality score of less than 25 was discarded.

Identification of candidate imprinted genes

In each cross, the SNVs—which were different between two F0 parent pigs—were chosen for further analysis, as the expression of two alleles in offspring could be defined. SNV phasing of F1 pigs from RNA-seq data was performed based on the lists of the parent SNVs. Specifically, for each F1 individual, if all of the phased SNVs as annotated by their parents' SNVs at same position were from a single parent, this SNV could be defined as monoallelically expressed. Otherwise, it could be defined as a biallelically expressed SNV. For imprinted gene identification, the SNVs located in exon–intron boundaries, UTR region and X/Y chromosomes of F1 piglets were removed. Then, only the transcripts that possessed monoallelically expressed SNVs were considered to be imprinted genes. In each cross, the intersection of four offspring's paternally expressed genes or maternally expressed genes were kept. The final imprinted genes were obtained by finding the intersection of paternally or maternally expressed genes between two crosses.

Prediction of CpG islands and motifs in the upstream region of imprinted genes

The upstream sequences (–2000 to +200 bp) of the imprinted genes were selected to predict the CpG islands using the Cpgplot software (https://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/) with the following parameters: observed/expected CpG ratio–0.6; minimum length of an island–200 nt; minimum G+C content–50% [16]. The imprint-related motifs were predicted by DREME (5.0.5) with default parameters (<http://meme-suite.org/tools/dreme>). All imprinted genes were considered to be members of the positive group. The same number of biallelically expressed genes selected randomly were considered to be members of the negative group. The motifs were discovered according to the upstream sequence of each gene. The statistical significance of each motif found in the positive group compared with the negative group was determined by Fisher's exact test. The significance threshold for motifs was $P < 0.05$.

GO enrichment analysis

A gene ontology (GO) enrichment analysis using the R package “GOstats” was performed on the candidate paternally and maternally imprinted genes [17]. The “ontology” parameters of biological process (BP), cellular component (CC) and molecular function (MF) were selected, and a “pvalueCutoff” parameter of 0.05 was chosen. The interactions of imprinted genes were analyzed using the STRING database. The interaction network was constructed using the Cytoscape software (3.7.1).

Results

Identification of imprinted genes by high-throughput sequencing

Reciprocal crosses of Yorkshire and Min were constructed to characterize the imprinted genes related to skeletal muscle development (Fig.1a). DNA-seq for F0 parents and RNA-seq for F1 progenies were performed to genotype the SNV and detect allelic bias of transcripts. There were 275,027,760–339,794,920 clean reads from DNA-seq and 174,931,928–208,251,498 clean reads from RNA-seq. The Q30 was about 90–94% for both DNA-seq and RNA-seq. The GC content was 40–44% for DNA-seq and 51–53% for RNA-seq (Additional file 1: Table S1).

DNA-seq and RNA-seq clean reads were aligned to the Sscrofa11.1 genome, and the aligned reads were used to identify the SNV with SAMtools. The observed SNVs—heterozygous in the F1 piglets, but homozygous in the F0 parents—were phased to determine the allele-specific expression (Fig.1b). A list of genes as the phased SNVs that exhibited a complete paternal or maternal expression bias in a certain transcript of an F1 individual were considered to be informative for further analysis. If all of the SNVs located in a transcript fragment were from the same parent-of-origin allele, the gene containing this fragment was selected. Then, the overlaps of the above genes among the four piglets in each cross were preserved. As shown in Fig.2a, a total of 6038 genes with paternal SNVs and 3944 genes with maternal SNVs were obtained in Y × M and 4665 genes with paternal SNVs and 6730 genes with maternal SNVs were obtained in M × Y.

All of the SNVs in an imprinted gene should originate from the single parent simultaneously. Therefore, the genes that processed both the paternal and maternal SNVs from the previous step were removed. Only the SNVs that existed in the exon were kept; those located in the X/Y chromosome, untranslated region (UTR), intron, and splicing sites between exon and intron were ignored. About 3,137 genes with paternal SNVs and 1,043 genes with maternal SNVs were obtained in Y × M, while 1,427 genes with paternal SNVs and 3,492 genes with maternal SNVs were obtained in M × Y (Fig.2b).

Finally, the intersection of genes with paternal SNVs (or maternal SNVs) between Y × M and M × Y were reserved to eliminate monoallelic expression caused by the breed type. After this filter step, we obtained 211 paternally imprinted genes and 417 maternally imprinted genes (Fig.2c and Additional file 2: Table S2).

