GENOME-WIDE IDENTIFICATION OF ENHANCERS, TRANSCRIPTION FACTORS, AND MECHANISMS THAT CONTROL SKELETAL MUSCLE DIFFERENTIATION

PENGCHENG LYU

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

In Animal and Poultry Sciences

Honglin Jiang, Chair
David E. Gerrard
Sally Johnson
Elizabeth Gilbert

Oral defense July 26, 2023

Blacksburg, VA

Keywords: Myogenesis, Transcription Factor, Enhancer

Copyright 2023, Pengcheng Lyu
Skeletal muscle development and growth involve significant changes in gene expression. The overall objective of this dissertation project was to identify transcription factors, enhancers, and mechanisms that control gene expression during skeletal muscle development and growth on a genome-wide scale. Three independent studies were conducted in this project. The objective of the first study was to identify potentially novel mechanisms that mediate myoblast differentiation, a process whereby the mononuclear muscle precursor cells myoblasts express skeletal muscle-specific genes and fuse with each other to form multinucleated myotubes. Comparing gene expression profiles in C2C12 cells, a widely used model of myoblasts, before and 6 days after induced myogenic differentiation by RNA sequencing (RNA-seq) revealed 11,046 differentially expressed genes, of which 5,615 and 5,431 were upregulated and downregulated, respectively. Functional enrichment analyses revealed that the upregulated genes were associated with biological processes or cellular components such as skeletal muscle contraction, autophagy, and sarcomere. In contrast, the downregulated genes were associated with biological processes or cellular components such as ribonucleoprotein complex biogenesis, mRNA processing, and ribosome. Western blot analyses showed an increased conversion of LC3-I to LC3-II protein during myoblast differentiation, further demonstrating the upregulation of autophagy during myoblast differentiation. Blocking the autophagic flux in C2C12 cells with chloroquine inhibited the expression of skeletal muscle-specific genes and the formation of myotubes, confirming a positive role of autophagy in myoblast differentiation and fusion. The aim of the second study was to identify enhancers and transcription factors
that regulate gene expression during the differentiation of bovine satellite cells, which are the myogenic precursor cells in adult skeletal muscle, into myotubes. In this study chromatin immunoprecipitation followed by sequencing (ChIP-seq) was used to identify active enhancers, i.e., genomic regions marked with histone modification H3K27ac (acetylation of lysine 27 of H3 histone protein). 19,027 and 47,669 H3K27ac-marked enhancers were identified from undifferentiated and differentiating bovine satellite cells, respectively. Of these enhancers, 5,882 and 35,723 were specific to undifferentiated and differentiating bovine satellite cells, respectively while 13,199 were shared by both undifferentiated and differentiating bovine satellite cells. Many of the H3K27ac-marked enhancers specific to differentiating bovine satellite cells were associated with muscle structure and development genes and were enriched with binding sites for MyoD, AP-1, AP-4, KLF, TEAD, and MEF2 transcription factors. Through siRNA-mediated knockdown, AP-4 was found to be essential for differentiation of bovine satellite cells into myotubes.

The objective of the third study was to identify enhancers and transcription factors that control differential gene expression in skeletal muscle between neonatal and adult cattle. First, RNA-seq was performed to compare gene expression profiles in skeletal muscle between neonatal calves and adult steers. This analysis identified 924 genes downregulated and 1,021 upregulated from calf to steer muscle. Among genes downregulated in steer muscle were myosin heavy chain3 (MYH3) and MYH8, and among genes upregulated in steer muscle were MYH7 and myoglobin. Surprisingly, many so-called adult muscle genes, such as MYH1 and MYH2, were not differentially expressed between calf and steer muscle. Gene ontology analyses showed that many genes downregulated in steer muscle are involved in protein synthesis and glycolysis and that many genes upregulated in steer muscle function in blood vessel development and immune cell activation. Next, ChIP-seq was performed to identify genomic regions marked with H3K27ac, i.e., active enhancers, in the skeletal muscle of neonatal calves and adult steers. This experiment led to the finding of 20,163 enhancers specifically active in the calf muscle, 14,909 enhancers specifically
active in the steer muscle, and 27,002 enhancers active in both the calf and steer muscle. Motif enrichment analyses revealed the enrichment of binding sites for the KLF family and TEAD family transcription factors in enhancers active specifically in the calf muscle, the enrichment of binding sites for the FOXO family and the SMAD family transcription factors in enhancers specifically active in the steer muscle, and the enrichment of binding sites for the MRF family and MEF2 family transcription factors in enhancers active in both the calf and steer muscle. These results shed light on the differences in gene expression and biology between newborn calf and adult steer skeletal muscle. These results also shed light on the enhancers and transcription factors that control these differences.
GENERAL AUDIENCE ABSTRACT

Muscle is the central part of meat. So, to improve meat yield, it is essential to know how muscle development is controlled. Muscle development, also called myogenesis, starts with muscle progenitor cells developing into myoblasts. Myoblasts then differentiate and fuse with each other to form myotubes. Myotubes undergo hypertrophy and form functional muscle fibers. During myogenesis, each step involves significant changes in gene expression. Gene expression is controlled mainly by proteins called transcription factors. The overall goal of this project was to identify transcription factors and DNA sequences bound by these factors that control gene expression during muscle development. This project consisted of three studies. In the first study, we used the RNA sequencing (RNA-seq) technique to find genes differentially expressed in myoblasts between before and after terminal differentiation. Analyzing the RNA-seq data led to the discovery that autophagy, a 'self-eating' biological process, is required for myoblast differentiation. In the second study, we used a technique called chromatin immunoprecipitation followed by sequencing (ChIP-seq) to identify genomic regions called active enhancers in differentiating bovine myoblasts. This work led to the identification of thousands of active enhancers and dozens of transcription factors binding to these genomic regions that control the differentiation of bovine myoblasts. In the third study, we combined RNA-seq and ChIP-seq to explore the genes and genomic regions controlling muscle transition from newborn calves to adult cattle. This part of the project led to the finding of thousands of genes differentially expressed and thousands of genomic regions differentially activated between newborn calf and adult steer muscle.
ACKNOWLEDGEMENTS

I want to acknowledge my advisor, Dr. Honglin Jiang, for his excellent guidance. He gave me the precious opportunity to study at Virginia Tech and learn new skills, including bioinformatics. He is an example of teaching me how to be a good researcher. He taught me how to write, and I would never be able to finish this dissertation without his revision. Besides Dr. Jiang, I sincerely thank the rest of my committee members: Dr. David Gerrard, Dr. Sally Johnson, and Dr. Elizabeth Gilbert. Special thanks go to Sasha Simpkins, who helps me with issues other than research. She gives me the courage to face the trouble during my study. Finally, I would like to thank my parents for their understanding and endless love. Their love has helped me go through hard times. They always encourage me and support me no matter where I am.
Table of Content

ABSTRACT .............................................................................................................................. II

GENERAL AUDIENCE ABSTRACT .................................................................................... V

ACKNOWLEDGEMENTS ....................................................................................................... VI

TABLE OF CONTENT ........................................................................................................ VII

LIST OF TABLES .................................................................................................................. X

LIST OF FIGURES ............................................................................................................... XII

1 CHAPTER I LITERATURE REVIEW TRANSCRIPTIONAL AND EPIGENETIC REGULATION OF GENE EXPRESSION DURING MYOGENESIS ....................... 1

ABSTRACT .......................................................................................................................... 1
INTRODUCTION ..................................................................................................................... 2

MYOGENIC REGULATORY TRANSCRIPTION FACTORS ......................................................... 3

Myogenic regulatory factor family of transcription factors ................................................. 3
E-protein family of transcriptional regulators ........................................................................ 6
ID family of transcriptional regulators .................................................................................. 7
PAX family of transcription factors ..................................................................................... 7
MEF2 family of transcription factors ................................................................................... 8
NFI family of transcription factors ...................................................................................... 9
NFAT family of transcription factors .................................................................................. 10
SMAD family of transcription factors ................................................................................ 12
KLF family of transcription factors .................................................................................... 13
TEAD family of transcription factors ................................................................................ 14
SIX family of transcription factors ..................................................................................... 15
Other transcription factors ................................................................................................. 15

MYOGENIC REGULATORY HISTONE MODIFICATIONS .................................................... 17
Histone acetylation and deacetylation in muscle development........................................18

Histone methylation and demethylation ...........................................................................21

CONCLUSION .....................................................................................................................27

REFERENCE .......................................................................................................................28

2 CHAPTER II* RNA-SEQUENCING REVEALS UPREGULATION AND A
BENEFICIAL ROLE OF AUTOPHAGY IN MYOBLAST DIFFERENTIATION AND
FUSION ............................................................................................................................43

ABSTRACT .........................................................................................................................43

INTRODUCTION ................................................................................................................44

MATERIALS AND METHODS ..........................................................................................46

RESULTS ...........................................................................................................................52

DISCUSSION ....................................................................................................................57

CONCLUSIONS ................................................................................................................59

REFERENCE .......................................................................................................................60

TABLE ...............................................................................................................................68

FIGURE .............................................................................................................................70

3 CHAPTER III* IDENTIFICATION OF ENHANCERS AND TRANSCRIPTION
FACTORS REGULATING THE MYOGENIC DIFFERENTIATION OF BOVINE
SATELLITE CELLS .............................................................................................................78

ABSTRACT .........................................................................................................................78

INTRODUCTION ................................................................................................................80

MATERIALS AND METHODS ..........................................................................................82

RESULT ...............................................................................................................................87

DISCUSSION ....................................................................................................................94

CONCLUSION ..................................................................................................................98

REFERENCE .......................................................................................................................99
4 CHAPTER IV IDENTIFICATION OF ENHANCERS AND TRANSCRIPTION FACTORS THAT CONTROL POSTNATAL SKELETAL MUSCLE DEVELOPMENT AND GROWTH IN CATTLE ................................................................. 125

ABSTRACT .............................................................................................................. 125

INTRODUCTION ....................................................................................................... 127

MATERIALS AND METHODS .................................................................................. 129

RESULTS .................................................................................................................. 132

DISCUSSION ............................................................................................................ 137

REFERENCE .......................................................................................................... 140

TABLE ....................................................................................................................... 147

FIGURE ................................................................................................................... 153
List of Tables

Table 2.1 Two examples of functional terms associated with genes upregulated during myoblast differentiation. ........................................................................................................ 68

Table 2.2 Two examples of functional terms associated with genes downregulated during myoblast differentiation. ........................................................................................................ 69

Table 3.1 Mapping summary of ChIP-seq libraries ........................................................................ 107

