Title: A causal SNP in the promoter of myogenin is essential for lean meat production

Authors: Zhuhu Lin [1, #], Qingshuang Zhou[1, #], Xiaoyu Wang [1], Ziyun Liang [1], Rong Xu [1], Xian Tong [1], Chenggan Li[1], Renqiang Yuan[1,2], Yaosheng Chen [1], Delin Mo [1, *]

Affiliation: [1] *State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510006, Guangdong, China.*

[2] *Guangxi Yangxiang Agriculture and Husbandry Co., LTD, Guigang, 537100, China.*

[#] These authors are co-first authors of the article.

*** Corresponding author:** Delin Mo, E-mail: modelin@mail.sysu.edu.cn.

Author information:

Zhuhu Lin, E-mail: linzhh55@alumni.sysu.edu.cn;

Qingshuang Zhou, E-mail: mamall0818@163.com;

Xiaoyu Wang, E-mail: wangxy067@163.com;

Ziyun Liang, E-mail: liangzy28@mail2.sysu.edu.cn;

Rong Xu, E-mail: xurong9@mail2.sysu.edu.cn;

Xian Tong, E-mail: tongxian_jiangsu@163.com;

Chenggan Li, E-mail: lichg33@163.com;

Renqiang Yuan, E-mail: yuanrq6@mail.sysu.edu.cn;

Yaosheng Chen, E-mail: chyaosh@mail.sysu.edu.cn.

Abstract

Single nucleotide polymorphisms (SNPs) widely existing in different breeds genome represent population-specific. Under the influence of long-term evolution and artificial selection, there are a large number of SNPs between western lean-type pig breeds and Chinese indigenous pig breeds, but until now, little is known about their roles in inter-breed differences. Our study revealed SNP rs3471653254 C>T generated from the two types of pigs mentioned above, located in the promoter shared by myogenin and Myoparr, played an important role in the differentiation of myoblast by influencing the enrichment of HOXA5 to regulate the transcription of myogenin and Myoparr. Meanwhile, Myoparr could be used as the sponge of miR-30b-3p which repressed myogenic differentiation and muscle regeneration through targeting MyoD. Our results indicated that SNP rs3471653254 C>T is essential for myogenic differentiation and regeneration and could be used as an ideal site for increasing lean meat production in pigs.

Keywords

SNP, Myoparr, myogenin, miR-30b-3p, muscle development, HOXA5

Introduction

The skeletal muscle, representing nearly 40% of adult tissue in mammal, is the most abundant and highly complex tissue(Bentzinger et al., 2012; Frontera and Ochala, 2015). The development of vertebrate skeletal muscle includes the specification of mesodermal precursor cells into myoblasts, followed by subsequent proliferation, differentiation and fusion of these cells into multinucleated myotubes during embryonic development, continually regenerating and self-renewal throughout life(Molkentin and Olson, 1996). Myogenesis is a highly ordered process regulated spatiotemporally by a lot of regulatory networks and signal pathways(Braun and Gautel, 2011). Myogenic regulatory factors (MRFs), including Myf5, MyoD, Myogenin (MyoG) and MRF4, are the core components regulating myogenesis. In early myogenesis, Myf5, MyoD are in charge of mediating the initial specification of skeletal myoblasts. MyoG, MRF4, and myocyte-specific enhancer factors induce differentiation of these specified cells(Nie et al., 2020; Zhu et al., 2021). Myosin heavy chain (MyHC) protein, also greatly functions in muscle contraction(Blum and Dynlacht, 2013; Buckingham and Rigby, 2014; Forcales et al., 2012).

In addition to the myogenic transcription factors, non-coding RNAs such as microRNAs (miRNAs), long non-coding RNAs (lncRNAs) and circular RNAs have been reported to be involved in the development of skeletal muscle and play critical roles in myogenesis(Gorospe et al., 2020; Liu et al., 2020; Zhang et al., 2018). LncRNAs are pervasively transcribed polyadenylated RNA molecules that are >200 bp

in length(Lee et al., 2016; Michalik et al., 2014). While a striking 75% of human genome was transcribed, only 2% of these genes encode proteins, and the vast majority of transcription produces lncRNAs, as reported by the ENCODE Project Consortium(Consortium, 2012). Increasing evidence indicates that promoter-associated lncRNAs act in cis to regulate the transcription of protein-coding genes, particularly as developmental regulators(Hamazaki et al., 2015). To give a prominent example, one of its functional mechanisms is provided by competing endogenous RNA (ceRNA), which bears miRNA binding sites. Then miRNAs would induce the decay of mRNA and translational suppression through the interaction with the complementary sequences in the 3'UTR of target gene(Akgul and Erdogan, 2018; Bartel, 2009), then enabling them to have important regulatory roles in diverse physiological and developmental processes(Friedman et al., 2009).

Single nucleotide polymorphism (SNP), is a kind of DNA sequence polymorphism caused by a single nucleotide variation at the genomic level(Brookes, 1999). On the other hand, nucleotide mutation is an irreversible sequence variation in DNA, which in essence includes all variations that occur in the species genome either spontaneously or non-spontaneously(Brown, 2002). The distribution of SNPs is population-specific, which may be one of the reasons for phenotypic differences among populations. Under the influence of long-term evolution and artificial selection, pig breeds from all over the world show differences in meat production, growth rate, feed conversion ratio etc(Wang et al., 2022). Western lean-type breeds, such as Duroc pigs (DR), have been

intensively selected over the past several decades for high lean meat production, while Chinese indigenous breeds, such as Guangdong small-ear spotted pig (GS), have lower growth rates and lean meat percentage without sustained artificially selection(Suzuki et al., 1991; Tang et al., 2007). Such a significant difference in meat production traits must be caused by genetic variation. Moreover, SNPs are widely distributed in the swine genome, with an average distribution interval of 300 to 400 bp (Jungerius et al., 2005). By comparing the genomes, it was found that there is a great quantity of SNPs between western lean-type pig breeds and Chinese indigenous pig breeds(Amaral et al., 2009; Kerstens et al., 2009; Ramos et al., 2009). Until now, several studies have indicated that SNPs are implicated in skeletal muscle formation, but little has been known about its roles and molecular mechanism on inter-breed differences(Khanal et al., 2021; Ma et al., 2022; Reguero et al., 2021; Tong et al., 2015). What's more, significant phenotypic variations occurred in Yuedong Black pigs (YDB) from 1975 to 2006. At present, YDB assume a new body-weight and body-shape, unexpectedly close to western lean-type pigs (Figure S1). By whole genome re-sequencing of Large White (LW), Landrace, Duroc and 16 kinds of Chinese indigenous pig breeds (including YDB), we found that the MyoG gene sequence of YDB is much closer to western lean-type pigs. Further, gene introgression analysis revealed that a long DNA fragment of approximately 950kb including myogenin gene originated from Duroc, which is line with YDB’s history of introducing hybridization (Wang, 2021). Based on the previous research, we hypothesize that the population-specific SNPs of myogenin sequence led to the body

change of YDB, but also serve as one of the main reasons for the difference in meat production between the two kinds of pig breeds. Therefore, we conduct research on SNPs among lean-type pig breeds and Chinese indigenous, providing theoretical knowledge support for explaining the differences of two kinds of pig breeds. Proceed to the next step, these SNPs could be expanded on techniques such as marker-assisted selection or gene editing for increasing lean meat production.

In this research, SNP rs3471653254 C>T, located in the promoter at 299 base pairs before translation start site (TSS) of MyoG, played an important role in myoblast differentiation. It regulates the transcription of MyoG and Myoparr by influencing the enrichment of transcription factor homeobox A5 (HOXA5) on the shared promoter. Meanwhile, Myoparr can be used as the sponge of miR-30b-3p, which constitutes a ceRNA. We further found miR-30b-3p overexpression represses myogenesis in C2C12 cells and muscle regeneration along with attenuating myogenic genes expression, whereas miR-30b-3p depletion has the opposite effects. Furthermore, we demonstrate miR-30b-3p regulates myoblast differentiation by targeting MyoD. In a word, SNP rs3471653254 C>T was found to be a vital and multifunctional gene site regulating myogenesis. Moreover, our results elucidate a novel myogenic inhibitor miR-30b-3p and its regulatory network in myogenesis.

Materials & methods

Animal

Wild-type mice (C57BL/6) were purchased from Cyagen Biosciences Co., Ltd. (Suzhou) and were housed under specific-pathogen-free conditions. All mice used in this study had a C57BL/6J genetic background and were housed in SPF conditions during the experiment. Guangdong small-ear spotted and Duroc pigs were acquired from the Guangdong YIHAO Food Co. Ltd (Guangzhou, Guangdong Province). All animals were kept in an environment in which they had sufficient food, were allowed to move freely, and had suitable lighting. Housing, husbandry and all experimental protocols for mice and pigs were approved and performed in accordance with the guidelines established by the Animal Care and Use Committee of Guangdong Province and conducted according to ethical standards.

Cardiotoxin injury

Cardiotoxin (CTX, Sigma) induced muscle injury/regeneration model in adult mouse is an established, reliable model to study muscle regeneration(Goetsch et al., 2003). CTX was dissolved in sterile saline to a final concentration of 10 mmol/L. The mice's hindlimbs were cleaned with 75% alcohol. 50 μL of 10 mmol/L CTX was injected into the left and right TA muscles which have been transfected with agomiR-30b-3p and NC (negative control) one day before, respectively. Let these mice naturally recover under specific-pathogen-free conditions. Then regenerating TA muscles were harvested 3 and

10 days after CTX injection. These tissues were cleaned in phosphate buffer saline (PBS) and were quickly frozen in liquid nitrogen for future experiments.

Primary myoblasts isolation and cells culture

Dorsal muscle of porcine E35 embryo was isolated and separated by 1 mg/ml type I collagenase (Sigma) in Dulbecco's modified Eagle's Medium (DMEM, Gibco) at 37℃ for 1.5-2 h. At E35, myoblasts are highly active in forming primary myofibers and express key myogenic regulatory factors. The suspension was fermented with ten milliliters of culture media within 20% fetal bovine serum (FBS, MIKX) and centrifuged at 1600 × g for 10 min (High-speed Refrigerated Centrifuge, Eppendorf). The pellet was resuspended in 10 ml of growth media (20% fetal bovine serum /DMEM + 2.5ng/ml bFGF), filtered through a 100 μm cell strainer, and plated on a 10 cm Matrigel coated culture dish. The culture is incubated in a small volume of PBS, gently tapping the plate to displace the myoblasts to enrich the myoblasts. When cells growing into 80% confluence, the culture medium was replaced to DMEM with 2% horse serum (FBS, MIKX) to induce differentiation.