Regulation and characteristics of imprinted genes

Mutation of the DNMT gene family causes genome-wide deficiencies in DNA methylation and alteration in imprinted gene expression, suggesting that DNA methylation plays an essential role in the regulation of imprinted gene expression [18]. We analyzed the CpG islands in the promoter regions of candidate imprinted genes to predict the imprinting control region (ICR) of imprinted genes. A total of 162 imprinted genes (50 paternally and 112 maternally imprinted genes) possessed CpG islands in the promoters (Fig.3 and Additional file 3: Table S3).

Two or more imprinted genes are often found in clusters on a chromosome over a region spanning nearly 1 MB or more [19]. Some of the candidate imprinted genes in this study existed in clusters (Fig.4). A total of 78 clusters—including 216 imprinted genes—were identified in this study. However, none of the imprinted genes matched known clusters (Additional file 4: Table S4).

ZFP57—recruited by the methylated TGCCGC motif—is a regulator of genomic imprinting, which suggests that the motif has an important role in the regulation of imprinted gene expression by DNA methylation [20]. Imprint-related motifs were predicted within the upstream sequences of the candidate imprinted genes in this study. About 38 imprinted-gene-specific motifs were obtained (Fig.5).

Functional analysis of imprinted genes for skeletal muscle development

To further understand the function of imprinted genes, a GO enrichment analysis of paternally or maternally imprinted genes was performed (Fig.6). The significant enrichment biological processes were homophilic cell adhesion via plasma membrane adhesion molecules, ion transport and regulation of postsynaptic cytosolic calcium ion concentration for paternally imprinted genes and immune response in mucosal-associated lymphoid tissue, cell–cell recognition and DNA methylation or demethylation for maternally imprinted genes (Additional file 5: Table S5 and Additional file 6: Table S6). However, only a maternally imprinted gene of EPHB1 was found to be involved in the biological process of negative regulation of skeletal muscle cell proliferation, which was related to skeletal muscle development. Myoblasts undergo proliferation, differentiation and fusion processes to form multinucleated myofibers. Therefore, the imprinted genes related to the biological processes of cell proliferation, differentiation and fusion were selected to illustrate their potential regulatory role in skeletal muscle development. As shown in Fig.7, 41 maternally imprinted genes and 25 paternally imprinted genes participated in the above three biological processes, some of which (e.g., E2F1, FBXO40, and WNT5A) are well known to play a role in skeletal muscle development.

Discussion

RNA-seq has been used to identify novel imprinted genes in neonatal mouse brain from the AKR × PWD reciprocal crosses [21]. The AKR allele proportion was defined as $p1$ or $p2$ in the PWD × AKR or AKR × PWD cross, respectively. The paternally imprinted genes were judged with a criteria of $p1 \geq 0.5$ and $p2 \leq 0.5$; maternally imprinted genes with $p1 \leq 0.5$ and $p2 \geq 0.5$. In this study, Yorkshire and Min pigs were used to construct reciprocal crosses for imprinted gene identification. Unlike in inbred animal models such as mice, individual differences among pigs are relatively large. Moreover, the porcine genome annotation

information is incomplete compared with that of other animal models. Therefore, we combined DNA-seq of F0 parents and RNA-seq of F1 reciprocal crosses to accurately trace the parent-of-origin monoallelically expressed SNVs and analyze imprinted genes. Gender effects and related genes were eliminated by using both male and female F1 individuals in each cross for sequencing and analysis and removing the transcripts from the X/Y chromosome during imprinted gene identification. The SNVs in the UTR that were removed as polymorphisms in the 5' UTR may have influenced the expression by variable interaction with transcription factors; those in the 3' UTR would affect mRNA stability and translation by binding with miRNA. We used four replications in each cross for imprinted gene analysis to distinguish between autosomal random monoallelic expression and parent-of-origin monoallelic expression. In each cross, the breed-specific genes could not be distinguished from parent-of-origin imprinted genes, which would be solved by keeping the intersection of parent-of-origin imprinted genes from both reciprocal crosses. Finally, a total of 211 paternally imprinted genes and 417 maternally imprinted genes were obtained. Wu et al. used DNA-seq of parental genomes and RNA-seq of offspring to identify porcine brain-imprinted genes [22]. However, they obtained fewer imprinted genes than we did, which may be due to fewer SNVs used in their study.