Table 3.2 Examples of H3K27ac-marked enhancers and associated genes in bovine satellite cells ........................................................................................................ 107

Table 3.3 Top 10 GO biological processes enriched in genes associated with H3K27ac-marked enhancers in differentiating bovine satellite cells .............................. 108

Table 3.4 Top 10 GO cellular components enriched in genes associated with H3K27ac-marked enhancers in differentiating bovine satellite cells .............................. 110

Table 3.5 Top 10 GO molecular functions enriched in genes associated with H3K27ac-marked enhancers in differentiating bovine satellite cells .............................. 111

Table 3.6 Top 30 motifs in differentiating bovine satellite cells ................................................. 112

Table 3.7 Examples of TFAP4 binding sites-containing enhancers in differentiating bovine satellite cells ........................................................................................................ 114

Table 3.8 Examples of biological processes enriched in genes associated with TFAP4 binding sites in differentiating bovine satellite cells ................................................. 116

Table 4.1 Mapping summary of ChIP-seq data ........................................................................ 147

Table 4.2 Top 10 motifs enriched in calf muscle-specific enhancers .............................................. 148

Table 4.3 Top 10 motifs enriched in steer muscle-specific enhancers .............................................. 149

Table 4.4 Top 10 motifs enriched in enhancers shared by both calf and steer muscle .................. 150

Table 4.5 Top 10 motifs enriched in enhancers associated with genes upregulated in calf muscle ........................................................................................................ 151

Table 4.6 Top 10 motifs enriched in enhancers associated with genes upregulated in steer
muscle ........................................................................................................ 152
List of Figures

Fig 1.1 Pathway of myogenesis. ................................................................. 3

Fig 1.2 MYOD1, MYOG, and MEF2 regulatory network. ............................................. 9

Fig 1.3 SMAD proteins affect myogenesis by regulating the expression of MYOD1 ..... 11

Fig 1.4 Key methylation on H3 and H4 histone proteins. ........................................ 21

Fig 2.1 Identification of differentially expressed genes during myoblast differentiation. 70

Fig 2.2 Top 10 biological processes, cellular components, and molecular functions associated with genes upregulated (A) and downregulated (B) during myoblast differentiation. ........................................................................................................... 71

Fig 2.3 Increased generation and accumulation of LC3-II in C2C12 cells during myogenic differentiation. ................................................................................................................................. 72

Fig 2.4 Effects of chloroquine on LC3-I and LC3-II protein expression in C2C12 cells during myogenic differentiation. ............................................................................................................................. 73

Fig 2.5 Effects of inhibiting autophagy with chloroquine on the morphological changes in C2C12 cells during myoblast differentiation. ..................................................................................................... 74

Fig 2.6 Effects of inhibiting autophagy with chloroquine on the expression of selected muscle-specific genes in C2C12 cells during myogenic differentiation. ......................... 76

Fig 3.1 Identification of H3K27ac-marked enhancers in before-differentiation and during-differentiation bovine satellite cells. ................................................................. 117

Fig 3.2 Examples of H3K27ac-marked enhancers in bovine satellite cells. ................ 118

Fig 3.3 Genomic distribution and location of H3K27ac-marked enhancers in bovine satellite cells. .......................................................................................................................... 119
Fig 3.4 Identification of TFAP4 binding site in H3K27ac-marked enhancers in differentiating bovine satellite cells ................................................................. 120

Fig 3.5 Effect of siRNA-mediated knockdown of TFAP4 on differentiation of bovine satellite cells ................................................................. 121

Fig 3.6 Effect of TFAP4 knockdown on gene expression in bovine satellite cells. .... 122

Fig 3.7 Effect of TFAP4 overexpression on differentiation of bovine satellite cells. .... 123

Fig 3.8 Effect of TFAP4 overexpression on protein expression of TFAP4, MYH, CKM, and MYOG in bovine satellite cells ....................................................... 124

Fig 4.1 Flowchart of experimental design and data analysis .................................. 153

Fig 4.2 RNA-seq analysis of transcriptomes in calf and steer muscle .................. 154

Fig 4.3 Top 20 biological processes enriched in genes upregulated in calf or steer muscle ............................................................. 155

Fig 4.4 Identification of H3K27ac-marked enhancers in calf and steer muscle by ChIP-seq ....................................................................................................... 156

Fig 4.5 Genomic distribution of calf muscle-specific, steer muscle-specific and common H3K27ac-marked enhancers ............................................................... 157

Fig 4.6 Examples of H3K27ac-marked enhancers identified from calf or steer muscle or both calf and steer muscle ...................................................... 158
Chapter I

LITERATURE REVIEW

Transcriptional and Epigenetic Regulation of Gene Expression during Myogenesis

ABSTRACT

Myogenesis is a complex process that involves significant changes in gene expression. Transcription factors are essential in determining gene expression level, spatiotemporal specificity, and phenotypes. Myogenic regulatory factors (MRFs) are well-studied transcription factors controlling myogenesis and include four members, myogenic factor 5 (MYF5), myogenic differentiation 1 (MYOD1), myogenin (MYOG), and MYF6. Recent studies demonstrate that gene expression during myogenesis is controlled by many other transcription factors, such as myocyte enhancer factor-2 (MEF2), serum response factor (SRF), and nuclear factor of activated T cells (NFAT). Recent studies also indicate that histone H3 modifications directly or indirectly regulate gene transcription in muscle cells. Acetylation of lysine 27 of H3 histone protein (H3K27ac) and di- and tri-methylation of lysine 4 of H3 histone protein (H3K4me2 and H3K4me3) are believed to promote the expression of myogenic genes. In contrast, tri-methylation of lysine 9 and tri-methylation of lysine 27 of H3 protein are thought to inhibit the expression of myogenic genes. Additionally, mono-methylation of lysine 4 of H3 histone protein is associated with both the promotion and repression of myogenic genes, depending on the developmental stage of muscle cells.

Keywords: myogenesis, transcription factor, histone modification
INTRODUCTION

Skeletal muscle, comprising approximately 40% of the body's mass [1], is the largest tissue in the body. The fundamental function of skeletal muscle is contraction. Skeletal muscle tissue is primarily made up of multinucleated cells called muscle fibers or myofibers. Myofibers can be classified into two categories based on their energy consumption patterns: fast fibers, also known as glycolytic fibers, and slow fibers, also known as oxidative fibers. Each fiber type possesses a unique set of structural proteins. Myosin heavy chain (MYH) is a major structural protein of muscle fibers. While MYH1 and MYH2 proteins are the major MYH proteins of fast fibers, MYH7 is the major MYH protein of slow fibers.

Muscle fibers are formed from their progenitor cells through a process called myogenesis. Myogenesis begins in the embryo. During embryonic myogenesis, muscle progenitor cells derived from the dermomyotome first commit to the myoblast lineage, followed by terminal differentiation, in which myoblasts migrate and align in a linear order and subsequently fuse to form multinucleated myotubes, the immature form of myofibers (Fig. 1.1). During embryonic myogenesis, a group of muscle progenitor cells retain stem cell-like features, migrate to the space between the sarcolemma of muscle fibers and the basal lamina, and become satellite cells [2]. These satellite cells are the progenitor cells for muscle regeneration and muscle hypertrophy in postnatal animals (Fig. 1.1). [3, 4]

Myogenesis is a complex process that involves significant changes in gene expression. Transcription factors (TFs), working collaboratively with epigenetic modifiers, are the major regulators of gene expression. The purpose of this review is to examine the literature detailing the function, expression, and regulation of key transcriptional and epigenetic regulators of myogenesis.
Fig 1.1 Pathway of myogenesis. Muscle progenitor cells originated from dermomyotome commit into myoblasts, and myoblasts migrate, proliferate, differentiate, and fuse, forming myotubes. Some muscle progenitor cells form satellite cells, which participate in muscle hypertrophy and regeneration during postnatal life.

Myogenic Regulatory Transcription Factors

Myogenic regulatory factor family of transcription factors

The myogenic regulatory factor (MRF) family of transcription factors, comprising myogenic factor 5 (MYF5), myogenic differentiation 1 (MYOD1), MRF4, and myogenin (MYOG), are considered the master transcription factors of myogenesis. Among them, MYF5 and MYOD1 are believed to be more critical for gene expression at the earlier stages of myogenesis. At the same time, MYOG serves as the essential factor for terminal differentiation of myoblasts and formation of myotubes. Transfecting MYOD1 into non-myogenic cells can induce the conversion of these cells into myoblast-like cells [5]. All MRFs are basic helix-loop-helix (bHLH) transcription factors. They exert their transcriptional function by binding to the E-box (CANNTG) on many myogenic genes, usually in combination with E-proteins, to form a functional heterodimer.
MYOD1

MYOD1, also known as MYF3, is highly conserved across species. During myogenesis, MYOD1 protein forms a heterodimer with an E-protein (including E12, E47, and HEB), and this heterodimer then binds to the E-box regulatory elements surrounding many myogenic genes [6]. MYOD1 is required for muscle precursor cells to commit to the myoblast lineage during embryonic myogenesis [7]. MYOD1 stops the cell cycle and facilitates terminal differentiation by activating cyclin-dependent kinase inhibitor 1A (p21 or CDKN1A) [8, 9]. Traditionally, MYOD1 is believed to bind to the regulatory elements flanking the target genes, including promoter or enhancer regions, to initiate or increase muscle-specific gene expression. Many myogenic genes contain MYOD1 binding sites, for example, MYOD1 itself, MYOG [10-12], creatine kinase, M-type (CKM) [13, 14], MYF5, and MRF4 [15, 16]. MYOD1 can also serve as a transcriptional guide by recruiting other transcription factors to its target genes. MYOD1 can also guide histone modifiers to its target genes, resulting in histone modification addition or removal and thereby changing the accessibility of chromatin to other transcriptional regulators [17]. MYOD1, a core transcriptional regulator of myogenic gene expression, is regulated by other upstream transcription factors. One of these factors is coflin2 (CFL2), which regulates the expression of not only MYOD1 but also MYOG and myocyte enhancer factor 2C (MEF2C). Such regulation by CFL2 plays a crucial role in governing the developmental fate, proliferation, and differentiation of myoblasts [18]. Transducin-like enhancer of split 3 (TLE3) inhibits myogenic differentiation by directly binding to MYOD1, disrupting the interaction of MYOD1 with E-proteins, and decreasing the transcriptional activity of MYOD1 and, subsequently, the expression level of MYOD1-targeted genes [19].