The C2C12 mouse myoblast cell line was supplied by American Type Culture Collection (ATCC). C2C12 myoblasts were cultured in DMEM containing 10% fetal bovine serum and 1% penicillin/streptomycin in a humidified atmosphere with 5% CO2 at 37 °C. When cells reached 90% confluence, the medium was replaced into DMEM with 2% horse serum and 1% penicillin/streptomycin to induce differentiation.

Overexpression

The sequences of MyoG and HOXA5 were cloned into pcDNA3.1 vector. C2C12 cells were seeded into 6-or 12-well plates at 12 hours before treatment and then transfected at 80% confluence then incubate 20 minutes for each well. Dilute expression plasmids DNA 1-4 μg per well of a 6-well plate in Opti-MEM. Dilute transfection reagent in Opti-MEM using MK40 (MIKX) according to the manufacture's instruction. Incubate for 5 minutes at RT. Add the DNA-reagent complex dropwise to the cells. Gently swirl the plate to distribute evenly. After transfection for 8-12 h，the culture medium was substituted by fresh DMEM containing 10% FBS. Transfections were performed at least in triplicate for each experiment.

AgomiR-30b-3p and antagomiR-30b-3p treatment

AgomiR-30b-3p and antagomiR-30b-3p (RiboBio), were used to specifically up-regulate or down-regulate miR-30b-3p activity. Following the instructions, agomiR-30b-3p and antagomiR-30b-3p dissolved to a concentration of 20μM were added directly into the culture medium of cells. After transfection for 12 h，the culture medium was substituted by fresh DMEM containing 10% FBS. After 2 days of treatment, the cells were harvest to detect the efficiency of activation or suppression.

RNA extraction and real-time quantitative PCR (qPCR)

The cells and regenerating tibialis anterior (TA) muscles were washed with PBS and harvested for total RNA extraction using TRIzol reagent (Vazyme). 1 μg of the total RNA was processed into single-stranded complementary DNA (cDNA) using a StarScript II First-strand cDNA Synthesis Kit (GenStar) following manufacturer's instructions. Quantitation of the mRNA levels by qPCR was performed on a real-time PCR system (Roche LightCycler 480 High-Throughput Fluorescence qPCR System) using SYBR Green Master Mix (GenStar). The relative mRNA expression level was normalized to that of GAPDH. Gene expression was quantified by comparative CT method. The primers used for qPCR are listed in Table 1.

Western blot

Protein extracts of cultured C2C12 cells or TA muscles were obtained using protein extraction buffer (FDbio) within 1% protease inhibitor cocktail. The lysed samples were centrifuged at 14, 000g and 4°C for 10 min and supernatants containing protein were collected. Then 5X protein loading buffer was added to the lysates prior to their full denaturation in boiling water for 10 min. Equivalent amount of protein samples were separated in 8% or 10% SDS-PAGE and transferred to polyvinylidene difluoride membrane (Millipore). After blocked with 4% bovine serum albumin (BSA) for 1 hour, membranes were incubated overnight at 4°C with the specific primary antibodies shown in Table 2. The membranes were washed with Tris-buffered saline with Tween

20 (TBST) and then incubated with corresponding secondary antibodies for 1 h at room temperature. Blots were visualized by the ECL chemiluminescence system (Guangzhou Biolight Biotechnology).

Droplet Digital Polymerase Chain Reaction (ddPCR) and data evaluation

After the optimization processes were completed, the ddPCR system was used to quantify the presence of suitable genes in the samples. For this purpose, 2×dPCR EvaGreen Master Mix (Rox), (Sniper, Suzhou, Chian) 0.5 μM of each primer F/R and cDNAs were mixed and the total volume of the reaction was adjusted to 22 μL with molecular grade dH2O. After the PCR conditions were selected, the device was started to amplify the genes. The samples were into the Sniper DQ24 Digital System for gene quantification. When the program finishes, the droplets were measured by the instrument, and data analysis was performed using the SightPro software program. The results were obtained by multiplying the data from the results of the study with the dilution coefficients. The result was shown in Supplemental table 3.

The times and temperatures used as thermal protocol are as the following: droplet stabilization 5 min at 60°C and enzyme activation 15 min at 95°C, followed by 40 cycle, denaturation 15 s at 95°C, binding 15 s at 60°C, extension 15s at 72°C.

Immunofluorescence

The cells cultured in 12-well plates or 24-well plates were fixed in 4%

paraformaldehyde for 10 minutes, followed by permeabilization in 0.5% Triton X-100 in PBS for 15 minutes. Samples were then blocked in 4% bovine serum albumin in PBS for 1 h at RT. Then, the cells were incubated with primary antibodies overnight at 4°C. Washed three times with PBS, the cells were incubated with secondary antibodies for 1 hour at room temperature. Finally, the cells were washed thrice in PBS, and nuclei were stained by DAPI. Antibodies are listed in Table 2. Immunostaining images were obtained via fluorescent reverse microscopy (Nikon).

Histology

Freshly isolated adult TA muscles were immediately fixed in 4% paraformaldehyde at 4 °C for 18 h and subsequently embedded in paraffin. 4 μm thick cross-sections of TA muscles were subjected to H&E staining and Masson's trichrome staining which was performed according to procedures provided by the Hematoxylin-eosin (H&E) staining kit and Masson's Trichrome Stain Kit.

Luciferase reporter assay

Luciferase reporter plasmids (psiCHECK2, Sangon) were transfected into 293T (ATCC). cells seeded in a 24-well plate. After 36 h transfection, cells were lysed and enzymic reactions were assayed by using the Dual Luciferase reporter assay system (Promega). The firefly luciferase activity was normalized to renilla luciferase internal control to exclude the differences of transfection efficiency.

Chromatin immunoprecipitation

293T cells transfected were cross-linked with 1% formaldehyde for 10 minutes at room temperature. The fixing solution was quenched for 5 min by adding glycine. Cells were lysed in lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5% SDS, 20 μg/ml proteinase K) on ice and chromatin was sonicated using an ultrasonic disruptor (Covaris Sonicator) for 8 min to generate chromatin fragments of 200-300 bp DNA. The clarified nuclear extracts were incubated with IgG antibody applied as negative control or target antibody shown in Table 2 with rotation overnight at 4°C and conjugated with Chromatin immunoprecipitation-grade protein G magnetic beads (Cell Signaling Technology). After extensive washing, bound DNA fragments were purified and eluted by elution buffer. The enrichment of DNA sequences was analyzed via qPCR using specific primers. Data were normalized to the respective control IgG values and assessed relative to the input DNA.

Isolation of longissimus dorsi muscle cells for scRNA-seq

Single-cell suspensions were prepared from fresh longissimus dorsi muscle specimens dissected from Yorkshire porcine embryos at E18, E28, E35 and postnatal individuals at E50 and E73 developmental time points. To generate each biological replicate, muscle tissues from five conceptuses or neonates derived from a single sow were combined into a composite sample. Immediately following collection, tissues were

preserved in chilled RPMI 1640 medium (GIBCO) containing 10% fetal bovine serum (HyClone). After three PBS washes, tissues were mechanically disaggregated using sterile blades and subsequently subjected to enzymatic digestion with 0.1 g/mL Collagenase I (Sigma) and 100 μg/mL DNase I (Sigma-Aldrich) in RPMI 1640 at 37°C for 30-40 minutes. The digestion process was intermittently agitated at 6-8 minutes intervals. The liberated cells were sequentially filtered through a 70-μm mesh (BD), collected in conical tubes, and evaluated for viability and concentration with Trypan Blue (Thermo Fisher) staining using an automated cell counting system (Countstar). All samples from distinct developmental stages were processed separately to maintain traceability to their origin.

10× Genomics single cell RNA sequencing (scRNA seq)

Single-cell sequencing libraries were constructed according to the manufacturer's protocol for the Chromium system (10× Genomics, PN120263). Following droplet formation, samples underwent reverse transcription in prechilled 8-well strips (Eppendorf) using a Veriti 96-well thermal cycler (Thermo Fisher). The synthesized cDNA was subsequently recovered with the proprietary Recovery Agent (10× Genomics) and purified using Silane Dynabeads (Thermo Fisher). After 12 amplification cycles, the cDNA products were cleaned up with SPRIselect beads (Beckman), quantified through Bioanalyzer (Agilent Technologies) analysis at 1:4 dilution, and finally processed into sequencing libraries following the Single Cell 3′

Reagent Kit v3 manual with PCR cycle optimization based on measured cDNA concentrations.

Single-cell RNA sequencing data analysis method for HOXA5 expression profiling

Bioinformatic processing commenced with alignment of raw sequencing reads to the Sscrofa11 reference genome (Ensemble annotations v11.1.98) through the CellRanger pipeline (10× Genomics). The resulting gene expression matrix underwent rigorous quality control in R v3.5.2 environment, implementing thresholds for cell inclusion (>200 genes detected, <30,000 UMIs, <10% mitochondrial gene content) and gene representation (detected in ⩾3 cells). Potential doublets exceeding 40,000 transcripts were excluded via DoubletFinder. Following normalization and scaling procedures, highly variable genes were selected for principal component analysis, with subsequent graph-based clustering and UMAP visualization. To investigate HOXA5 expression dynamics during myofiber maturation, we combined UMAP visualization, violin plots, and differential expression analysis within a statistical framework. Our analysis traced the developmental trajectory from myoblasts through secondary and primary fiber formation, culminating in myotube assembly, revealing stage-specific expression patterns of HOXA5.