IGF2 is a well-known imprinted gene transcribed from the paternal allele in humans and rodents [23, 24]. IGF2 is also reported to have biallelic expression in chicken embryo [25]. We also found IGF2 to be biallelically expressed in the F1 piglets. SLC22A2, FBXO40, CTNNA3 and IDO1 are known imprinted genes that were also identified in our reciprocally crossed piglets. SLC22A2 is from the IGF2R imprinted gene cluster and is maternally expressed in the placenta but not imprinted in adult mouse kidney [26]. FBXO40 is imprinted and maternally expressed in the mouse placenta but not imprinted in E17.5 in the mouse brain [27]. CTNNA3 is subject to genomic imprinting with maternal allele expression in human placental tissue [28]. IDO1 is maternally expressed in mouse E9.5 placentae as the paternal allele is hypermethylated [29]. According to our results, SLC22A2 and FBXO40 are maternally imprinted genes, while CTNNA3 and IDO1 are paternally imprinted genes. The conflicting manner of expression of SLC22A2 and FBXO40 to previous studies may be caused by tissue and temporal specificity.

DNA methylation is a conserved epigenetic-silencing mechanism that functions by adding a methyl group to cytosine residues in CpG islands. The region of differential methylation (DMR) is also found at most imprinted locations and participates in the regulation of imprinted gene expression [30]. The CpG islands in the promoters of the imprinted genes in this study were predicted to elucidate the potential regulation of genomic imprinting. A subset of imprinted genes, including TDRD9, DPPA5, GFI1, MTTP, CDX1, WNT2, ITGA2, TCF7, MAML2 and WNT5A, possessed CpG islands in their promoter, which showed hypermethylation or hypomethylation of CpG islands from previous studies. The TDRD9 CpG island is hypomethylated and has a high expression level in non-small cell carcinomas, whereas hypermethylation will lead to nonexpression of TDRD9 [31]. The silencing of DPPA5 is associated with increased methylation in its CpG island promoter in differentiated embryonic carcinoma cells [32]. Hypomethylation at one CpG locus in the GFI1 promoter on the maternal side of the placenta is due to maternal smoking [33]. The promoter of MTTP is rich in CpG islands, and its DNA hypermethylation causes a decrease in mRNA expression levels in the livers of rats treated with a high-fat diet [34]. There is a typical CpG island

located near the transcription start site of the CDX1 gene. Hypermethylation of the CDX1 promoter reduces this gene's expression in colorectal cell lines [35]. High levels of WNT2 promoter methylation exist in placental tissue and are significantly associated with reduced WNT2 expression [36]. The ITGA2 gene contains CpG islands in the promoter, the hypomethylation of which is found in prostate tumors compared with that of normal prostate tissue [37]. The upstream region of the TCF7 promoter shows strongly reduced methylation in blood samples from patients with Down's syndrome [38]. The DNA methylation level at CpG sites of MAML2 promoters shows hypomethylation in cancer cells versus normal cells [39]. The WNT5A promoter possesses a CpG island, the methylation of which could make it silenced in several cell lines [40]. The expression of porcine imprinted genes with predicted CpG islands located in the promoter may be mediated by DNA methylation.

Tandem-repeat motifs have been identified in imprinted regions, which are involved in the regulation of the epigenetic status of imprinted genes through differential methylation or their insulator role [41]. Previous studies have shown that ZFP57—a member of KRAB zinc-finger protein family—can recognize a hexameric motif and maintain genomic imprinting in embryonic stem (ES) cells [42]. CCCTC-binding factor (CTCF) is a specific zinc-finger protein that can bind to the CAGCCC motif and mediate the expression of imprinted genes [43, 44]. In this study, the imprint-specific motifs in the promoter regions were predicted. Two motifs—m18nc_AWTCAAA and m19nc_RTTGWT—were enriched in the imprinted gene promoters and could be recognized by PBX1. PBX1 can bind to the ATCAATCAA motif; deletion of PBX1 will lead to deregulation of imprinted genes in ES cells [45]. The PBX1 motif may be a candidate element in the regulation of the expression of imprinted genes during skeletal muscle development.