MYF5

Among all four MRFs, MYF5 is the first myogenic transcription factor expressed during
myogenesis [20]. MYF5 was first found in 1987 [5]. At the early developmental stage, the expression or activity of MYF5 is regulated by the Sonic hedgehog pathway [21] and the Wnt pathway [22, 23]. As a critical transcription factor regulating myogenic gene expression, MYF5 is also controlled by other transcription factors, for instance, the transcription factors of the paired box family (PAX), the eyes absent family (EYA), and the SIX homeobox (SIX) family [24]. In addition to its role in transcriptional regulation, MYF5 controls myogenesis at the translational level by binding to specific mRNAs. Research has shown that MYF5 binds to the mRNA of cyclin D1 (CCND1) and facilitates the initiation of cyclin D1 translation [25].

**MYOG**

Myogenin (MYOG) was found in 1989 as a homolog of MYOD1 [26]. Although the other three members of the MRF family have redundant functions in the early stages of myogenesis, myogenin has a unique role in terminal differentiation that MYOD1, MYF5, or MRF4 cannot replace. Mice with the deletion of MYOG died at birth due to muscle development failure [27]. Myogenin is required to form the transcriptional initiation complex on muscle-specific genes [28]. MYOG targets many genes involved in myogenesis, including genes encoding muscle structural proteins, such as myosin heavy chain (MHC) and muscle creatine kinase (CKM). Significantly, myoblast fusion as an essential process during terminal differentiation to form multinucleated myotubes is controlled by MYOG via directly enhancing the transcription of two target genes, myomaker (MYMK or TMEM18C1 in bovine) and myomixer (MYMX), two membrane proteins required for myoblast fusion [29]. MYOG does not merely direct terminal differentiation, but also controls the shift of slow muscle fiber toward the fast type. Mice with myogenin deletion in adulthood exhibited better duration of exercise due to the increased number of slow, fatigue-resistant muscle fibers [30]. Many other transcription factors control MYOG expression itself. As a MYOG regulator, MYOD1 can principally
increase its expression [31]. Moreover, MEF2 transcription factors can bind to enhancer regions of MYOG to increase its expression, which in turn promotes the expression of MEF2s, forming positive feedback, an important mechanism at the early stage of myoblast differentiation [32].

**MRF4**

MRF4 is the last expressed MRF among all four members, which participates in the maturation of myofibers [33]. The other three MRFs (MYF5, MYOD1, and MYOG) can all enhance the expression of MRF4 by binding to the E-box in the promoter region [34]. Besides the MRFs, transcription factors of the MEF2 family also bind to the promoter region of MRF4 to increase its expression [34]. By transfecting MRF4 into MYOG-null myoblasts, MRF4 was discovered to partly overcome the myogenesis impairment brought by MYOG depletion [35]. Recently, MRF4 has been reported to exert a repressive effect on adult muscle hypertrophy via interaction with MEF2 factors and recruitment of histone deacetylases (HDAC) 3 and 4 to myogenic genes [36]. The absence of MRF4 leads to enhanced muscle mass gain, as it disrupts the repression complex comprised of MEF2C and HDAC, which generally acts on the regulatory region of the myogenic genes [36].

**E-protein family of transcriptional regulators**

The E-protein family contains E12 (TCF3), E47 (TCF4) and HEB (TCF12). E-proteins combine with bHLH transcription factors to form heterodimers, which exert regulatory functions by directly altering target gene expression or recruiting other transcription factors. Because the interaction between E-proteins and MRFs is necessary for myoblast differentiation, many myoblast differentiation inhibitors inhibit myoblast differentiation by either hampering the formation of an MRF-E-protein complex or preventing the interaction of this complex with chromatin. TCF12 is essential for embryonic muscle development, particularly during the early stages of terminal differentiation [37]. TCF12 increases
muscle terminal differentiation by enhancing MYOG and MYOD1 expression, facilitating the transcription of TCF12 itself, which builds a positive cycle that boosts muscle differentiation [38]. TCF12-MYOD1 guides histone modifiers including the methyltransferase, EZH2, and the demethylase, UTX, to cis-elements that alter chromatin accessibility and stabilize the transcription of target genes [39].

**ID family of transcriptional regulators**

The inhibitor of the DNA-binding/differentiation (ID) family has four members, ID1, ID2, ID3, and ID4. The proteins in this family inhibit myoblast differentiation by prohibiting the interaction of the E-protein or E-protein-MYOD1 complex with cis-regulatory DNA elements [40, 41]. Although binding weakly to MYOD1[42], ID proteins possess a strong affinity to the E-protein, which ultimately inhibits the binding between MRFs and E-proteins [40]. C2C12 cells with ectopic expression of an ID protein stopped entering into differentiation compared to the cells without ID overexpression [40]. During myoblast differentiation, ID mRNA or protein levels drop dramatically [43], partly resulting from the withdrawal of serum or growth factors [44, 45]. Besides the E-proteins, ID protein also inhibits the DNA binding ability of other bHLH transcription factors, such as TFAP4 [46].

**PAX family of transcription factors**

Two members of the Paired Box (PAX) family of transcription factors, PAX3 and PAX7, play significant roles in myogenesis. Exogenous expression of PAX3 alone is enough to activate muscle progenitor cells into later myogenesis [47], and PAX3 does that by directly regulating the expression of MYF5 and MYOD1 [48]. PAX7 is required for satellite cell formation [49, 50]. PAX7-null mice showed impaired muscle regeneration due to the loss of satellite cells [49, 51]. Overexpression of PAX7 in myoblasts caused downregulation of MYOD1 [52, 53] and prevented ectopic MYOD1 from initiating the myogenic program in non-muscle cells [5, 54]. The differentiation-repressing effect of PAX7 keeps the cells with
proliferation capability, which further maintains the number of quiescent satellite cells for muscle regeneration in adulthood. PAX7 downregulates MYOD1 expression by directly increasing the expression of ID2 and ID3 genes [55].

**MEF2 family of transcription factors**

The MEF2 family of transcription factors (MEF2A, B, C and D) plays an essential role in maintaining normal myogenesis by acting as transcriptional enhancers. Unlike the MRFs, which can work alone, MEF2 proteins regulate muscle development by cooperating with the MRFs [56]. Expression of the MEF2 transcription factors varies with developmental stages [57, 58]. In C2C12 cells, MEF2A, B, and C are only present in the differentiated myotubes, but MEF2D is expressed both in myoblasts and myotubes with greater expression in undifferentiated myoblasts [59]. Due to the similar DNA-binding domain, MEF2 proteins share similar binding sites on the chromatin; as such, it is challenging to dissect their unique regulatory functions during myogenesis [57]. MEF2A ablation is associated with loss of muscle regeneration ability and a reduction in the numbers of PAX7+ satellite cells, indicating PAX7 as a probable target of MEF2A [60]. Depletion of MEF2B, C, or D or any combination does not cause an aberrant phenotype [61]. Postnatal deletion of MEF2C causes hypoplastic myofibers with unorganized M lines, due to reduced expression of myomesin [62].

The MEF2 family also serves as a regulatory platform in muscle atrophy. MEF2C is the target of SMAD2 and 3, which mediate the effect of the TGF-β family of growth factors on muscle atrophy [63]. In cancer-induced cachexia, increased SMAD2 and 3 inhibit the expression of MEF2C and thereby reduce MEF2C-regulated slow fiber expression, for example, MYH7 expression [64]. MEF2C also affects the expression of PGC-1A, another essential node governing slow fiber formation [65]. The phosphorylation status is a crucial factor affecting the function of MEF2C in skeletal muscle differentiation. Skeletal myosin
light chain kinase (MLCK) phosphorylates MEF2C at the T80 position, and this phosphorylation facilitates the recruitment of p300/CBP, a histone acetyltransferase complex, to the enhancer regions of MEF2-targeted genes [66]. Acetylation is another post-translational modification that affects the activity of MEF2 proteins. HDAC3 represses the transcriptional activity of MEF2 proteins via removing the acetyl group adjacent to the DNA binding domain, resulting in the declined efficiency of MEF2s binding to their target chromatin [67].

**Fig 1.2** MYOD1, MYOG, and MEF2 regulatory network.

**NFI family of transcription factors**

Nuclear factor I (NFI) family of transcription factors contains NFIA, NFIB, NFIC, and NFIX [68]. NFIX inhibits the expression of slow MHC genes by repressing the expression of NFATC4, a positive regulator of slow MHC genes [69]. Adult mouse skeletal muscle with decreased NFIX expression exhibits smaller myofibers indicating the positive role of NFIX in myoblast differentiation and muscle hypertrophy during adulthood [70]. Interestingly, NFIX is the only NFI protein that binds to MEF2A to form a protein dimer, which is required to activate CKM transcription [69]. Another NFI family member, NFIA,
represses myogenesis by competitively interacting with KLF5 at the MYOD1 enhancer, reducing MYOD1 expression and thereby down-regulating MYOG expression [71].

**NFAT family of transcription factors**

Nuclear factors of activated T-cells (NFAT) family of transcription factors include NFATC1, NFATC2, NFATC3, NFATC4, and NFAT5 [72]. NFATC1 and NFATC4 affect myofiber size and the number of myoblasts [73]. NFATC1 drives the formation of slow muscle fibers [74, 75]. NFATC1 enhances the accumulation of slow fibers by repressing MYOD1 from interacting with p300 histone acetyltransferase, thereby hampering the recruitment of other transcription factors and the activation of enhancers found within fast fiber-type genes. NFATC1 is required for fast-to-slow fiber transformation but also acts synergistically with MYOD1 [76]. The combination of NFATC1 and MYOD1 acts upstream of the target gene MYH7 and recruits p300, MEF2D, and other transcription factors or co-factors to increase MYH7 transcription [76]. Additionally, NFATC1 regulates myoblast fusion by upregulating the expression of a membrane protein, stabilin 2 (STAB2), which plays a vital role in myoblast fusion and differentiation. [77]. NFATC2 is a transcription factor closely associated with myogenesis and directly interacts with MYOD1 [78]. During the postnatal stage, the collaboration between NFATC2 and MYOD1 influences the expression of MYH8, the neonatal form of MHC [78]. Unlike NFATC1 and NFATC3, which prefer binding to embryonic MHC genes, NFATC2 exclusively binds to the neonatal MHC. This unique binding pattern suggests that NFATC2 plays a distinct role in regulating neonatal MHC expression compared to other NFAT family members [78]. NFATC2-null mice have a significant impairment in muscle fiber size and a decrease in nuclei within each fiber [78]. These observations suggest that NFATC2 regulates genes associated with myoblast fusion. Furthermore, NFATC2 participates in secondary myogenesis [78]. During the early stages of myoblast differentiation, NFATC3 translocates to the nucleus before NFATC1 and NFATC2 [79]. Notably, NFATC3 has a more
prominent role in enhancing myoblast differentiation than NFATC1 and NFATC2 [79]. However, it is important to note that NFATC3 does not contribute to the modulation of slow or fast fiber expression balance, suggesting that its function is specific to other aspects of myoblast differentiation [79]. NFAT5 plays a significant role in myoblast migration [80]. Additionally, NFAT5 facilitates cell-cell contact, a crucial requirement for proper myoblast differentiation. The involvement of NFAT5 in these processes highlights its importance in orchestrating the initial stages of myoblast differentiation and promoting the necessary cellular interactions for successful myofiber development [80]. Other transcription factors or non-nuclear proteins can regulate myogenesis by affecting NFAT expression or activity. Prox1 is a transcription factor that activates the NFAT pathway [81]. SHP-2 (cytoplasmic SH2 domain-containing protein tyrosine phosphatase) is another protein regulating NFAT activity through an unclear mechanism. Decreased SHP-2 expression impairs NFAT expression and causes the formation of smaller myotubes, a similar result as the deletion of NFAT genes [82]. Taken together, the NFAT transcription factors participate in slow fiber expression, myoblast differentiation, and myoblast fusion.