Statistical analysis

All dot plots and graphs were generated using the Prism 9 software (GraphPad Software). Experiments were done with a minimum of three biological replicates. For all experiments where "n=3"or "n=5" is indicated, it represents three or five independent biological replicates. For each of these biological replicates, measurements were performed in technical replicates, and the data presented are the mean values calculated from the three biological replicates. Data points represent the mean of technical replicates for each biological replicate, and the values presented in the figures are the mean ± SD derived from the three biological replicates. Data are presented as mean ± SD, and the statistical significance analysis was performed using an unpaired two-tailed Student's t test to test differences between groups. The level of significance is indicated as follows: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Result

Pig breed specific SNPs at myogenin promoter regulate myoblast differentiation

By comparing the genome sequences of myogenin between 4 western lean-type pig breeds and 16 Chinese indigenous pig breeds, a total of 11 SNPs whose gene frequency and genotype frequency showed significant differences between the two kinds of pig breeds were selected. Then two kinds of MyoG expression plasmid (including a 5 ' flank sequence of 894 bp) were constructed according to SNPs based on western lean-type pig breed and Chinese indigenous pig breed (Figure 1A). The MyoG plasmid derived from the western lean-type pig breed was abbreviated as plasmid-DR (Duroc, a representative western lean-type pig breed), and the MyoG plasmid derived from the Chinese indigenous pig breed was abbreviated as plasmid-GS (Guangdong small-ear spotted pig, a representative Chinese indigenous breed) in the following text and figure. The whole MyoG was overexpressed in porcine primary myoblasts and 293T cells using plasmid transfection (Figure 1B-1D and Figure 1H-1I). As a result, the cells transfected with plasmid-DR show remarkably higher myogenic differentiation efficiency as determined by immunofluorescence staining of MyHC (Figure 1E). Compared with the plasmid-GS, the fusion index and differentiation index of porcine primary myoblasts transfected with plasmid-DR increased by nearly 13% (Figure 1E). In addition, the expression of genes related to myogenesis was monitored by qPCR and Western blot. MyHC and MyoD showed higher expression in the cells transfected with plasmid-DR. (Figure 1F-1G). In the same way for mouse species, when DR' MyoG

overexpressed in C2C12 cells, significant promotion was observed in both myoblast differentiation and myogenic marker genes expression (Figure S2A-S2F). To explore which SNP plays the key role, the plasmid-DR were respectively designed to mutate on the upstream or downstream (Figure S2G). The mRNA and protein expression level of MyoG shows almost no difference between downstream mutation and DR's wild MyoG, but decrease significantly when mutations occurred in upstream (Figure 1J). These results prove that SNPs in the upstream region of MyoG affect the expression of myogenic genes and myoblast differentiation.

Identification of lncRNA Myoparr and acts as a ceRNA sponging miR-30b-3p

In order to figure out why SNPs at promoter can affect the expression of myogenin and myoblast differentiation, we analyzed the upstream sequence of human myogenin and found human Myoparr (Gene ID: 114004358), which is a lncRNA. Then the porcine Myoparr (approximately 4183 bp long) was obtained through homologous comparison, and its transcript is detectable in pigs (Figure 2A). Like humans, Myoparr and myogenin in pigs are also located on the positive and negative chains, respectively, with large repeat regions that cover both exons and introns of myogenin (Figure S3A). More importantly, Myoparr shares the same promoter together with MyoG and presented significantly higher expression after transfection with plasmid-DR over plasmid-GS (Figure 2B, Figure S3B). An online Coding Potential Calculator (CPC) analysis indicated that the lncRNA Myoparr has no coding potential. This analysis used MyoG

as a positive control and the well-documented non-coding RNA HOTAIR as a non-coding control (Figure 2C). (Hajjari and Salavaty, 2015). QPCR analyses have shown that Myoparr expression was predominantly enriched in skeletal muscle tissue among the organs of 180-day-old pigs as same as MyoG and MyoD expression (Figure 2D). In addition, during skeletal muscle development, the expression trends of Myoparr were similar to the key myogenic genes, indicating its involvement in embryonic myogenesis (Figure 2I). During C2C12 differentiation, the expression level of Myoparr was peaked on DM 2d and subsequently decreased from DM 2d to DM 8d, as the same expression trend as MyoG and MyoD (Figure 2J). Considering that lncRNAs usually sponge miRNAs to modulate gene expression, an online prediction website miRBase was employed, and miR-30b-3p was screened as a candidate target for it closely bound to the base of Myoparr and evolutionally conserved among multiple species (Figure 2E-2F, Figure S3C). Moreover, the amount of Myoparr and miR-30b-3p are quantified by ddPCR, implying the enough potential for their combination (Figure S3D-S3E). To verify whether miR-30b-3p can directly target Myoparr, we cloned the Myoparr DNA segment containing either the wild-type (WT) binding sites or mutant (Mut) binding sites of miR-30b-3p into luciferase reporter plasmids (Figure 2G). Dual-luciferase reporter assays showed that overexpressing miR-30b-3p could reduce the luciferase activity of Myoparr-WT, but not affect the luciferase activity of the mutant (Figure 2H), which confirmed that miR-30b-3p was sponged with the predicted target site by Myoparr. As expected, the expression trend of miR-30b-3p was contrary to Myoparr

(Figure 2I-2J)，indicating their strong negative correlation.

MiR-30b-3p impairs myogenic differentiation

In order to investigate the function of miR-30b-3p during myogenic differentiation, porcine primary myoblasts were transfected with agomiR-30b-3p or negative control (NC) at 80% confluence and then induced to differentiation. As shown in Figure 3A, the expression of miR-30b-3p was highly up-regulated. MyHC immunofluorescence staining assay (Figure 3B) demonstrated that overexpressing miR-30b-3p prevented myognic differentiation, along with fewer myotubes and almost no thick muscle fibers (Figure 3B). Consistent with this result, the protein and mRNA expression levels of myogenic factors, including MyoD, MyoG, MyHC and Myoparr, were all decreased (Figure 3C-3L). Furthermore, the same phenomenon also appeared in C2C12 cells (Figure S4). The above data indicated that overexpression of miR-30b-3p impairs the differentiation of myoblasts. AntagomiR-30b-3p and NC were respectively transfected into porcine primary myoblasts, and the expression of miR-30b-3p was decreased nearly 70% (Figure 4A). MiR-30b-3p knockdown in primary myoblasts promoted myoblast differentiation into myotubes, evidenced by an increased number of MyHC+ cells and fusion rate with a lot of extraordinary thick myofibers in the visual field (Figure 4B). Correspondingly the expression levels of myogenic factors, such as MyoD, MyoG, MyHC and Myoparr, were markedly improved (Figure 4C-4L), which was also observed in C2C12 cells (Figure S5). The above results indicated that depletion of miR-

30b-3p promotes myogenic differentiation.

MiR-30b-3p delays skeletal muscle regeneration

To further confirm the function of miR-30b-3p, muscle regeneration model was established, and the treatment was illustrated in Figure 5A. AgomiR-30b-3p or NC was injected into TA muscle regeneration model every four days. Then, regenerating TA samples were harvested at days 3 and days 10, respectively. The staining on TA cross-sections revealed that overexpression of miR-30b-3p lead to less efficient regeneration characterized by the smaller myofiber cross-sectional area and more inflammatory cells, which is consistent with directly visible TA muscle injury status (Figure 5B-5D). The protein and mRNA of myogenic key factors were dramatically down-regulated in TA muscles administrated with agomiR-30b-3p at day 3 (Figure 5E-5N). These results indicate that superfluous miR-30b-3p is detrimental for muscle regeneration. In support of this notion, at 10 days after injection, TA muscle transfected by agomiR-30b-3p showed less efficient regeneration than NC according to the number and size of myofibers (Figure S6A-S6M). Together, above evidence confirms that miR-30b-3p does regulate muscle development and regeneration of injury muscle.

MiR-30b-3p targets the 3'UTR of MyoD

Given that miR-30b-3p makes a difference in myogenic differentiation, we further explored the molecular mechanisms of miR-30b-3p in myogenic regulation. In general,

miRNAs are usually involved in the post-transcriptional level and bind to 3' UTR of their target mRNA to suppress expression. The software TargetScan and miRDB were used to predict the target mRNA of miR-30b-3p. The overlapping target genes obtained from these two databases were subjected to GO and KEGG enrichment analyses to assess their potential functional roles (Figure S7A-S7B). Although GO and KEGG enrichment analyses of the overlapping targets did not yield significantly enriched terms directly linked to myogenesis, further examination of the underlying data table identified MyoD as a key target gene present in the dataset. Based on the analysis results and its established relevance to our research context, MyoD was selected for further experimental validation (Figure 6A-6B). Then the luciferase reporter vectors were constructed containing wild-type binding site of 3' UTR (MyoD 3' UTR-WT) or mutated binding site of 3' UTR (MyoD 3' UTR-MUT) (Figure 6C). As a result, only the luciferase activity of MyoD 3' UTR-WT but not MyoD 3' UTR-MUT could be reduced by agomiR-30b-3p (Figure 6D).

SNP rs3471653254 regulates myogenin via affecting the binding of HOXA5

Above results have shown that SNPs in the upstream of MyoG affected myoblasts differentiation, but we wondered how the SNPs play the regulatory role. Therefore, based on the sequence of the promoter, we predicted SNP rs3471653254 C>T, which located at 299 sites before the TSS, is an important site for HOXA5 binding using the JASPAR database. The predicted HOXA5 binding affinity is significantly higher in the

445 promoter segment of DR than that of GS (Figure S8A). And we found HOXA5 keeps
446 evolutionally conserved in multiple species and its binding site motif (Figure 7A-7B).
447 Furthermore, we examined the differences in HOXA5 expression levels during
448 embryonic myogenic differentiation between the two pig breeds, as well as the
449 expression patterns of HOXA5 during myoblast differentiation (Figure 7C-7D). Our
450 analysis traced the developmental trajectory from myoblasts through fiber formation,
451 culminating in myotube assembly, revealing stage-specific expression patterns of
452 HOXA5 (Figure S8B). HOXA5 exhibits relatively high expression in myoblasts and
453 during the cell differentiation stage, whereas its expression level decreases significantly
454 at the myofiber stage which suggests that HOXA5 may be involved in the process of
455 myogenesis. To verify this speculation, the promoter segments harboring different
456 SNPs were cloned into luciferase reporter vector (Figure 7E-7F). Reporter vectors were
457 co-transfected into 293T cells with empty or plasmids expressing HOXA5. As
458 illustrated in Figure 7G, HOXA5 obviously elevated the activity of promoters which
459 contain the first SNP keeping T base. Moreover, the interaction between the promoter
460 region and HOXA5 was confirmed by chromatin immunoprecipitation. Significantly
461 higher HOXA5 enrichment levels were detected in promoter segment of DR than that
462 of GS (Figure 7H). The mRNA and protein levels of HOXA5 were detected to
463 authenticate the efficiency of overexpression. (Figure 7I-7K). Along with the up-
464 regulation of HOXA5 on the promoter activity, MyoG and Myoparr expressing levels
465 were significantly increased by 30%-70% in C2C12 cells (Figure 7L-7O). Overall, our

findings elucidate that upstream binding site in which SNP rs3471653254 C>T exists is an enhancer of MyoG and Myoparr, and HOXA5 promotes their expression by being recruited.