To elucidate the roles of imprinted genes during skeletal muscle development, a GO enrichment analysis was performed. The results showed that EPHB1—expressed from the paternal allele—was involved in the regulation of skeletal muscle cell proliferation and differentiation. EPHB1 is expressed in postnatal satellite cells and has a role in promoting self-renewal and reducing the activation and differentiation of satellite cells [46]. The imprinted genes involved in the biological processes of cell proliferation, differentiation and fusion were also analyzed to disclose the molecular mechanisms in the regulation of skeletal muscle development. The interaction network suggested that more maternally imprinted genes than paternally imprinted genes took part in the skeletal muscle development. Some of the above mentioned maternally imprinted genes—including ADRA1D, E2F1, FBXO40, GHRH and WNT5A—have been implicated in skeletal muscle development in previous studies. Silencing of ADRA1D in myoblasts has a negative impact on myotube formation [47]. Constitutive expression of E2F1 is critical to the cell cycle in myoblasts as it activates the expression of E2F1-dependent cell cycle genes [48, 49]. FBXO40 is considered to be a muscle-specific gene and influences the growth rates and muscle mass of young mice by regulating IRS1 [50]. GHRH has survival and antiapoptotic effects in myotubes by inhibiting TNF- α -induced apoptosis and promotes muscle regeneration by inducing the expression of myogenin in TNF- α -treated myotubes [51]. WNT5A—belonging to the Wnt family—can activate both MYF5 and MYOD [52]. Overexpression of WNT5A can also promote satellite-cell proliferation [53].

Conclusions

In this study, a total of 628 imprinted genes were identified in porcine skeletal muscles. The CpG islands and motifs in the promoters of imprinted genes that may be involved in the regulation of imprinted gene expression were predicted. Function analysis showed that more maternally imprinted genes were involved in the regulation of skeletal muscle development than paternally imprinted genes. This study may facilitate novel insights into the potential use of imprinted genes in pig breeding.

Abbreviations

FMP: fat meat percentage; CL:carcass length; BP:bone percentage; ABT:average backfat thickness; TFT:thorax-waist fat thickness; BFT:buttock fat thickness; SNV:single nucleotide variations; UTR:untranslated region; ICR:imprinting control region; DMR:region of differential methylation

Declarations

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request

Ethics approval and consent to participate

The experimental procedures were approved by the Science Research Department of the Institute of Animal Sciences, Chinese Academy of Agricultural Sciences. There was no use of human participants, data or tissues.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by National Natural Science Foundation (31872337), National Swine Industry Technology System (CARS-35), and Agricultural Science and Technology Innovation Project (ASTIP-IAS02). The funding bodies had no role in the design of the study and collection, analysis and interpretation of data and writing the manuscript.

Authors' contributions

LW and LZ conceived and designed the experiment. XH and ZW analyzed data and wrote the manuscript. LW, FZ, XL, HG, LS, and HY collected tissue sample. All authors read and approved the final manuscript.

Acknowledgements

We are grateful for the dedication and hard work of the members of our laboratory.

References

1. Barlow DP: **Genomic imprinting: a mammalian epigenetic discovery model.** *Annual review of genetics* 2011, **45**:379-403.
2. Prickett AR, Oakey RJ: **A survey of tissue-specific genomic imprinting in mammals.** *Molecular Genetics and Genomics* 2012, **287**(8):621-630.
3. Berg JS, Lin KK, Sonnet C, Boles NC, Weksberg DC, Nguyen H, Holt LJ, Rickwood D, Daly RJ, Goodell MA: **Imprinted genes that regulate early mammalian growth are coexpressed in somatic stem cells.** *PloS one* 2011, **6**(10):e26410.
4. Lui JC, Finkelstein GP, Barnes KM, Baron J: **An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs.** *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 2008, **295**(1):R189-R196.
5. Lefaucheur L, Edom F, Ecolan P, Butler-Browne G: **Pattern of muscle fiber type formation in the pig.** *Developmental Dynamics* 1995, **203**(1):27-41.
6. Mascarello F, Stecchini M, Rowleron A, Ballocci E: **Tertiary myotubes in postnatal growing pig muscle detected by their myosin isoform composition.** *Journal of animal science* 1992, **70**(6):1806-1813.
7. Yan Z, Choi S, Liu X, Zhang M, Schageman JJ, Lee SY, Hart R, Lin L, Thurmond FA, Williams RS: **Highly coordinated gene regulation in mouse skeletal muscle regeneration.** *Journal of Biological Chemistry* 2003, **278**(10):8826-8836.
8. Guo L, Qiao M, Wang C, Zheng R, Xiong Y-Z, Deng C-Y: **Imprinting analysis of porcine MAGEL2 gene in two fetal stages and association analysis with carcass traits.** *Molecular biology reports* 2012, **39**(1):147-155.
9. Cheng H, Zhang F, Jiang C, Li F, Xiong Y, Deng C: **Isolation and imprinting analysis of the porcine DLX5 gene and its association with carcass traits.** *Animal genetics* 2008, **39**(4):395-399.
10. Jiang C, Li S, Deng C: **Assessment of genomic imprinting of PPP1R9A, NAP1L5 and PEG3 in pigs.** *Russian Journal of Genetics* 2011, **47**(4):471-476.
11. Li S, Li J, Tian J, Dong R, Wei J, Qiu X, Jiang C: **Characterization, tissue expression, and imprinting analysis of the porcine CDKN1C and NAP1L4 genes.** *Journal of Biomedicine and Biotechnology*