**Fig 1.3** SMAD proteins affect myogenesis by regulating the expression of MYOD1. TGF-
β or other extracellular stimuli like myostatin (MSTN) activate their receptors, causing the autophosphorylation and phosphorylation of SMAD2 and SMAD3. Meanwhile, SMAD4 interacts with SMAD2 or SMAD3 and leads the SMAD2-SMAD3 dimer into the nucleus to inhibit the transcription of MYOD1.

**SMAD family of transcription factors**

The proteins in the SMAD family play crucial roles in the TGF-β superfamily signaling [83]. Many growth factors of the TGF-β superfamily negatively affect the terminal differentiation of myoblasts [84, 85]. The SMAD proteins can be separated into three categories: R-SMAD (SMAD1-3,5), co-SMAD (SMAD4), and I-SMAD (SMAD6,7). Upon stimulation by extracellular TGF-β, R-SMAD is phosphorylated by the type II TGF receptor allowing it to form a heterodimer with SMAD4, which can translocate to the nucleus and alter gene transcription. During muscle differentiation, the expression of TGF-β decreases, coinciding with the reduction of SMAD 2, 3, and 4. Due to reduced TGF-β activation, the phosphorylation level of the SMAD proteins also decreases during muscle differentiation [86]. SMAD2 negatively affects the expression of MYOG and MYOD1 [87]. siRNA knockdown of SMAD2 or SMAD3 increases MYOG expression [86]. However, it is worth pointing out that there is research suggesting that SMAD2 enhances myoblast differentiation independent of the TGF-β signaling pathway. Overexpression of SMAD2 in C2C12 cells increased myotube size and fusion index, while loss of SMAD2 reduced myotube size [88]. SMAD2 regulates C2C12 differentiation partly by promoting KLF4 transcription [88]. Unlike SMAD2, SMAD3 can interact with the MYOD1 bHLH domain, blocking the dimerization of the latter with E-proteins [89]. Furthermore, SMAD3 hampers the interaction between MYOD1 and MEF2 [90]. Interestingly, although phosphorylated SMAD2 and 3 have adverse effects on muscle differentiation, they can recruit the p300 histone acetyltransferase to specifically add acetyl groups to H3K9 and H3K18, thereby opening chromatin and activating gene expression [91]. SMAD2 and SMAD3 play roles...
in adulthood muscle atrophy and autophagy without the involvement of the mTOR pathway [63]. SMAD4 acts as a negative regulator of myoblast differentiation. Disruption of SMAD4 enhances myotube formation [92, 93]. Meanwhile, SMAD4 is an essential factor associated with muscle regeneration. ID1 is considered a target gene of SMAD4, and it can interfere with the myogenic function of MYOG. When SMAD4 is decreased, it leads to the downregulation of ID1, consequently rescuing the function of MYOG and driving the cells into the differentiation process [94]. During aging, the expression of SMAD4 decreases in mice, reducing the satellite cell pool available for muscle regeneration [94]. SMAD1, 5, and 8 are the main transducers of intracellular signaling activated by bone morphogenetic proteins (BMPs), a subfamily of the TGF-β superfamily. SMAD7, an I-SMAD, promotes myogenesis and counteracts the negative effects induced by TGF-β and myostatin by inhibiting phosphorylation of SMAD2 [95]. Another way by which SMAD7 affects myogenesis is to enhance MYOD1 mRNA expression [95]. Depletion of SMAD7 causes a shift towards more oxidative fibers in muscle [96]. Taken together, SMAD2 and SMAD3 mediate the inhibitory effects of members of the TGF-β superfamily on myoblast differentiation by interfering with the MRFs or MEF2s. SMAD7 promotes myoblast differentiation by counteracting SMAD2 and SMAD3.

**KLF family of transcription factors**

The Krüppel-like (KLF) family of transcription factors contains 17 members, KLF1-17, all zinc-finger proteins. KLF3 binds to cis-elements within the promoter regions of CKM and MYH2 and other genes associated with myogenesis [97]. At the CKM promoter, KLF3 exhibits synergy with serum response factor (SRF) [97]. Furthermore, KLF3 can interact with the transcriptional co-repressors C-terminal-binding protein 1 (CTBP1) and C-terminal-binding protein 2 (CTBP2) [97]. Because CTBP1 and 2 knockout mice show muscle development defects [98], KLF3-SRF complexes likely work with CTBP1 and 2 to regulate skeletal muscle development. KLF4 promotes myoblast differentiation by
enhancing the expression of SMAD7, which inhibits SMAD2 and 3 phosphorylation [99]. KLF4 and KLF2 might be involved in myoblast fusion because KLF4 and KLF2 can transactivate the expression of nephronectin (NPNT), which drives myoblast fusion [100]. KLF5 is crucial for myoblast fusion and differentiation. There is a significant overlap between the DNA binding sites of KLF5 and those of MYOD1 and the MEF2s, in particular, MEF2D, supporting the notion that KLF5 regulates myoblast fusion and differentiation through interaction with these transcription factors [101]. KLF5 mediates the transmission of signals from several pathways, such as p38 MAPK, ERK1/2, and IL-1β/HIF-1α from the cytosol to the nucleus [102, 103]. KLF6 is not only targeted by MEF2D but shares the same binding sites with MEF2D on the genome in undifferentiated myoblasts [104].

**TEAD family of transcription factors**

The transcriptional enhanced associate domain (TEAD) family consists of TEAD1-4, [105]. These factors are involved in the Hippo signaling pathway and exert their regulatory effects as functional partners of the transcription coactivators Yes-associated protein (YAP) and WW domain containing transcription regulator 1 (WWTR1) [106, 107]. The TEAD proteins have both redundant and distinct functions in myogenesis. TEAD1 facilitates the exit of C2C12 cells from the cell cycle. It drives their entry into the differentiation process, partly by regulating the expression of mitochondrial ribosomal protein L2 (MRPL2) and cyclin E1 (CCNE1), which possess TEAD1 DNA binding sites in their promoters [108]. As a direct target gene of MYOD1 and MYOG, TEAD4 expression is controlled by these two MRFs during C2C12 differentiation [109]. MYOD1 and MYOG also possess TEAD4 DNA binding sites in their promoters, suggesting their expression is regulated by TEAD4 [110]. Additionally, through co-occupied binding sites with MYOD1 and MYOG, TEAD4 engages in synergistic interactions promoting myoblast differentiation [105]. Vestigial-like family member 3 (VGLL3) acts as a co-regulator and interacts with TEAD proteins. It
binds to different TEAD family members depending on the developmental stages of myoblasts. Specifically, it binds to TEAD1 and TEAD3 before myoblast differentiation but interacts with TEAD4 after myotube formation [111].

**SIX family of transcription factors**

The SIX family of transcription factors contains 6 members, SIX1-6 [112]. There is evidence suggesting that they cooperate with MRFs in regulating myogenic development. MRFs-targeted genes contain a large number of DNA binding sites for SIX factors suggesting their involvement in regulating gene expression during myogenic differentiation [109, 113]. One mechanism by which SIX1 promotes myoblast differentiation is through increased expression of MYOG [114]. SIX1 also has a DNA binding site at the core enhancer region of MYOD1 [115]. Another mechanism by which SIX1 regulates myoblast differentiation is through SKI (SKI proto-oncogene), a protein that is essential for myoblast terminal differentiation and that can induce the accumulation of fast-type myofibers [116]. As a non-DNA binding protein, SKI has been shown to regulate MYOG expression through its association with SIX1 and EYA3. Contrary to the stimulatory function of SIX1 in muscle development, SIX4 and 5 repress myogenesis. Exogenous overexpression of SIX4 or 5 inhibits the expression of MYOG in myoblasts [115].

**Other transcription factors**

**AP-1 complex**

The activator protein 1 (AP-1) complex consists of two subunits from the FOS and JUN families. The AP-1 complex in myoblasts and myotubes exhibits distinct characteristics, and the genes regulated by this complex vary depending on the composition of AP-1. The presence of FRA2 in the AP-1 complex increases during myotube formation. FRA2-
containing AP-1 has been shown to modulate the core enhancer activity of MYOD1 [117]. Differentiation of bovine satellite cells into myotubes is associated with increased DNA binding of FOS and FOSB-containing AP-1 complexes to the genome [118]. Besides forming the AP-1 complex, each subunit can also serve as the transcription factor exerting distinct regulatory functions. FRA1, another possible subunit of the AP-1 complex, acts as an upstream regulator and promotes the expression of cyclin D [119], which maintains myoblasts in the cell cycle and prevent them from undergoing terminal differentiation [119]. JUN alone also inhibits myoblast differentiation by hampering the transcription of CKM and MYOD1 [120].

**SRF**

Serum response factor (SRF) is essential for MYOD1 expression in myoblasts as removing SRF mRNA or protein rapidly eliminates MYOD1 expression in C2C12 cells [121]. Besides regulating MYOD1 expression, SRF participates in satellite cell fusion by facilitating actin organization, although this function is not necessary for satellite cell differentiation or proliferation [122]. SRF partially regulates satellite cell proliferation and migration by regulating the expression of interleukin-6 (IL-6) and IL-4 [123].