Discussion

The domestic pig is an essential protein source for humans. Genetic and phenotypic differentiation between pig breeds result from geographical divergence, local adaptation, and artificial selection(Ai et al., 2015; Li et al., 2017). Compared to Chinese indigenous pig breeds, western lean-type pig breeds have been undergone nearly a century of artificial selection aimed at generating faster growth and higher lean meat production (Li et al., 2017; Zhao et al., 2011). These differences make them ideal research models to elucidate the potential mechanisms of phenotypic differentiation of lean meat production (Tang et al., 2007; Zhao et al., 2011; Zhao et al., 2015). Large-scale whole-genome resequencing revealed that fatty-type pig and lean-type pig breeds host distinct alleles at several SNP sites in the promoter region of the MyoG gene (Wang, 2021). Thereby, overexpression plasmids carrying different MyoG genotype were constructed respectively, and the DR' MyoG has a stronger promoting effect on myogenic differentiation. These findings are consistent with research showing the higher meat production and growth rate of western lean-type pig breeds, confirming the rationality and reliability of our results. As the DR' MyoG is more conducive to the differentiation of myoblasts, we would like to explore which SNP play a major regulatory role. Through a series of experiments, we found that two SNPs in the upstream may have an important impact on myogenesis.

Meanwhile, the promoter not only transcripts MyoG, but also transcripts lncRNA Myoparr, which is a recently discovered regulator of myogenesis in mouse and human

transcribed from the upstream region of myogenin(Hitachi et al., 2019). To unravel the underlying regulatory mechanism, we predicted the possible miRNAs that may be sponged by Myoparr and found miR-30b-3p. In previous studies, miR-30b-3p functions in anticancer treatment by targeting multiple genes(Chen et al., 2022b; Gao et al., 2019; Li et al., 2020). Here, when miR-30b-3p was overexpressed, the expression of Myoparr WT was obviously inhibited, but fluorescence activity of Myoparr MUT was not affected. In accordance with our expectation, miR-30b-3p holds an opposite expressing tendency to Myoparr no matter during the differentiation stage of C2C12 or during embryonic muscle development. To explore the precise effect of miR-30b-3p on skeletal muscle, we explore its function by employing overexpression and knockdown strategy. It was found that miR-30b-3p not only prevents myoblast differentiation, but also delays the muscle regeneration. From the above, the conclusion can be reached that Myoparr forms a ceRNA regulatory network with miR-30b-3p, and regulates the myoblast differentiation together.

Increasing evidence has demonstrated that miRNAs were involved in the muscle development via interacting with 3'UTR of target mRNAs leading to mRNA degradation or translational repression(Yang et al., 2021). MyoD, the key myogenic differentiation factor, has been well known as a master TF in myogenic cell-lineage specification during development and trans-differentiation(Buckingham and Tajbakhsh, 1999). In this study, it was confirmed as the target of miR-30b-3p which is a novel myogenic inhibitor.

Promoters are specific DNA sequences that RNA polymerase II binds to and initiate transcription. RNA polymerase II requires transcription factors to be recruited to the transcription start site as part of a large transcription pre-initiation complex(Harris and Marles-Wright, 2019). The high conservatism of HOXA5 suggests that it serves a vital function. More importantly, HOXA5 is a crucial transcriptional factor in both tumor suppression and oncogenesis via the interaction with various target genes and signaling pathways(Fan et al., 2022; He et al., 2022; Jin et al., 2023; Liao et al., 2020). Once SNP rs3471653254 is changed from C to T, the enrichment level of HOXA5 to the promoter is up-regulated, which subsequently enhances the promoter activity of MyoG and Myoparr (Figure 8).

We are currently planning to employ CRISPR base editing to introduce a C-to-T single base mutation at SNP rs3471653254 in GDSS pigs. This targeted approach aims to determine the causal effect of this specific genetic variant on lean meat yield and growth rate. Previous studies have established HOXA5 as a critical regulator in mammalian development, directing the development of multiple organs-including the digestive tract, thyroid, and mammary gland-primarily by orchestrating mesenchymal–epithelial interactions that guide cell fate and tissue organization(Jeannotte et al., 2016). In adipose tissue, HOXA5 promotes adipogenic differentiation and lipid deposition in the subcutaneous compartment while suppressing preadipocyte proliferation(Chen et al., 2022a; Parrillo et al., 2023). In our current study, we demonstrated that HOXA5 promotes myogenic differentiation by binding to and activating the MyoG promoter.

Critically, this regulatory process is modulated by the SNP rs3471653254 C>T. This SNP is located on the MyoG promoter, which is specifically expressed in muscle tissue. Except for skeletal muscle tissue, MyoG expression is almost undetectable in other tissues such as liver, brain, kidney, etc. Although SNP affects the enrichment of HOXA5, which is widely expressed, there is currently no evidence to suggest that this SNP on the myogenin promoter affects the expression of HOXA5 in other tissues. Upon successful base editing, we anticipate an enhanced transcriptional activation of MyoG by HOXA5. Given that MyoG expression is spatially restricted to skeletal muscle tissue, this targeted enhancement is expected to be tissue-specific and unlikely to disrupt normal development in other organs. This specificity presents a significant advantage for subsequent breeding applications by minimizing pleiotropic risks. However, this premise requires rigorous validation through further observation and experimentation.

Taken together, our study provides a theoretical basis for explaining the differences in meat production between western lean-type pig breeds and Chinese indigenous pig breeds. These mechanisms can now be interpreted in molecular detail and provide an exciting theoretical framework for past and future functional studies of SNP, giving us unique insight into the complex process of species evolution.

Abbreviations

MyoG	myogenin

MRFs	Myogenic regulatory factors
MyHC	Myosin heavy chain
miRNAs	microRNAs
lncRNAs	long non-coding RNAs
ceRNA	competing endogenous RNA
SNP	Single nucleotide polymorphism
DR	Duroc pigs
GS	Guangdong small-ear spotted pig
YDB	Yuedong Black pigs
TSS	translation start site
HOXA5	transcription factor homeobox A5
CTX	Cardiotoxin
TA	tibialis anterior

Acknowledgments

This work was supported by the National Key Research and Development Program (2023YFD1300200, 32472858), Modern Agricultural Industrial Technology System Innovation Team of Guangdong Province (2024CXTD22), Selection and Breeding of New Local Pig Breeds and Promotion of Industrialization (2024-XPY-00-001), Selection and Breeding of Guangdong small-ear spotted pig (2022-440000-4301030202-9510), and the earmarked fund for CARS-35.

Conflict of Interest Statement

The authors declare no competing interests.

Authors' contributions

Delin Mo: Conceptualization, Methodology, Resources, Supervision, Funding acquisition, Project administration, Writing–review & editing. Zhuhu Lin: Methodology, Formal analysis, Validation, Investigation, Visualization, Writing – original draft. Qingshuang Zhou: Validation, Visualization. Xiaoyu Wang: Methodology, Formal analysis. Ziyun Liang: Project administration. Rong Xu: Resources. Yaosheng Chen: Funding acquisition. Rong Xu: Data curation. Xian Tong: Data curation. Ziyun Liang: Data curation.

Ethics Approval

The animal experimental procedures used in this experiment were approved by the Animal Care and Use Committee of Guangdong Province, China. Approval ID or permit numbers are SYXK (yue) 2023-0313.

Data Availability Statement

All data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Reference

Ai H, Fang X, Yang B, et al. 2015. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat Genet*, **47**(3): 217-225.

Akgul B & Erdogan I. 2018. Intracytoplasmic Re-localization of miRISC Complexes. *Front Genet*, **9**: 403.

Amaral AJ, Megens HJ, Kerstens HH, et al. 2009. Application of massive parallel sequencing to whole genome SNP discovery in the porcine genome. *BMC Genomics*, **10**: 374.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell*, **136**(2): 215-233.

Bentzinger CF, Wang YX, Rudnicki MA. 2012. Building muscle: molecular regulation of myogenesis. *Cold Spring Harb Perspect Biol*, **4**(2).

Blum R & Dynlacht BD. 2013. The role of MyoD1 and histone modifications in the activation of muscle enhancers. *Epigenetics*, **8**(8): 778-784.

Braun T & Gautel M. 2011. Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nat Rev Mol Cell Biol*, **12**(6): 349-361.

Brookes AJ. 1999. The essence of SNPs. *Gene*, **234**(2): 177-186.

Brown T. 2002. Mutation, Repair, and Recombination. Mutation, Repair, and Recombination.

Buckingham M & Rigby PW. 2014. Gene regulatory networks and transcriptional

mechanisms that control myogenesis. *Dev Cell*, **28**(3): 225-238.

Buckingham M & Tajbakhsh S. 1999. 41 - Myogenic Cell Specification during Somitogenesis. *In*: Moody SA. Cell Lineage and Fate Determination. San Diego: Academic Press, 617-633.

Chen D, Lin Y, Zhao N, et al. 2022a. Hoxa5 Inhibits the Proliferation and Induces Adipogenic Differentiation of Subcutaneous Preadipocytes in Goats. *Animals (Basel)*, **12**(14).

Chen L, Chen X, Liu L, et al. 2022b. [miR-30b-3p Inhibits the Proliferation and Invasion of Lung Adenocarcinoma by Targeting COX6B1]. *Zhongguo Fei Ai Za Zhi*, **25**(8): 567-574.

Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**(7414): 57-74.

Fan F, Mo H, Zhang H, et al. 2022. HOXA5: A crucial transcriptional factor in cancer and a potential therapeutic target. *Biomed Pharmacother*, **155**: 113800.

Forcales SV, Albini S, Giordani L, et al. 2012. Signal-dependent incorporation of MyoD-BAF60c into Brg1-based SWI/SNF chromatin-remodelling complex. *Embo j*, **31**(2): 301-316.

Friedman RC, Farh KK, Burge CB, et al. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res*, **19**(1): 92-105.

Frontera WR & Ochala J. 2015. Skeletal muscle: a brief review of structure and function. *Calcif Tissue Int*, **96**(3): 183-195.