2012, 2012.

12. Zhang F, Cheng H, Jiang C, Deng C, Xiong Y, Li F, Lei M: **Imprinted status of pleomorphic adenoma gene-like I and paternal expression gene 10 genes in pigs.** *Journal of animal science* 2007, **85**(4):886-890.
13. Wang M, Zhang X, Kang L, Jiang C, Jiang Y: **Molecular characterization of porcine NECD, SNRPN and UBE3A genes and imprinting status in the skeletal muscle of neonate pigs.** *Molecular biology reports* 2012, **39**(10):9415-9422.
14. Zhang F-W, Deng C-Y, He H-J, Gu N, Han Z-B, Chen Y, Wu Q: **Molecular cloning, mRNA expression and imprinting status of PEG3, NAP1L5 and PPP1R9A genes in pig.** *Genes & genetic systems* 2011, **86**(1):47-52.
15. Cheng H-C, Zhang F-W, Deng C-Y, Jiang C-D, Xiong Y-Z, Li F-E, Lei M-G: **NNAT and DIRAS3 genes are paternally expressed in pigs.** *Genetics Selection Evolution* 2007, **39**(5):1-9.
16. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey AR, Potter SC, Finn RD: **The EMBL-EBI search and sequence analysis tools APIs in 2019.** *Nucleic acids research* 2019, **47**(W1):W636-W641.
17. Falcon S, Gentleman R: **Using GOstats to test gene lists for GO term association.** *Bioinformatics* 2007, **23**(2):257-258.
18. Hirasawa R, Chiba H, Kaneda M, Tajima S, Li E, Jaenisch R, Sasaki H: **Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development.** *Genes & development* 2008, **22**(12):1607-1616.
19. Verona RI, Mann MR, Bartolomei MS: **Genomic imprinting: intricacies of epigenetic regulation in clusters.** *Annual review of cell and developmental biology* 2003, **19**(1):237-259.
20. Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, Baglivo I, Pedone PV, Grimaldi G, Riccio A: **In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions.** *Molecular cell* 2011, **44**(3):361-372.
21. Wang X, Sun Q, McGrath SD, Mardis ER, Soloway PD, Clark AG: **Transcriptome-wide identification of novel imprinted genes in neonatal mouse brain.** *PloS one* 2008, **3**(12):e3839.
22. Wu Y-Q, Zhao H, Li Y-J, Khederzadeh S, Wei H-J, Zhou Z-Y, Zhang Y-P: **Genome-wide identification of imprinted genes in pigs and their different imprinting status compared with other mammals.** *Zoological research* 2020:1-5.
23. DeChiara TM, Robertson EJ, Efstratiadis A: **Parental imprinting of the mouse insulin-like growth factor II gene.** *Cell* 1991, **64**(4):849-859.
24. Giannoukakis N, Deal C, Paquette J, Goodyer CG, Polychronakos C: **Parental genomic imprinting of the human IGF2 gene.** *Nature genetics* 1993, **4**(1):98-101.
25. O'Neill MJ, Ingram RS, Vrana PB, Tilghman SM: **Allelic expression of IGF2 in marsupials and birds.** *Development genes and evolution* 2000, **210**(1):18-20.