**ZBED6**

Zinc finger BED-type containing 6 (ZBED6) acts as an inhibitor of myoblast differentiation by downregulating the expression of IGF2. The first intron of the IGF2 gene has a specific ZEBD6 DNA binding site; when this binding site is erased, the expression of IGF2 significantly increases, mimicking the effect of ZBED6 ablation [124]. Loss of ZBED6 leads to a significant increase in the expression of MEF2C, even though there is no DNA binding site located in this gene [125]. Moreover, ZBED6 inhibits the expression of MYOD1 via upregulating the expression of TWIST2, an inhibitor of MYOD1 [125]. Interestingly, ZBED6 DNA binding sites are enriched in chromatin regions marked by
active histone markers, such as acetylation at lysine 27 on H3 histone protein (H3K27ac),
tri-methylation at lysine 4 on H3 histone protein (H3K4me3), and di-methylation at lysine
4 on H3 histone protein (H3K4me2), in myoblasts. This indicates that ZBED6 controls the
expression of many genes during myogenesis [125].

ZEB1

Zinc finger e-box binding homeobox 1 (ZEB1) is a competitor of the E-box binding sites
for MYOD1, thereby regulating the expression of genes usually targeted by MYOD1 [126].
In proliferating myoblasts, ZEB1 binds to more E-box sites than MYOD1, while in
differentiating myoblasts, MYOD1 binds to more E-box sites than ZEB1. ZEB1 reduces
MEF2C expression by recruiting HDAC9 through CtBP-mediated interactions [127].
ZEB1 participates in muscle differentiation and plays a role in muscle atrophy. F-box
protein 32 (FBXO32) is a significant mediator of protein degradation in muscle atrophy
[128]. ZEB1 binds to the promoter of FBXO32, resulting in increased expression of
FBXO32 under atrophic conditions [127]. Additionally, ZEB1 is involved in muscle
regeneration by governing the interaction between muscle fibers and macrophages [126].

FOXO transcription factors

Forkhead box O1 (FOXO1) affects myogenesis by regulating the expression of myostatin,
an important negative regulator of muscle development. The promoter region of myostatin
contains a specific DNA binding site for FOXO1 [129]. FOXO3, another member of the
FOXO family of transcription factors, may similarly regulate myostatin expression [129].

MYOGENIC REGULATORY HISTONE MODIFICATIONS

Histone modification controls the accessibility of chromatin to transcription factors and co-
factors [130, 131]. Increasing evidence suggests that histone modification plays an
essential role in regulating gene expression during myogenesis. The remainder of the review will discuss the functions of the two most studied histone modifications, histone acetylation and histone methylation, in regulating muscle development.

**Histone acetylation and deacetylation in muscle development**

Adding an acetyl group to histone proteins introduces an additional negative charge to the protein core, increasing repulsion between histones and DNA binding, which leads to a more relaxed chromatin structure for gene transcription. Histone acetylation is generally associated with increased transcriptional activity, while removal of acetyl groups or deacetylation indicates a loss of transcriptional potential. H3K27ac is found at transcriptionally active enhancer regions [132]. Histone acetyltransferases (HATs) are responsible for adding and maintaining acetyl modifications on chromatin, while histone deacetylases (HDACs) act as antagonists to HATs by removing acetyl groups from histones and reducing the accessibility of chromatin to transcription factors [133]. Both HATs and HDACs are involved in gene expression during myogenesis.

**Histone acetyltransferases**

Two essential participants in regulating the overall histone acetylation status are p300 and CREB-binding protein (CBP). These two HATs play critical roles in initiating and maintaining the transcription of myogenic genes. Knockdown experiments targeting p300 and CBP have revealed distinct effects on myogenesis. Deletion of p300 results in the failure of the myogenic process, indicating its indispensable role in myogenesis, while deficiency of CBP leads to less pronounced consequences, suggesting that CBP may have a more subtle or compensatory role in regulating myogenesis [134, 135]. p300 specifically induces H3K27ac accumulation at the MYF5 enhancers, leading to the expression of MYF5 at the early myoblast developmental stage [136]. p300 is also required for MYOD1 expression. The expression of MYOD1 is initiated by the activation of its three enhancer
regions: the core enhancer region (CER), proximal enhancer region (PER), and distal enhancer region (DER) [137]. During myoblast differentiation, p300 is predominantly concentrated at the CER, while the PER and DER have lower levels of p300 binding. The acetylation patterns differ between these regions. H3K27ac is enriched at the CER, while H3K18ac and H3K9ac are more prominent at the PER and DER, respectively [138]. p300/CBP-associated factor (PCAF) is another acetyltransferase in myogenesis. PCAF is recruited by MYOD1 to the enhancers of target genes, resulting in additional histone acetylation and, consequently, increased activation of transcription [139].

**Histone deacetylases**

Histone deacetylases (HDACs) are classified as I, II, III, and IV, with class II and III HDACs more active during muscle development [140]. Class I HDACs include HDAC1, HDAC2, HDAC3, and HDAC8. HDAC1 and HDAC2, as transcriptional corepressors, are enriched in genes related to myoblast differentiation. HDAC1 and 2 are recruited by the SNAI1 (snail family transcriptional repressor 1) protein, which serves as a transcription corepressor cooperating with HDACs. During myoblast differentiation, the HDAC1/2-SNAI1 complex identifies chromatin regions bound by MYOD1 and decreases the histone acetylation level nearby, thereby impairing MYOD1-dependent gene expression [141]. Both HDAC1 and 3 can inhibit the expression of myogenic genes without directly interacting with the MEF2 family of transcription factors [142]. HDAC1 and 3, and SIRT1, a class III HDAC, can directly interact with PCAF, a histone acetyltransferase. This interaction reduces the autoacetylation capability of PCAF, consequently impairing PCAF-dependent histone acetylation and decreasing gene transcription [143, 144].

Class II histone deacetylases include HDAC4, HDAC5, HDAC7, and HDAC9. HDAC4 and 5 inhibit the expression of myogenic genes by removing acetyl groups around the MYOD1 and MEF2 DNA binding sites [142]. During myoblast differentiation, HDAC5
undergoes nuclear-cytoplasmic shuttling, and the amount of HDAC5 within the nucleus determines the degree of myoblast differentiation [145]. Calcium/calmodulin-dependent kinase (CaMK) phosphorylates HDAC5, leading to its expulsion from the nucleus. This phosphorylation-mediated relocation of HDAC5 overcomes its repressive effects on the MYOD1 and MEF2 transcription factors, allowing for proper activation of myogenic genes during myoblast differentiation [145].

Class III HDACs, also known as the sirtuins (SIRT1-7), are NAD+-dependent deacetylases [146]. Among the sirtuins, SIRT1, SIRT6, and SIRT7 are predominantly located in the nucleus, while SIRT2 is primarily found in the cytoplasm [147]. SIRT1 has the activity to deacetylate H3K16, and this activity plays an essential role in satellite cell activation [148]. During satellite cell activation, the metabolic switch from fatty acid oxidation to glycolysis increases the ratio of NAD+/NADH, resulting in the activation of SIRT1 and initiation of deacetylation driven by SIRT1 [148]. SIRT1-null mice showed greater amounts of acetylated H3K16, resulting in premature differentiation and the formation of smaller muscle fibers [148]. During myoblast proliferation, SIRT2 removes the acetyl groups from H3K9 and H3K14 localized to the MYOG and MHC loci. This deacetylation activity controls the progression of myogenesis. SIRT2 usually works as a component of a complex that includes MYOD and PCAF [149]. As an inhibitor of muscle differentiation, the catalytic activity of SIRT2 is repressed when the NAD+/NADH ratio decreases during differentiation. This repression helps maintain a hypoacetylation state, contributing to the regulation of gene expression during muscle differentiation [149].
**Fig 1.4** Key methylation on H3 and H4 histone proteins. Histone methylation markers in blue and red have negative and positive effects on gene transcription, respectively. The effect of the histone methylation marker highlighted in gray depends on the genes associated with the marker.

**Histone methylation and demethylation**

Unlike histone acetylation, histone methylation does not significantly alter the charge of chromatin, leading to more complex outcomes that depend on the specific histone residue being modified. Complicating matters further, methylation patterns can vary, with one residue (typically lysine or arginine) being modified by the addition of one, two, or three methyl groups [150]. Furthermore, methylation can be asymmetric or symmetric in the case of di-methylation. Consequently, a growing body of research has investigated the participants involved in histone methylation and the regulatory mechanisms that impact myogenesis.

**Methylation of H3K27**

Tri-methylation of lysine 27 on H3 histone protein (H3K27me3) serves as a repressive histone mark during muscle development [151]. EZH2, a polycomb-group (PcG) family
member, forms a functional complex with the polycomb repressive complex 2 (PRC2). In undifferentiated myoblasts, the PRC2-EZH2 complex maintains H3K27me3 marks on chromatin, thereby preventing the expression of genes at inappropriate time points. However, during differentiation, the occupancy of PRC2-EZH2 complexes on chromatin significantly decreases, coinciding with a reduction in the percentage of H3K27me3 marks. This dynamic regulation of EZH2 and H3K27me3 contributes to the proper timing of gene expression during cellular differentiation [152]. YY1 is a transcription factor that is present in myoblasts both before and after differentiation. It has been observed that YY1 recruits EZH2 and HDAC1 to repressive regions, suppressing gene expression through trimethylation and deacetylation of the H3K27 residue. Upon initiation of differentiation, the repressive complex consisting of YY1, EZH2, and HDAC1 is replaced by an activating complex that involves SRF and MYOD1. This switch from a repressive to an activating complex promotes myoblast differentiation [153, 154]. In contrast, the replacement of EZH2 by its homolog, EZH1, leads to the formation of PRC2-EZH1 which has minimal effects on myoblasts but plays a significant regulatory role during myoblast differentiation. EZH1 binds to the promoter region of the MYOG gene and facilitates the recruitment of MYOD1 to this position. Interestingly, after the initiation of MYOG transcription, EZH1 associates with SUZ12 (SUZ12 polycomb repressive complex 2 subunits) and exhibits a function like EZH2 that prevents the continuous increase in MYOG expression. This regulatory mechanism helps maintain the appropriate level of MYOG during muscle differentiation [155]. SET7, as a histone demethylase, removes the methyl group on lysine residues, which reduces the presence of the H3K27me3 marker on chromatin to influence gene transcription directly. Meanwhile, SET7 inhibits a major histone methyltransferase responsible for the introduction of H3K27me3, SUV39H1, binding to its target genes, which declines the trimethylation on H3K27 indirectly [156]. By decreasing H3K27me3, SET7 contributes to the transcriptional changes associated with myoblast differentiation [156]. Indeed, SET7 is guided to genes involved in myoblast differentiation by
transcription factors MYOD1 and MEF2 [157]. H3K27me3 is also involved in satellite cell myogenesis. During satellite cell differentiation, there is a significant reduction in the amount of H3K27me3 modification [158], indicating a more active transcriptional state in differentiating satellite cells. UTX, a histone-specific demethylase, is essential for muscle regeneration and likely involved in satellite cell proliferation [159]. In myogenesis, UTX removes the H3K27me3 mark at specific genomic regions, such as MYOG and CKM, allowing their expression [160]. The recruitment of UTX to these regulatory regions is guided by SIX4, a transcription factor that facilitates demethylation at the CKM and MYOG loci. Furthermore, once UTX reaches its target genes, it can be carried along by RNA polymerase II, spreading its demethylase activity throughout the coding region [160].