628 Gao D, Zhou Z, Huang H. 2019. miR-30b-3p Inhibits Proliferation and Invasion of
629 Hepatocellular Carcinoma Cells via Suppressing PI3K/Akt Pathway. *Front Genet*, **10**:
630 1274.
631 Goetsch SC, Hawke TJ, Gallardo TD, et al. 2003. Transcriptional profiling and
632 regulation of the extracellular matrix during muscle regeneration. *Physiol Genomics*,
633 **14**(3): 261-271.
634 Gorospe M, Abdelmohsen K, De S, et al. 2020. circSamd4 represses myogenic
635 transcriptional activity of PUR proteins. *Nucleic Acids Research*, **48**(7): 3789-3805.
636 Hajjari M & Salavaty A. 2015. HOTAIR: an oncogenic long non-coding RNA in
637 different cancers. *Cancer Biol Med*, **12**(1): 1-9.
638 Hamazaki N, Uesaka M, Nakashima K, et al. 2015. Gene activation-associated long
639 noncoding RNAs function in mouse preimplantation development. *Development*,
640 **142**(5): 910-920.
641 Harris JR & Marles-Wright JJS-CB. 2019. Macromolecular Protein Complexes II:
642 Structure and Function.
643 He ZC, Liu Q, Yang KD, et al. 2022. HOXA5 is amplified in glioblastoma stem cells
644 and promotes tumor progression by transcriptionally activating PTPRZ1. *Cancer Lett*,
645 **533**: 215605.
646 Hitachi K, Nakatani M, Takasaki A, et al. 2019. Myogenin promoter-associated lncRNA
647 Myoparr is essential for myogenic differentiation. *EMBO Rep*, **20**(3).
648 Jeannotte L, Gotti F, Landry-Truchon K. 2016. Hoxa5: A Key Player in Development

649 and Disease. *J Dev Biol*, **4**(2).

650 Jin WY, Zhang Y, Tian J, et al. 2023. HOXA5 inhibits the proliferation and metastasis
651 of cervical squamous cell carcinoma by suppressing the beta-catenin/Snail signaling.
652 *Neoplasma*, **70**(1): 82-93.

653 Jungerius BJ, Gu J, Crooijmans RP, et al. 2005. Estimation of the extent of linkage
654 disequilibrium in seven regions of the porcine genome. *Anim Biotechnol*, **16**(1): 41-54.

655 Kerstens HH, Kollers S, Kommadath A, et al. 2009. Mining for single nucleotide
656 polymorphisms in pig genome sequence data. *BMC Genomics*, **10**: 4.

657 Khanal P, Williams AG, He L, et al. 2021. Sarcopenia, Obesity, and Sarcopenic Obesity:
658 Relationship with Skeletal Muscle Phenotypes and Single Nucleotide Polymorphisms.
659 *J Clin Med*, **10**(21).

660 Lee S, Kopp F, Chang TC, et al. 2016. Noncoding RNA NORAD Regulates Genomic
661 Stability by Sequestering PUMILIO Proteins. *Cell*, **164**(1-2): 69-80.

662 Li M, Chen L, Tian S, et al. 2017. Comprehensive variation discovery and recovery of
663 missing sequence in the pig genome using multiple de novo assemblies. *Genome Res*,
664 **27**(5): 865-874.

665 Li Y, Zhou J, Wang J, et al. 2020. Mir-30b-3p affects the migration and invasion
666 function of ovarian cancer cells by targeting the CTHRC1 gene. *Biol Res*, **53**(1): 10.

667 Liao Y, Wang C, Yang Z, et al. 2020. Dysregulated Sp1/miR-130b-3p/HOXA5 axis
668 contributes to tumor angiogenesis and progression of hepatocellular carcinoma.
669 *Theranostics*, **10**(12): 5209-5224.

Liu Y, Wang J, Zhou X, et al. 2020. miR-324-5p Inhibits C2C12 cell Differentiation and Promotes Intramuscular Lipid Deposition through lncDUM and PM20D1. *Mol Ther Nucleic Acids*, **22**: 722-732.

Ma L, Qin M, Zhang Y, et al. 2022. Identification and functional prediction of long non-coding RNAs related to skeletal muscle development in Duroc pigs. *Anim Biosci*, **35**(10): 1512-1523.

Michalik KM, You X, Manavski Y, et al. 2014. Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ Res*, **114**(9): 1389-1397.

Molkentin JD & Olson EN. 1996. Defining the regulatory networks for muscle development. *Current Opinion in Genetics & Development*, **6**(4): 445-453.

Nie Y, Cai S, Yuan R, et al. 2020. Zfp422 promotes skeletal muscle differentiation by regulating EphA7 to induce appropriate myoblast apoptosis. *Cell Death & Differentiation*, **27**(5): 1644-1659.

Parrillo L, Spinelli R, Longo M, et al. 2023. The Transcription Factor HOXA5: Novel Insights into Metabolic Diseases and Adipose Tissue Dysfunction. *Cells*, **12**(16).

Ramos AM, Crooijmans RP, Affara NA, et al. 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS One*, **4**(8): e6524.

Reguero M, Gomez De Cedron M, Reglero G, et al. 2021. Natural Extracts to Augment Energy Expenditure as a Complementary Approach to Tackle Obesity and Associated Metabolic Alterations. *Biomolecules*, **11**(3).

691 Suzuki A, Kojima N, Ikeuchi Y, et al. 1991. Carcass composition and meat quality of
692 Chinese purebred and European × Chinese crossbred pigs. *Meat Science*, **29**(1): 31-41.
693 Tang Z, Li Y, Wan P, et al. 2007. LongSAGE analysis of skeletal muscle at three prenatal
694 stages in Tongcheng and Landrace pigs. *Genome Biol*, **8**(6): R115.
695 Tong B, Li GP, Sasaki S, et al. 2015. Association of the expression levels in the skeletal
696 muscle and a SNP in the CDC10 gene with growth-related traits in Japanese Black beef
697 cattle. *Anim Genet*, **46**(2): 200-204.
698 Wang C. 2021. Mitochondrial DNA diversity and origin of indigenous pigs in South
699 China and their contribution to western modern pig breeds (vol 18, pg 2338, 2019).
700 *JOURNAL OF INTEGRATIVE AGRICULTURE*, **20**(7): VI-VI.
701 Wang X, Xu R, Tong X, et al. 2022. Characterization of different meat flavor
702 compounds in Guangdong small-ear spotted and Yorkshire pork using two-dimensional
703 gas chromatography–time-of-flight mass spectrometry and multi-omics. *Lwt*, **169**.
704 Yang J, Do-Umehara HC, Zhang Q, et al. 2021. miR-221-5p-Mediated Downregulation
705 of JNK2 Aggravates Acute Lung Injury. *Front Immunol*, **12**: 700933.
706 Zhang ZK, Li J, Guan D, et al. 2018. A newly identified lncRNA MAR1 acts as a miR-
707 487b sponge to promote skeletal muscle differentiation and regeneration. *J Cachexia*
708 *Sarcopenia Muscle*, **9**(3): 613-626.
709 Zhao X, Mo D, Li A, et al. 2011. Comparative analyses by sequencing of transcriptomes
710 during skeletal muscle development between pig breeds differing in muscle growth rate
711 and fatness. *PLoS One*, **6**(5): e19774.

712 Zhao Y, Li J, Liu H, et al. 2015. Dynamic transcriptome profiles of skeletal muscle
713 tissue across 11 developmental stages for both Tongcheng and Yorkshire pigs. *BMC*
714 *Genomics*, **16**(1): 377.
715 Zhu Q, Liang F, Cai S, et al. 2021. KDM4A regulates myogenesis by demethylating
716 H3K9me3 of myogenic regulatory factors. *Cell Death & Disease*, **12**(6): 514.