26. Zwart R, Sleutels F, Wutz A, Schinkel AH, Barlow DP: **Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes.** *Genes & development* 2001, **15**(18):2361-2366.
27. Brideau CM, Eilertson KE, Hagarman JA, Bustamante CD, Soloway PD: **Successful computational prediction of novel imprinted genes from epigenomic features.** *Molecular and cellular biology* 2010, **30**(13):3357-3370.
28. van Dijk M, Mulders J, Könst A, Janssens B, Van Roy F, Blankenstein M, Oudejans C: **Differential downregulation of α T-catenin expression in placenta: trophoblast cell type-dependent imprinting of the CTNNA3 gene.** *Gene expression patterns* 2004, **5**(1):61-65.
29. Spinelli P, Latchney SE, Reed JM, Fields A, Baier BS, Lu X, McCall MN, Murphy SP, Mak W, Susiarjo M: **Identification of the novel Ido1 imprinted locus and its potential epigenetic role in pregnancy loss.** *Human molecular genetics* 2019, **28**(4):662-674.
30. Constância M, Pickard B, Kelsey G, Reik W: **Imprinting mechanisms.** *Genome research* 1998, **8**(9):881-900.
31. Guijo M, Ceballos-Chávez M, Gómez-Marín E, Basurto-Cayuela L, Reyes JC: **Expression of TDRD9 in a subset of lung carcinomas by CpG island hypomethylation protects from DNA damage.** *Oncotarget* 2018, **9**(11):9618.
32. Dansranjavin T, Krehl S, Mueller T, Mueller LP, Schmoll H-J, Dammann RH: **The role of promoter CpG methylation in the epigenetic control of stem cell related genes during differentiation.** *Cell cycle* 2009, **8**(6):916-924.
33. van Otterdijk SD, Binder AM, Michels KB: **Locus-specific DNA methylation in the placenta is associated with levels of pro-inflammatory proteins in cord blood and they are both independently affected by maternal smoking during pregnancy.** *Epigenetics* 2017, **12**(10):875-885.
34. Chang X, Yan H, Fei J, Jiang M, Zhu H, Lu D, Gao X: **Berberine reduces methylation of the MTTP promoter and alleviates fatty liver induced by a high-fat diet in rats.** *Journal of lipid research* 2010, **51**(9):2504-2515.
35. Suh ER, Ha CS, Rankin EB, Toyota M, Traber PG: **DNA methylation down-regulates CDX1 gene expression in colorectal cancer cell lines.** *Journal of Biological Chemistry* 2002, **277**(39):35795-35800.
36. Ferreira JC, Choufani S, Grafodatskaya D, Butcher DT, Zhao C, Chitayat D, Shuman C, Kingdom J, Keating S, Weksberg R: **WNT2 promoter methylation in human placenta is associated with low birthweight percentile in the neonate.** *Epigenetics* 2011, **6**(4):440-449.
37. Chin SP, Marthick JR, West AC, Short AK, Chuckowree J, Polanowski AM, Thomson RJ, Holloway AF, Dickinson JL: **Regulation of the ITGA2 gene by epigenetic mechanisms in prostate cancer.** *The Prostate* 2015, **75**(7):723-734.
38. Kerkel K, Schupf N, Hatta K, Pang D, Salas M, Kratz A, Minden M, Murty V, Zigman WB, Mayeux RP: **Altered DNA methylation in leukocytes with trisomy 21.** *PLoS Genet* 2010, **6**(11):e1001212.