**Methylation of H3K9**

The common methylation on lysine 9 of H3 histone protein is H3K9me2 and H3K9me3. H3K9me3 is a marker associated with heterochromatin, reflecting a condensed and transcriptionally inactive chromatin state [161]. During myoblast differentiation, genes involved in cell cycle regulation, such as RB/E2F, are enriched with H3K9me3 marks in their promoter regions [162], indicating the inhibition of these loci to drive the shift of cell status from pro-proliferation to pro-differentiation. KDM4A is a demethylase that explicitly removes the methyl group from H3K9me3 during myoblast differentiation [163]. At the MYOD1 locus, KDM4A increases chromatin accessibility and hence, MYOD1 expression. Deletion of KDM4A leads to the overaccumulation of H3K9me3, resulting in impaired muscle differentiation [163]. Interestingly, KDM4A also regulates myoblast proliferation by removing the H3K9me3 modification at the MYF5 locus, increasing the expression of CCD1, a key cell cycle regulator [25, 163]. LSD1, another demethylase, modifies the methylation state at H3K9 [164]. During myoblast differentiation, LSD1 is recruited to the core enhancer region of the MYOD1 locus and removes the di-/tri-methyl group from H3K9, thereby increasing the transcriptional activity at the MYOD1 core
enhancer. This modification is necessary to initiate transcription of MYOD1 and other genes during myoblast differentiation.[137, 165]. LSD1 also recruits RNA Polymerase II at the MYOD1 locus [166]. SUV39H1, a methyltransferase, adds and maintains methyl groups at H3K9. Once recruited by MYOD1, the complex consisting of MYOD1 and SUV39H1 induces H3K9 methylation at MYOD1 target genes, thereby decreasing gene transcription [17]. One of the target genes of the SUV39H1-MYOD1 complex is MYOG, highlighting how SUV39H1 inhibits terminal differentiation of myoblasts [17].

**Methylation of H3K4**

Mono-, di-, and tri-methylation can be introduced to lysine 4 on the H3 histone protein. Most investigators believe methyl modification at H3K4 represents positive effects on gene transcription. Indeed, H3K4 methylation is enriched in active enhancer, Pol-II binding, and promoter region [167-169]. H3K4me1 is reported as the pioneer of H3K27ac-marked enhancers [132, 170], indicating the importance of H3K4me1 in initiating a transcriptional active region. H3K4me1 is also enriched at active promoter regions [171], consistent with its pro-transcriptional function. However, the H3k4me1 marker has also been reported to serve as a conditional repressor to the transcription of myogenic genes [172]. Unlike H3K4me1, di- and tri-methylation of H3K4 are usually associated with enhanced gene transcription. These different methylation states of H3K4 contribute to the dynamic regulation of gene expression during myoblast differentiation[172]. The induction and maintenance of H3K4 methylation in myocytes are carried out by different members of the Mixed lineage leukemia 1 (MLL) family, which contain 6 proteins, MLL 1-4, and SET1A and 1B [173, 174]. MLL1 plays a role in regulating myogenesis by inducing H3K4me3 at MYF5, which promotes the expression of CCND1, thereby myoblast proliferation [175]. MLL4, as a histone mono-/di-methyltransferase [176], is required for myoblast differentiation. Even though MLL3 contributes to the induction of H3K4me1/2 at enhancers, MLL4 is primarily responsible for this epigenetic modification. During
myoblast differentiation, MLL4 binds to enhancer regions of myogenic genes and facilitates their transcription by introducing H3K4me1/2 modification [176]. Furthermore, MLL3 and MLL4 maintain H3K4 methylation at the promoter regions in undifferentiated myoblasts, while MLL1 and MLL2 contribute to the same modification in differentiated myoblasts [175]. Besides the MLL family, SET7, another histone methyltransferase, induces monomethylation at H3K4 [177, 178]. During myoblast differentiation, SET7 is recruited by MYOD1 and MEF2 to the pro-differentiation genes, causing monomethylation at enhancer regions [157, 171].

**Methylation of H3K36**

Another common methylation on H3 histone protein is at lysine 36, H3K36. Like others, this lysine residue can potentially receive multiple methyl groups. Ash1L, also known as KMT2H, functions as a methyltransferase and specifically adds two methyl groups to the H3K36 position [179]. Ash1L promotes myoblast fusion by positively regulating its direct target gene, CDON (cell-adhesion-molecule-related/downregulated by oncogenes), through K3K36me2. CDON is a membrane protein crucial for cell-cell contact. Through genomic analysis, H3K36me2 is enriched at the transcription start sites of target genes [180]. Interestingly, this mark is also found in regions previously occupied by the H3K27me3 modification, indicating a dynamic interplay between different histone modifications [180]. SETD2, a trimethyltransferase-inducing tri-methylation of H3K36 residue [181], is also involved in myogenesis. Reduced expression of SETD2 in C2C12 myoblasts impairs their proliferation, as evidenced by a decrease in the expression level of cell cycle markers such as CCND1, CDK4, and CCNE2[182]. Consistently, the expression of the cyclin-dependent kinase inhibitor, p21, increases when SETD2 is lost in C2C12 myoblasts [182]. Reduced expression of SETD2 is associated with the activation of differentiation-related genes, such as MYOG, in C2C12 cells [182]. These findings suggest that the H3K36me3 modification induced by SETD2 plays a role in facilitating myoblast
entry into differentiation.

**Methylation of H4K20**

H4K20 undergoes multiple methylation states. Mono-methylation of H4K20 is catalyzed by KMT5A [183, 184], while di-methylation of H4K20 is induced by SUV4-20H1/KMT5B, which also plays a role in introducing tri-methylation to K4K20 [185]. The majority of H4K20me3, however, is caused by KMT5C. These methyltransferases and demethylases dynamically regulate the methylation status of H4K20, contributing to the epigenetic regulation of gene expression and various cellular processes [186-188]. During induced differentiation of C2C12 myoblasts, there is a notable decrease in the abundance of H3K4me1 and H4K20me3 on the chromatin, culminating three days after the induction of differentiation [187, 189]. The role of H4K20me3 in regulating myogenesis is partially attributed to its control over the transcription of MYF5, MYOD1, and MYOG [190].

**Methylation of histone arginine residues**

Protein arginine methyltransferases (PRMT) comprise two classes, class I and class II, including PRMT4 and PRMT5, respectively, which are responsible for adding methyl groups to histone arginine residues during myogenesis [191-193]. PRMT5 adds methyl groups to H3R8 and H4R3, while PRMT4 targets H3R17 and H3R26 [191-193]. PRMT5 adds two methyl groups to H3R8, causing di-methylation in the MYOG promoter during the early stages of myoblast differentiation. PRMT5 also acts as a coactivator of MYOD1 [194]. PRMT4 plays a role in myoblast differentiation at a later stage, following the initiation of MYOG expression [195]. Unlike PRMT5, which is required for the transcription of MYOG, the loss of PRMT4 has minimal effects on MYOG expression [195]. PRMT5 plays a crucial role in maintaining the activity of the SWI/SNF complex, ensuring chromatin remodeling, which is necessary for the transcription of myogenin [194].
CONCLUSION

Many transcription factors participate in gene regulation during myogenesis. These factors directly influence gene expression or act as scaffolds to recruit other transcription factors essential for myogenesis. During myogenesis, some transcription factors establish a positive feedback loop, accelerating the expression of myogenic genes. Notably, MYOD1 and MYOG are part of this loop, as both contain binding sites for each other. Many transcription factors control myogenesis by affecting the expression or activity of MYOD1 or MYOG, underscoring the role of MYOD1 and MYOD as master transcription factors in myogenesis. Besides transcription factors, histone modifications, particularly histone methylation and acetylation, control gene expression during myogenesis too. H3K27ac, H3K4me2, and H3K4me3 represent histone modifications associated with gene activation, while H3K9me3 and H3K27me3 are the two major histone modifications associated with gene repression. Histone phosphorylation, SUMOylation, and ubiquitination are additional histone modifications that can influence the transcriptional state of chromatin, and their roles in controlling gene expression during myogenesis should be further studied.


15. Carvajal JJ, Keith A, Rigby PW: Global transcriptional regulation of the locus


42. Langlands K, Yin X, Anand G, Prochownik EV: Differential interactions of Id


56. Molkentin JD, Black BL, Martin JF, Olson EN: Cooperative activation of


82. Fornaro M, Burch PM, Yang W, Zhang L, Hamilton CE, Kim JH, Neel BG, Bennett AM: **SHP-2 activates signaling of the nuclear factor of activated T**


94. Paris ND, Soroka A, Klose A, Liu W, Chakkalakal JV: **Smad4 restricts differentiation to promote expansion of satellite cell derived progenitors during skeletal muscle regeneration.** *Elife* 2016, **5**.


120. Bengal E, Ransone L, Scharffmann R, Dwarki VJ, Tapscott SJ, Weintraub H,


ABSTRACT

This study aimed to identify potentially novel mechanisms that mediate myoblast differentiation. We first compared transcriptomes in C2C12 myoblasts before and six days after induction of myogenic differentiation by RNA-seq. This analysis identified 11,046 differentially expressed genes, of which 5615 and 5431 were upregulated and downregulated, respectively, from before differentiation to differentiation. Functional enrichment analyses revealed that the upregulated genes were associated with skeletal muscle contraction, autophagy, and sarcomeres, while the downregulated genes were associated with ribonucleoprotein complex biogenesis, mRNA processing, ribosomes, and other biological processes or cellular components. Western blot analyses showed an increased conversion of LC3-I to LC3-II protein during myoblast differentiation, further demonstrating the upregulation of autophagy during myoblast differentiation. Blocking the autophagic flux in C2C12 cells with chloroquine inhibited the expression of skeletal-muscle-specific genes and the formation of myotubes, confirming a positive role for autophagy in myoblast differentiation and fusion.