Figure legends

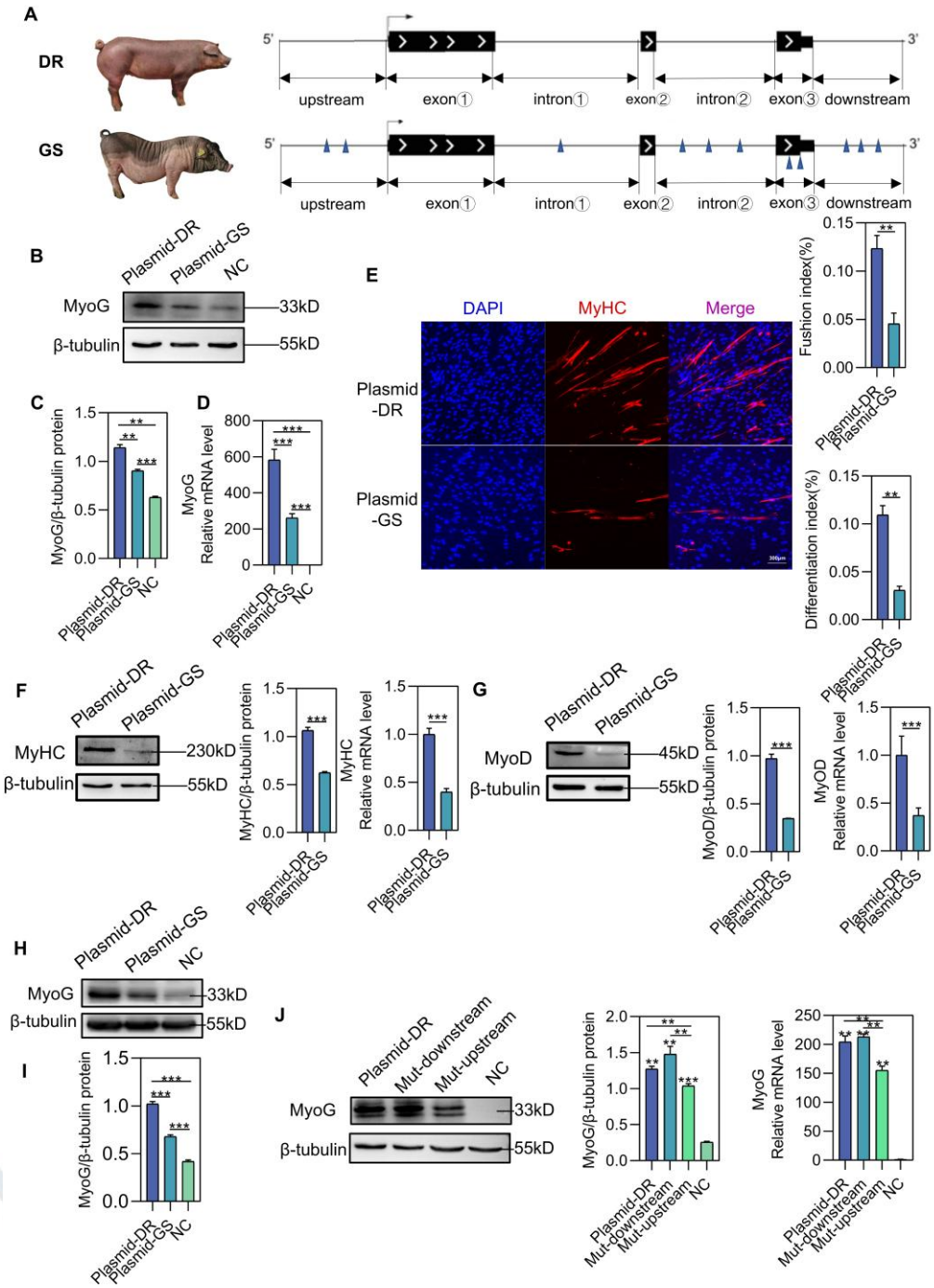
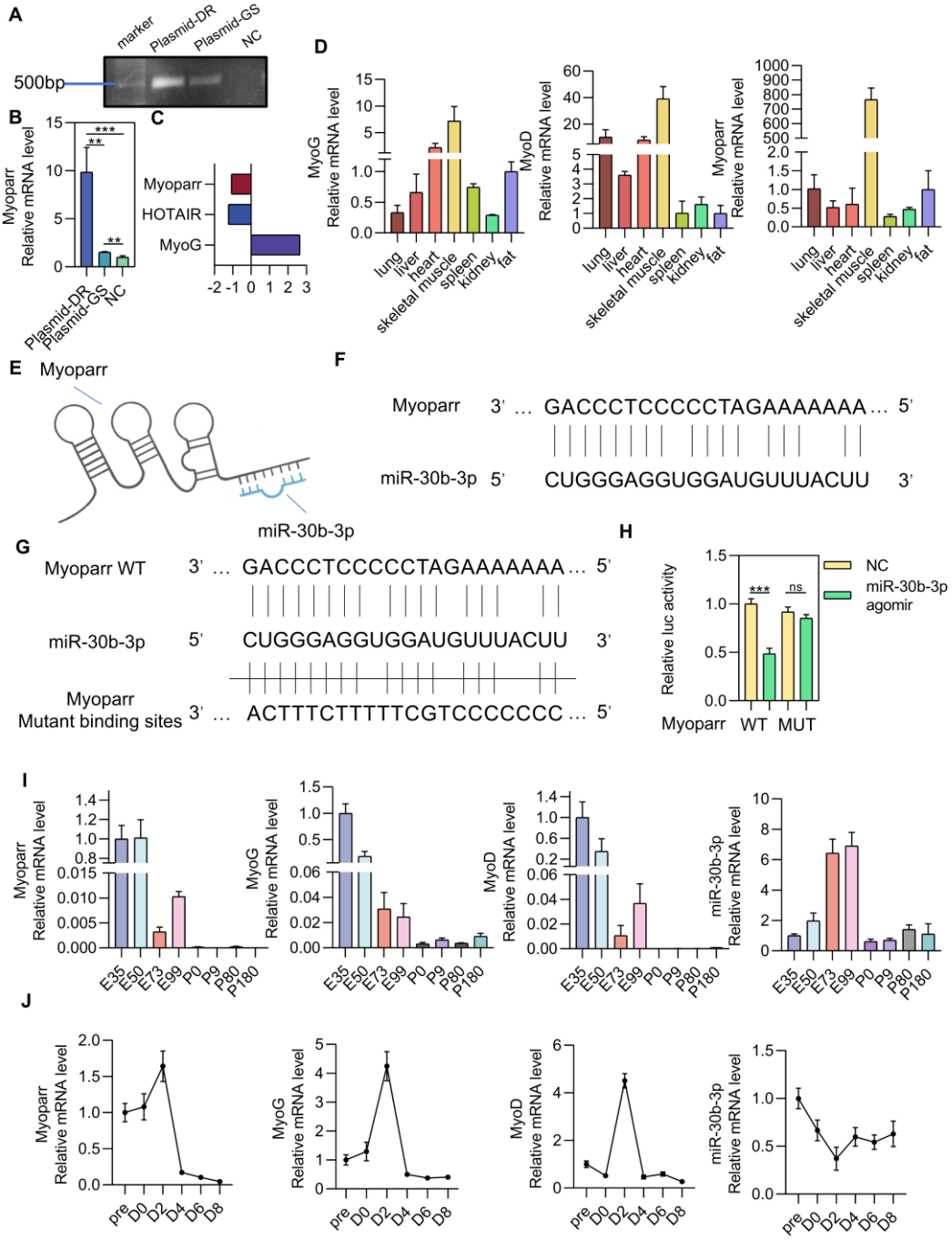


Figure 1. The Expression of Different MyoG Vectors

(A) Construction of two kinds of MyoG vectors including selected 11 SNPs that maintaining stable difference between the two pig breeds. The triangular arrows indicate SNP mutation sites. (B-C) Western blot analysis of MyoG overexpression in porcine primary myoblasts transfected with

control pcDNA3.1 or two kinds of MyoG vectors from Figure 1A after 48 h (n=3). (D) QPCR analysis of MyoG overexpression in porcine primary myoblasts transfected with control pcDNA3.1 or two kinds of MyoG vectors after 48 h (n=3). (E) Representative photographs of MyHC immunofluorescence staining in porcine primary myoblasts differentiated for 6 days after overexpression of two kinds of MyoG vectors. Statistical results of myotube fusion index (n=3). Myotube fusion index was determined as the distribution of nucleus number in total myotubes. Statistical results of differentiation index. The differentiation index was calculated as the percentage of MyHC-positive nuclei among total nuclei(n=3). (F-G) Western blot and qPCR analysis for myogenic marker genes expression from differentiated porcine primary myoblasts for 6 days after overexpression of two kinds of MyoG vectors. (H-I) Western blot analysis of MyoG overexpression in 293T cells transfected with control pcDNA3.1 or two kinds of MyoG vectors from Figure 1A after 48 h (n=3). (J) Western blot and qPCR analysis for MyoG expression from 293T 2 cells transfected with control pcDNA3.1 or 3 kinds of MyoG vectors from Figure S2K after 48 h (n=3). Data are represented as mean ± SD. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$ (Student's *t* test).



737

738 **Figure 2. The validation of lncRNA Myoparr as a ceRNA sponging miR-30b-3p**

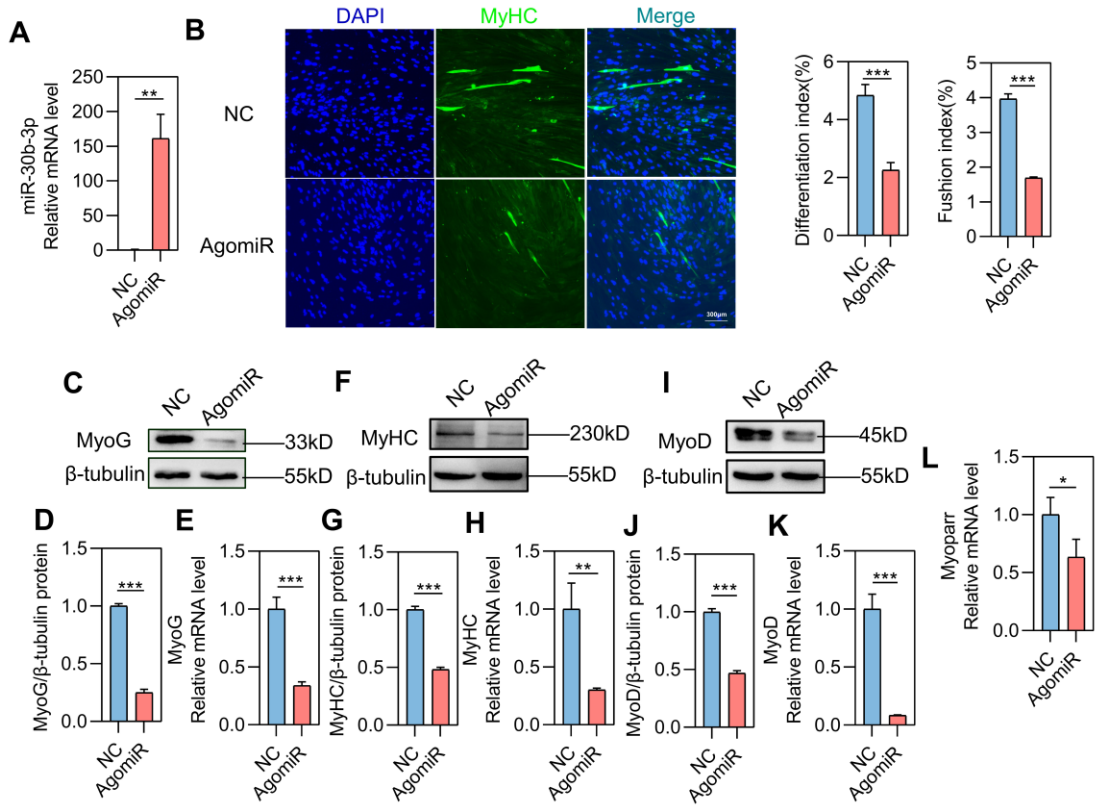
739 (A) C2C12 cells were transfected with two kinds of MyoG vectors, pcDNA3.1 vector acted as a

740 control. Then after 48 h PCR validated the 500bp sequence expressed by Myoparr from the extracted

741 RNA. (B) QPCR analysis of Myoparr expression in porcine primary myoblasts transfected with

742 control pcDNA3.1 or two kinds of MyoG vectors from Figure 1A after 48 h (n=3). (C) In silico

analysis of the coding potential of Myoparr comparing with coding genes MyoG and the noncoding control HOTAIR by using CPC. (D) Expression pattern of MyoG, MyoD and Myoparr in various tissues in 180-day-old pigs was detected through qPCR (n=3). (E) Sketch on potential miR-30b-3p binding site in Myoparr. (F) Details of the alignment of miR-30b-3p to the binding site of Myoparr. (G) Schematic of the double luciferase assay vector with upper wildtype (WT), and lower mutant (Mut) Myoparr binding sites. (H) AgomiR-30b-3p or NC were co-transfected with psiCHECK2-Myoparr WT or psiCHECK2-Myoparr Mut vectors into 293T cells. After 48 h, Dual-Luciferase reporter assay was quantified and normalized (n=4). (I) QPCR detection of MyoD, MyoG, Myoparr and miR-30b-3p expression during pigs' skeletal muscle development. (J) QPCR detection of MyoD, MyoG, Myoparr and miR-30b-3p expression during C2C12 cell differentiation. Data are represented as mean ± SD. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (Student's *t* test).