39. Lubecka K, Kurzava L, Flower K, Buvala H, Zhang H, Teegarden D, Camarillo I, Suderman M, Kuang S, Andrisani O: **Stilbenoids remodel the DNA methylation patterns in breast cancer cells and inhibit oncogenic NOTCH signaling through epigenetic regulation of MAML2 transcriptional activity.** *Carcinogenesis* 2016, **37**(7):656-668.
40. Ying J, Li H, Yu J, Ng KM, Poon FF, Wong SCC, Chan AT, Sung JJ, Tao Q: **WNT5A exhibits tumor-suppressive activity through antagonizing the Wnt/ β -catenin signaling, and is frequently methylated in colorectal cancer.** *Clinical cancer research* 2008, **14**(1):55-61.
41. Walter J, Hutter B, Khare T, Paulsen M: **Repetitive elements in imprinted genes.** *Cytogenetic and genome research* 2006, **113**(1-4):109-115.
42. Takikawa S, Wang X, Ray C, Vakulenko M, Bell FT, Li X: **Human and mouse ZFP57 proteins are functionally interchangeable in maintaining genomic imprinting at multiple imprinted regions in mouse ES cells.** *Epigenetics* 2013, **8**(12):1268-1279.
43. Bell AC, Felsenfeld G: **Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene.** *Nature* 2000, **405**(6785):482-485.
44. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM: **CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus.** *Nature* 2000, **405**(6785):486-489.
45. Jürgens AS, Kolanczyk M, Moebest DC, Zemojtel T, Lichtenauer U, Duchniewicz M, Gantert MP, Hecht J, Hattenhorst U, Burdach S: **PBX1 is dispensable for neural commitment of RA-treated murine ES cells.** *In Vitro Cellular & Developmental Biology-Animal* 2009, **45**(5-6):252-263.
46. Alonso-Martin S, Rochat A, Mademtoglou D, Morais J, De Reyniès A, Auradé F, Chang TH-T, Zammit PS, Relaix F: **Gene expression profiling of muscle stem cells identifies novel regulators of postnatal myogenesis.** *Frontiers in cell and developmental biology* 2016, **4**:58.
47. Saini A, Al-Shanti N, Stewart C: **C2 skeletal myoblast survival, death, proliferation and differentiation: regulation by Adra1d.** *Cellular Physiology and Biochemistry* 2010, **25**(2-3):253-262.
48. Chen G, Lee EY-H: **Phenotypic differentiation without permanent cell-cycle arrest by skeletal myocytes with deregulated E2F-1.** *DNA and cell biology* 1999, **18**(4):305-314.
49. Rao VK, Ow JR, Shankar SR, Bharathy N, Manikandan J, Wang Y, Taneja R: **G9a promotes proliferation and inhibits cell cycle exit during myogenic differentiation.** *Nucleic acids research* 2016, **44**(17):8129-8143.
50. Shi J, Luo L, Eash J, Ibebunjo C, Glass DJ: **The SCF-Fbxo40 complex induces IRS1 ubiquitination in skeletal muscle, limiting IGF1 signaling.** *Developmental cell* 2011, **21**(5):835-847.
51. Gallo D, Gesmundo I, Trovato L, Pera G, Gargantini E, Minetto MA, Ghigo E, Granata R: **GH-releasing hormone promotes survival and prevents TNF- α -induced apoptosis and atrophy in C2C12 myotubes.** *Endocrinology* 2015, **156**(9):3239-3252.
52. Tajbakhsh S, Borello U, Vivarelli E, Kelly R, Papkoff J, Duprez D, Buckingham M, Cossu G: **Differential activation of Myf5 and MyoD by different Wnts in explants of mouse paraxial mesoderm and the later activation of myogenesis in the absence of Myf5.** *Development* 1998, **125**(21):4155-4162.

53. Otto A, Schmidt C, Luke G, Allen S, Valasek P, Muntoni F, Lawrence-Watt D, Patel K: **Canonical Wnt signalling induces satellite-cell proliferation during adult skeletal muscle regeneration.** *Journal of cell science* 2008, **121**(17):2939-2950.

Figures

Figure 1

Pipeline for identification of imprinted genes. (a) Crossing scheme used to generate reciprocal F1 piglets. (b) schematic descriptions of the outline for detecting imprinted genes.

Figure 2

Information used during the identification of imprinted genes. (a) Overlap of genes with either paternal or maternal SNVs among the four F1 individuals in each cross. (b) removing the genes containing the paternal and maternal SNVs simultaneously in each cross. (c) imprinted genes were obtained by keeping the intersection of genes with paternal or maternal SNVs between two crosses.

P2RX2

E2F1

Figure 3

CpG islands in the promoters of the imprinted genes

Chr4: 89,375,545 - 90,539,670

Chr5: 61,181,192 - 62,820,173

Figure 4

Imprinted genes tend to occur in clusters.

Figure 5

Specific motifs in the promoters of imprinted genes.

Figure 6

Gene ontology analysis of imprinted genes

Figure 7

Interaction of imprinted genes related to skeletal muscle development. Rounded rectangles represent candidate biological processes. Green nodes represent paternally imprinted genes. Red nodes represent maternally imprinted genes. Circles indicate the imprinted genes not possessing predicted CpG islands. Triangles indicated the imprinted gene possessing predicted CpG islands.

Supplementary Files

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

- [TableS6G0forpaternalimprintedgenes.xls](#)

- [TableS5GOformaternalimprintedgenes.xls](#)
- [TableS4.xls](#)
- [TableS3.xls](#)
- [TableS2.xls](#)
- [TableS1.xls](#)