Keywords: autophagy; differentiation; myoblast; myotube; RNA-seq

*This chapter has been published as Lyu P, Jiang H. RNA-Sequencing Reveals Upregulation and a Beneficial Role of Autophagy in Myoblast Differentiation and Fusion. Cells. 2022 Nov 10;11(22):3549. doi: 10.3390/cells11223549. PMID: 36428978; PMCID: PMC9688917.
Skeletal muscle is the largest tissue in the body, accounting for nearly half of the body’s entire weight, and has many important functions, including contraction [1]. The basic structural units of skeletal muscle are myofibers, which are multinucleated muscle cells differentiated and fused from the muscle precursor cells myoblasts [2-5]. Significant changes occur in gene transcription during myoblast differentiation and fusion into myofibers, and these changes are regulated in the nucleus by cis-regulatory elements, including enhancers and promotors, and trans-regulatory factors, including transcription factors, co-factors, and other DNA-targeted regulatory proteins [6, 7]. Four master transcription factors have been identified to regulate gene transcription during myogenesis, including myogenic factor 5 (MYF5), MYF6 (also known as MRF4), MyoD (MYOD1), and myogenin (MYOG), of which MYOG is the major transcription factor that regulates gene transcription during myoblast differentiation [8, 9]. DNA methylation plays a vital role in determining the expression of genes critical for myoblast differentiation via changing the accessibility of chromatin to transcription factors or co-factors [10, 11]. Within the cytoplasm, multiple signal transduction pathways, including the mTOR and AMPK pathways, ensure accurate epigenetic control of gene expression during myoblast differentiation [12, 13]. Additionally, microRNA and lncRNA control the protein buildup in differentiating myoblasts partly via altering the stability of mRNA [14, 15]. On the cell surface, various receptors respond to extracellular signals, such as amino acids, hormones, cytokines, and mechanical damage, and initiate the intracellular signal transduction pathways involved in myoblast differentiation [9, 16-18].

It can be easily imagined that significant protein turnover occurs in the cytoplasm during myoblast differentiation and fusion and that myoblast differentiation and fusion might be controlled by protein degradation. Autophagy is a major intracellular protein degradation process by which unnecessary or damaged proteins or organelles are degraded and reused.
as raw materials to reconstruct new cellular components [19]. This process includes packaging the target objects in the phagophores derived from the endoplasmic reticulum [20]; the expansion and sealing of the phagophores with various proteins, including different Atg isoforms [21]; the establishment of autophagosomes; and the fusion of autophagosomes with lysosomes to form autolysosomes, where the target material for recycling is degraded by lysosomal hydrolases. Autophagy has been linked to muscle regeneration following injury [12], muscle aging [22], muscle atrophy [23], and myoblast differentiation [24-29].

The objective of this research was to uncover potentially novel mechanisms that regulate myoblast differentiation. Through an RNA-seq analysis of transcriptomes in C2C12 myoblasts, we identified many biological processes, including autophagy, associated with myoblast differentiation. Through biochemical and immunocytochemical analyses, we confirmed the beneficial role of autophagy-mediated protein degradation in myoblast differentiation and fusion.
MATERIALS AND METHODS

Major Reagents Used in this Study

Growth medium (GM) was composed of 89% Dulbecco’s Modified Eagle Medium (DMEM), 10% fetal bovine serum (FBS), and 1% Antibiotic-Antimycotic (ABAM). Differentiation medium (DM) consisted of 97% DMEM, 2% horse serum, and 1% ABAM. Radioimmunoprecipitation assay (RIPA) buffer consisted of 150 mM NaCl, 5 mM ethylenediaminetetraacetic acid (EDTA), 50 mM Tris-base pH 8.0, 1% Nonidet P40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), and 1% Halt Protease Inhibitor Cocktail. Western blot running buffer consisted of 25 mM Tris-base pH 8.0, 190 mM glycine, and 3.5 mM SDS. Western blot transfer buffer consisted of 25 mM Tris-bas pH 8.0, 190 mM glycine, and 20% methanol. Tris-buffered saline with Tween 20 (TBST) buffer consisted of 20 mM Tris-base pH 8.0, 50 mM NaCl, and 0.05% Tween20. The blocking buffer was TBST, with 5% non-fat milk added.

C2C12 cells (catalog number CRL-1772) were purchased from ATCC (Manassas, VA, USA), FBS (S11150) from Atlanta Biologicals (Minneapolis, MN, USA), ABAM (30-004-Cl) from Corning (Corning, NY, USA), DMEM (11965118) from Thermo Fisher Scientific (Waltham, MA, USA), non-fat milk (M0841) from Labscientific (Highlands, NJ, USA), chloroquine (14194) from Cayman (Ann Arbor, MI, USA), TRIzol Reagent (15596026) from Invitrogen (Waltham, MA, USA), a Direct-zol RNA MicroPrep kit (R2062) from Zymo Research (Irvine, CA, USA), and nitrocellulose membranes (1620147) from BIO-RAD (Hercules, CA, USA). Random primers (C1181), ImProm-II Reverse Transcriptase (A3803), RNasin Ribonuclease Inhibitor (N2615), and Deoxyribonucleoside Triphosphate (dNTP)

Mix (U1515) was purchased from Promega (Madison, WI, USA). Fast SYBR Green Master Mix (4385612) was purchased from Applied Biosystems (Waltham, MA, USA).
The NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (E7760S) was purchased from New England BioLabs (Ipswich, MA, USA). All other reagents and materials used in this study were purchased from either Thermo Fisher Scientific or Sigma-Aldrich (St. Louis, MO, USA) unless otherwise indicated.

**Cell Culture**

C2C12 cells were cultured in GM to expand. When they reached approximately 70% confluency, they were split into 6-well plates at a density of $5 \times 10^5$ cells/well. After one day of culture in GM or when they reached more than 90% confluency, GM was replaced with DM to induce myogenic differentiation, as described previously [30]. To block autophagy in C2C12 cells, 10 µM or 20 µM chloroquine, which inhibits autophagy by blocking the formation of autolysosomes from autophagosomes and lysosomes [31-33], was added to the culture medium. Pilot experiments showed that at concentrations higher than 20 µM, chloroquine caused significant death in C2C12 cells. To control wells, 0.1% water was added, the solvent for chloroquine. The culture medium and treatment were refreshed every two days.

**RNA Extraction and RNA-Seq Library Construction and Sequencing**

Total RNA from C2C12 cells was extracted using TRIzol reagent and Direct-zol RNA Microprep kit, according to the suppliers’ instructions. RNA concentration and purity were measured with a Nanodrop Spectrophotometer (Thermo Fisher Scientific). RNA integrity was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). All RNA samples for RNA-seq library preparation had RNA integrity numbers (RINs) greater than 7.7. RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (E7760S) (New England BioLabs), according to the manufacturer’s instructions. Each library was uniquely indexed. RNAseq libraries were paired-end sequenced on an Illumina sequencing system at Novogene
Corporation (Sacramento, CA, USA).

**Gene Expression Analysis and Gene Ontology Analysis of RNA-Seq Data**

The raw sequencing reads were first filtered with Trimmomatic [34] to remove low-quality reads and reads containing adapters. Clean reads were mapped to the mouse reference genome (mm10) using the STAR (v2.5) program [35]. The uniquely mapped reads were used in the quantification of gene expression levels, which was performed using the HTSeq v0.6.1 program [36]. Gene expression levels were calculated as reads per kilobase of exon model per million mapped reads (FPKM) [37]. A differential expression analysis was performed using the DESeq2 R package (2_1.6.3) [38], and the resulting p-values were adjusted using the Benjamini–Hochberg approach for controlling the false discovery rate (FDR). Genes with adjusted p-values (Padj) < 0.05 were considered as differentially expressed genes (DEGs). Gene ontology (GO) analysis of DEGs was performed in three categories: biological process (BP), cellular compound (CC), and molecular function (MF), using the R package clusterProfiler [39, 40]. In this analysis, the p-value and q-value cutoffs were 0.01 and 0.05, respectively.

**Reverse Transcription Quantitative PCR (RT-qPCR)**

One microgram of total RNA was denatured by incubation with random primers at 70 °C for 10 min, followed by cooling on ice for 5 min. Reverse transcription of total RNA was performed by incubating 5 μL RNA–random primer mix with 4 μL 5 × reverse transcription buffer, 4.8 μL MgCl2, 1 μL dNTP, 1 μL ImProm-II reverse transcriptase, 0.5 μL RNasin Ribonuclease Inhibitor, and 3.7 μL RNase free H2O at 42 °C for 90 min, at 70 °C for 10 min, and then at 4 °C for 5 min. Quantitative PCR of 20 ng cDNA was performed in duplicate using Fast SYBR Green Master Mix on a 7500 Fast Real-Time PCR system (Applied Biosystems). The PCR primers used in this study were designed in a previous study [41]. Relative gene expression levels were calculated using the ΔΔCt method [42],

48
using the HMBS gene as a reference gene, which, based on its Ct value (data not shown), was stably expressed in C2C12 cells during differentiation.

**Immunocytochemistry**

C2C12 cells were first fixed by incubating them with 4% formaldehyde solution for 15 min at room temperature and then rinsed twice with PBS. Cells were then permeabilized by incubation with 0.25% Triton X-100 for 10 min at room temperature. After that, cells were rinsed again with PBS and then incubated with 1% BSA and 0.05% Tween-20 in PBS for one hour at room temperature with shaking to block nonspecific antibody binding. Cells were then incubated with the antibody for myosin heavy chains (NA-4, DSHB, Iowa City, IA, USA) at 1:100 in PBS at 4 °C overnight. Cells were rinsed twice with PBS and incubated with an anti-mouse IgG FITC antibody (Sigma-Aldrich) at 1:200 dilution for 1h at room temperature. Nuclei were stained by incubating cells with 1 µg/mL of 40,6diamidino-2-phenylindole (DAPI) for 10 min at room temperature. Images of stained cells were taken with a fluorescence microscope.