755

756 **Figure 3. Overexpression of miR-30b-3p impairs myogenic differentiation in porcine**

757 **primary myoblasts**

758 (A) Overexpression efficiency of agomiR-30b-3p targeting miR-30b-3p was detected in porcine

759 primary myoblasts through qPCR (n=3). (B) Immunofluorescent staining for MyHC in NC or

760 agomiR-30b-3p porcine primary myoblasts was performed to detect myotube formation. Statistical

761 results of differentiation index. The differentiation index was calculated as the percentage of MyHC-

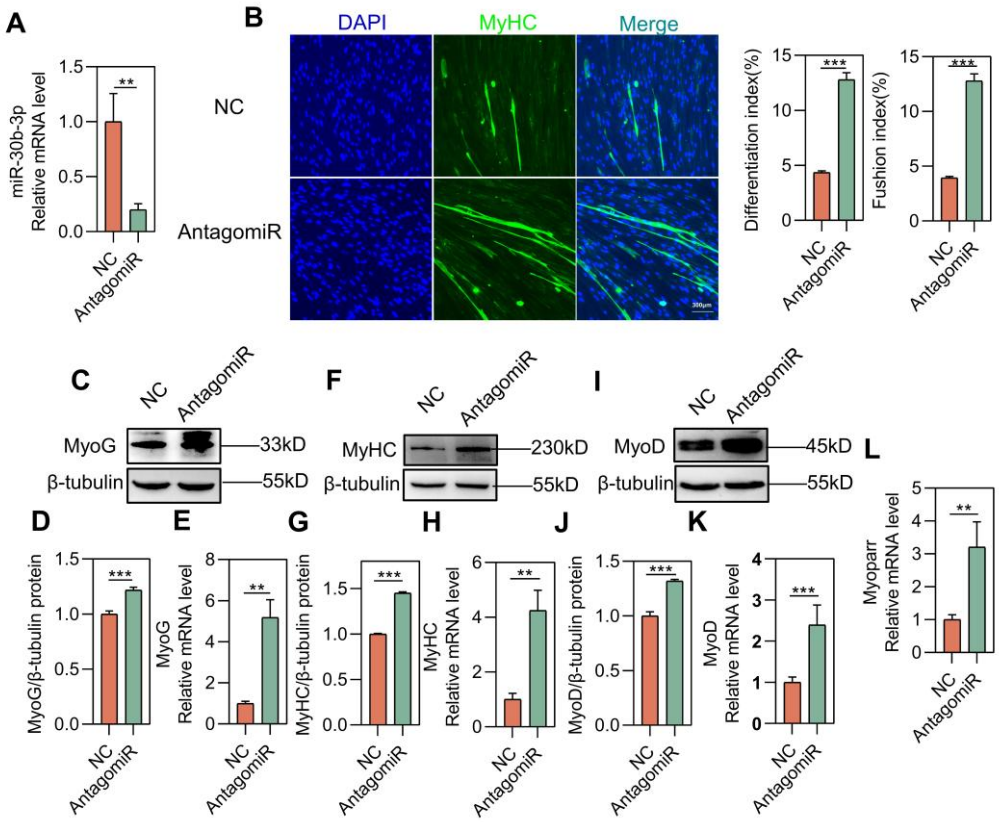
762 positive nuclei among total nuclei(n=3). Statistical results of myotube fusion index (n=3). Myotube

763 fusion index was determined as the distribution of nucleus number in total myotubes. (C-L) Western

764 blot and qPCR analysis for myogenic marker genes expression from differentiated porcine primary

765 myoblasts for 6 days after transfected with NC or agomiR-30b-3p. Data are represented as mean ±

766 SD. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (Student's *t* test).



767

768 **Figure 4. miR-30b-3p deficiency facilitates myogenic differentiation in porcine primary**

769 **myoblasts**

770 (A) Knockdown efficiency of antagomiR-30b-3p targeting miR-30b-3p was detected in porcine

771 primary myoblasts through qPCR (n=3). (B) Immunofluorescent staining for MyHC in NC or

772 antagomiR-30b-3p porcine primary myoblasts was performed to detect myotube formation.

773 Statistical results of differentiation index. The differentiation index was calculated as the percentage

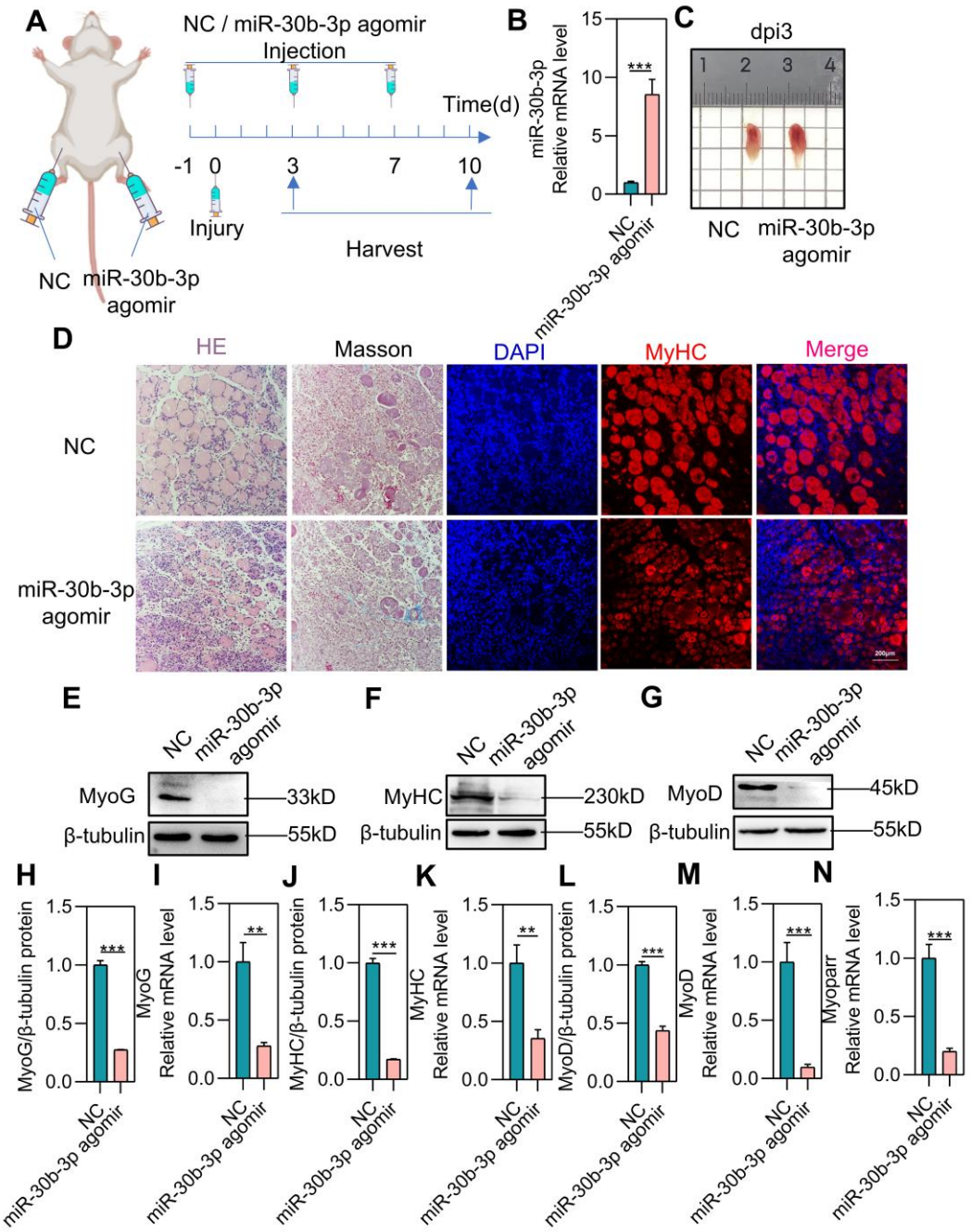
774 of MyHC-positive nuclei among total nuclei(n=3). Statistical results of myotube fusion index (n=3).

775 Myotube fusion index was determined as the distribution of nucleus number in total myotubes. (C-

776 L) Western blot and qPCR analysis for myogenic marker genes expression from differentiated

777 porcine primary myoblasts for 6 days after transfected with NC or antagomiR-30b-3p. Data are

778 represented as mean ± SD. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (Student's *t* test).



779

780 **Figure 5. Elevated miR-30b-3p delays skeletal muscle regeneration (3dpi)**

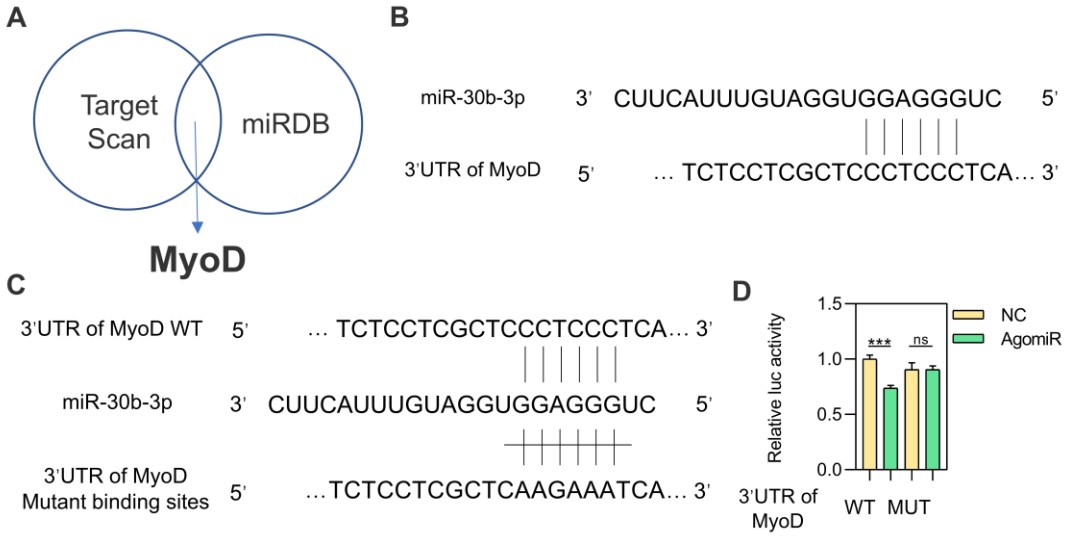
781 (A) A schematic outlining the experimental protocol followed to analyze muscle regeneration. CTX

782 was injected into the TA muscles and when the TA muscles were treated with CTX, it was defined

783 as day 0. The left and right TA muscles were treated with NC and agomiR-30b-3p respectively, then

784 harvested and analyzed at 3 days and 10 days post-injury (dpi). The number of mice in each group

is 5 (n=5). (B) Overexpression efficiency of agomiR-30b-3p targeting miR-30b-3p was detected from the extracted RNA of TA through qPCR after 3 days injury (n=3). (C) Representative images of regenerating TA muscles treated with NC and agomiR-30b-3p at day 3. (D) Immunofluorescence staining of MyHC, H&E staining and Masson trichrome staining were performed on cross sections of TA muscles after 3 days injury. (E-N) Western blot and qPCR detection of myogenic genes expression in the regenerating TA muscles at day 3 (n=3). Data are represented as mean ± SD. *P < 0.05; **P < 0.01; ***P < 0.001 (Student's t test).



799

800 **Figure 6. The Validation of miR-30b-3p Targeting MyoD 3'UTR**

801 (A) The binding target of miR-30b-3p predicted by TargetScan and miRDB was intersected. MyoD

802 was selected for closely related to the myogenesis process. (B) Details of the alignment of miR-30b-

803 3p to the binding site of MyoD 3'UTR. (C) Schematic of the double luciferase assay vector with

804 upper wildtype (WT), and lower mutant (Mut) MyoD 3'UTR binding sites. (D) AgomiR-30b-3p or

805 NC were co-transfected with psiCHECK2-MyoD 3'UTR WT or psiCHECK2- MyoD 3'UTR Mut

806 vectors into 293T cells. After 48 h, Dual-Luciferase reporter assay was quantified and normalized

807 (n=4). Data are represented as mean ± SD. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (Student's *t* test).

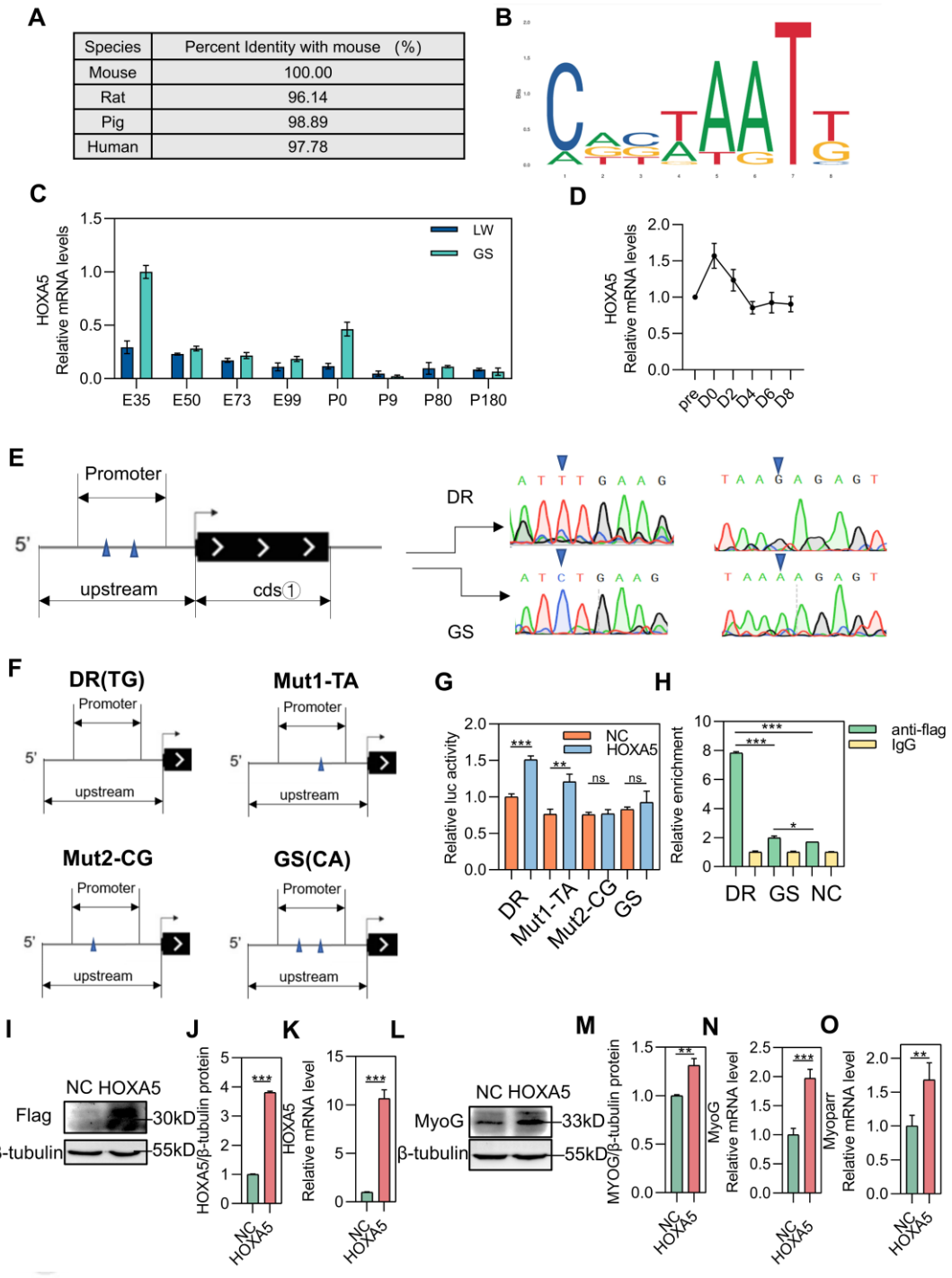
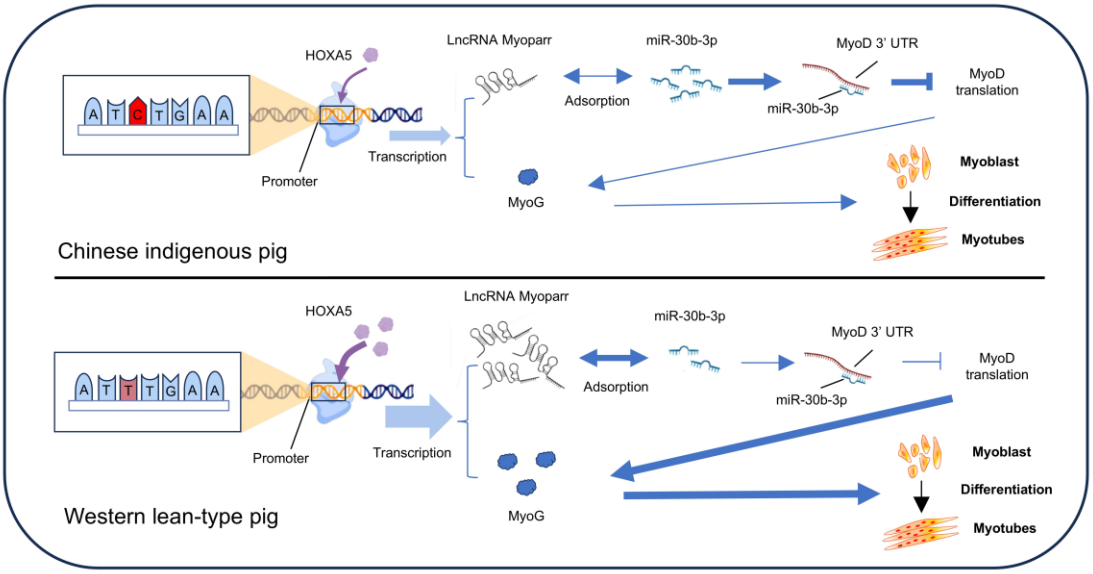


Figure 7. SNP rs3471653254 C>T regulates HOXA5 Binding and Transcriptional Activity.

(A) The sequence similarity of HOXA5 was compared across four species. (B) Schematic of the possibility on HOXA5 combination motif. (C) QPCR detection of HOXA5 expression between the large white and Guangdong small-ear spotted pigs' skeletal muscle development. (D) QPCR detection of HOXA5 expression during C2C12 cell differentiation. (E) The mutation loci

information of the two pig breeds obtained by sequencing. The triangular arrows indicate SNP mutation sites. (F) WT or two SNPs mutated in various combinations of promoter binding sites were inserted into reporter vectors. The triangular arrows indicate SNP mutation sites. (G) HOXA5 overexpressed plasmid or NC were co-transfected with four above-mentioned reporter vectors from D into 293T cells. After 48 h, Dual-Luciferase reporter assay was quantified and normalized (n=4). (H) ChIP-qPCR analysis of the binding of HOXA5 in 293T cells transfected with two kinds of promoter segment vectors and NC. (I-K) The efficiency of HOXA5 overexpression compared with that of the NC detected by Western blot and qPCR. (L-O) Western blot and qPCR detection of MyoG and Myoparr expression in C2C12 cells transfected with HOXA5 overexpression vectors. Data are represented as mean ± SD. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (Student's *t* test).



834

835 **Figure 8. SNP rs3471653254 C>T regulates myogenic differentiation**

836 Schematic diagram of the mechanism by which SNP rs3471653254 C>T modulates myogenesis.

837 SNP rs3471653254 C>T influences the enrichment of HOXA5 to regulate the transcriptional

838 expression level of MyoG and Myoparr. Myoparr can be used as the sponge of mir-30b-3p for

839 adsorption. Afterwards miR-30b-3p inhibits the expression of MyoD through interaction 3'UTR of

840 MyoD, then down-regulating the process of myoblast differentiation.