**Determination of Fusion Index, Myotube Length, and Myotube Area**

Numbers of nuclei were counted, and lengths of myotubes and areas of myotubes were measured from randomly taken immunofluorescent images of myotubes using ImageJ[43]. A myotube was defined as a myosin-heavy-chain-positive (“green”) cell containing two or more DAPI-stained nuclei (“blue”); the length of a myotube was defined as the longest distance between two ends of a myotube, and fusion index was defined as the ratio of the number of nuclei in myotubes to the number of total nuclei counted. Around 1000 total nuclei were counted, and 15 myotubes were measured for each treatment in each experiment.
Total Cellular Protein Extraction and Western Blot Analysis

Total protein lysates from C2C12 cells were prepared by scraping and incubating them in the RIPA buffer with proteinase inhibitors added for 30 min on ice, followed by centrifugation at 16,000×g, 4 °C, for 15 min. Protein concentrations were determined using a Pierce BCA Protein Assay Kit (23227, obtained from Thermo Fisher Scientific). For the Western blot analysis, 20 µg of total protein lysates were separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with 6% stacking and 15% separating gel. Gel electrophoresis was run at 80 V for 30 min and then 120 V for 80 min at 4 °C. Following separation, the protein was transferred from the gel to a nitrocellulose membrane by electrophoresis at 70 V, 4 °C, for 60 min. To block nonspecific antibody binding, the membrane was incubated with 5% non-fat milk in the TBST buffer for 1 h at room temperature. The membrane was then incubated with a primary antibody overnight at 4 °C. The membrane was rinsed twice with TBST and then incubated with a secondary antibody for 1 h at room temperature. The following primary antibodies were used in Western blot analyses: β-tubulin at 1:1000 dilution (E7 from DSHB), LC3B at 1:1000 dilution (NB100-2220 from Novus Biologicals, Centennial, CO, USA), and myogenin at 1:25 dilution (F5D from DSHB). The following secondary antibodies were used in Western blot analyses: IRDye 800CW goat anti-rabbit IgG secondary antibody (926-32211, LI-COR Biosciences, Lincoln, NE, USA) at 1:15000 dilution and IRDye 800CW goat anti-mouse IgG secondary antibody (926-32210, LI-COR Biosciences) at 1:15,000 dilution. Western blots were visualized with the LI-COR Odyssey Infrared Image System 9120, and the fluorescent intensities were quantified with the Image Studio Lite software (LI-COR Biosciences).

Statistical Analyses

Data were analyzed by ANOVA, followed by multiple comparisons using Tukey’s test or
Dunnett’s test. Two-group comparisons were carried out via t-tests. All statistical analyses were performed in JMP Pro 16 (SAS, Cary, NC, USA) or GraphPad Prism 9 (San Diego, CA, USA). All data are expressed as means ± standard errors.
RESULTS

More Than Ten Thousand Genes Were Differentially Expressed during Myoblast Differentiation

We performed an RNA-seq analysis to identify genes differentially expressed during myogenic differentiation in C2C12 cells. The sequencing of five RNA-seq libraries constructed from C2C12 cells immediately before induction of myogenic differentiation and five RNA-seq libraries constructed from C2C12 cells on day 6 of induced myogenic differentiation generated more than 36 million unique mapped sequencing reads per library. Gene expression quantification and subsequent differential expression analysis revealed 11,046 genes differentially expressed (Padj < 0.05) between the two conditions, of which 5615 genes were upregulated, and 5431 genes downregulated from the day before induction of differentiation to day 6 of differentiation (Figure 1A). Examples of genes upregulated in C2C12 cells during induced myogenic differentiation were Myh1, Myh2, Myh3, Myh4, Myog, Ckm, Mymk, and Mb, which are all known to be specifically or preferentially expressed in skeletal muscle [44, 45]. Examples of genes downregulated during this differentiation included Ccna2, Ccnd1, Ccne1, Ccne2, Id1, Id3, and Cdc6, which are known to function in cell proliferation [46, 47]. To confirm the gene expression differences quantified by RNA-seq, we measured the expression levels of Myh1, Myh2, Myh3, Myh4, and Myog mRNAs in the two conditions by RT-qPCR. As shown in Figure 1B, average fold changes measured by RT-qPCR were similar to those quantified by RNA-seq for all of these mRNAs.

Autophagy Was Upregulated during Myoblast Differentiation

Gene ontology enrichment analyses of genes upregulated during myogenic differentiation in C2C12 cells revealed the biological processes (BPs), cellular components (CCs), and molecular functions (MFs) associated with these genes. Shown in Figure 2A are the top 10
upregulated BPs, CCs, and MFs during myogenic differentiation in C2C12 cells. Shown in Table 2.1 are two examples of top-upregulated BPs and associated genes. As expected, most of these upregulated BPs were related to skeletal muscle development and maturation, including muscle cell differentiation, muscle system process, striated muscle cell differentiation, muscle contraction, and muscle tissue development (Figure 2.2A). Interestingly, four of the top-upregulated BPs were related to autophagy (Figure 2.2A, Table 2.1). Similar to the BPs, most of the top-upregulated CCs during myogenic differentiation in C2C12 cells were related to skeletal muscle structures, such as sarcomeres, I-bands, and Z-discs (Figure 2.2A). Two of the top-upregulated CCs were related to organelles that are involved in protein degradation, namely, lysosomes and lytic vacuoles (Figure 2.2A). The top-upregulated MFs during myogenic differentiation in C2C12 cells included actin binding, motor activity, and calmodulin binding, which are all known as functions of mature skeletal muscle [48, 49]. It is also interesting to note that small GTPase binding, Ras GTPase binding, coenzyme binding, Rab GTPase binding, protein serine/threonine kinase activity, and enzyme activator activity were among the MFs upregulated during myogenic differentiation in C2C12 cells (Figure 2.2A).

GO analyses of genes downregulated during myogenic differentiation in C2C12 cells indicated that many of them were associated with BPs, CCs, and MFs related to DNA replication and RNA processing. Examples of the top-downregulated BPs during myogenic differentiation in C2C12 cells were ribonucleoprotein complex biogenesis and DNA replication; examples of top-downregulated CCs were ribosomes and chromosomal regions; and examples of top downregulated MFs were mRNA binding and helicase activity (Figure 2.2B, Table 2.2).

**Myoblast Differentiation Was Associated with Increased Conversion of LC3-I to LC3-II**
LC3 protein, encoded by the microtubule-associated proteins 1A/1B light chain 3B (Map1lc3b) gene, is the most widely used marker of autophagy [50]. During autophagy, the cytosolic LC3 protein LC3-I is conjugated with phosphatidylethanolamine to form the autophagosomal membrane protein LC3-II [51]. Western blot analyses (Figure 3A) showed that the protein expression ratio of LC3-II to LC3-I in C2C12 cells increased rapidly from the day before induction of differentiation (day 0 in the figure) to day 1 of differentiation and remained high on day 2 and day 4 of differentiation (Figure 3B). These changes indicated that the autophagic flux increased in C2C12 cells during myogenic differentiation, confirming what was revealed by the RNA-seq analysis described above.

**Inhibition of Autophagy Reduced the Number, Length, and Size of Myotubes Formed during Myoblast Differentiation**

To determine the role of increased autophagy in myoblast differentiation, we induced C2C12 cells to differentiate in the presence of chloroquine (CQ), a widely used inhibitor of autophagic flux that inhibits autophagy by blocking the fusion of autophagosomes with lysosomes [32]. We first performed Western blot analyses to determine whether CQ is effective in inhibiting autophagy in C2C12 cells. As shown in Figures 4A, B, the ratios of LC3-II to LC3-I in C2C12 cells treated with CQ were significantly higher than those in control C2C12 cells on days 1, 2, and 4 of myogenic differentiation. These increases confirmed the effectiveness of CQ in blocking autophagosome–lysosome fusion and hence the degradation of LC3-II protein by the lysosomal hydrolases. These Western blot analyses also confirmed the increases in the generation of LC3-II in C2C12 cells during myogenic differentiation (Figure 4A).

Based on the morphology, multinucleated myotubes began to form in C2C12 cells two days after the initiation of myogenic differentiation, and multinucleated myotubes could be easily observed by day 4 of differentiation (Figure 5A). C2C12 cells treated with CQ
appeared to form fewer and smaller myotubes than untreated C2C12 cells on the same day of differentiation (Figure 5A). The higher concentration (20 µM) of CQ apparently caused greater decreases in the number and size of formed myotubes than the lower concentration (10 µM) of CQ; myotubes were barely seen in C2C12 cells treated with 20 µM CQ on day 4 of differentiation (Figure 5A). Interestingly, at high magnification, “dark rings” around the nuclei could be observed in CQ-treated C2C12 cells (Figure 5A). These “dark rings” were probably formed by autophagosomes which accumulated due to their blocked fusion with lysosomes and hence their blocked degradation by lysosomal enzymes. To more clearly show the morphological differences between CQ-treated and control C2C12 cells, we stained the myosin heavy-chain proteins and nuclei in C2C12 cells on day 4 of differentiation and quantified the number of nuclei fused into myotubes and the lengths and areas of myotubes. As shown in Figure 5B, myotubes in CQ-treated C2C12 cells were clearly fewer and smaller than in control C2C12 cells, and these differences were more obvious with the higher concentrations of CQ. The average fusion indexes of C2C12 cells treated with 10 µM and 20 µM CQ were approximately 30% and 60%, respectively, lower than that of control C2C12 cells (p < 0.05, Figure 5C). The average lengths of myotubes formed from C2C12 cells treated with 10 µM and 20 µM CQ were approximately 31% and 43%, respectively, shorter than that of myotubes formed from control C2C12 cells (p < 0.05, Figure 5C). The average areas of myotubes formed from C2C12 cells treated with 10 µM and 20 µM CQ were 40% and 50%, respectively, smaller than that of myotubes formed from control C2C12 cells (p < 0.05, Figure 5C).

Inhibition of Autophagy Reduced the Expression of Muscle-Specific Genes during Myoblast Differentiation

To further determine the effect of the inhibition of autophagy on myoblast differentiation, we quantified the expression levels of several skeletal-muscle-specific genes in CQ-treated and control C2C12 cells on day 4 of myogenic differentiation. These genes included Myh1,
Myh3, Tnnt3, Mb, Ckm, Myog, and Mymk, which encode either structural or functional components of skeletal muscle (Myh1, Myh3, Tnnt3, Mb, and Ckm) or are master regulators of myoblast terminal differentiation and fusion (Myog and Mymk) [44]. As shown in Figure 6A, the expression of most of these genes was inhibited in C2C12 cells treated with 10 µM CQ, and the expression of all of these genes was inhibited in C2C12 myoblasts treated with 20 µM CQ compared with their expression in untreated C2C12 cells. We also confirmed the effect of the inhibition of autophagy on the expression of myogenin at the protein level. As expected, myogenin protein expression in untreated C2C12 cells was significantly higher on days 1, 2, and 4 of differentiation than on the day before the induction of differentiation (Figure 6B, C). Compared to the control, CQ significantly reduced myogenin protein expression in C2C12 cells on days 1 and 2 of differentiation (Figure 6D).
DISCUSSION

In this study, we first compared the gene expression profiles between undifferentiated and differentiating C2C12 myoblasts. Our RNA-seq analysis revealed that the expression of many skeletal-muscle-specific genes increased while the expression of many genes involved in the cell cycle decreased during myoblast differentiation. These results were consistent with those of earlier microarray-based gene expression profiles for C2C12 cells [52-54], indicating that myoblast differentiation involves a shift in gene expression from genes functioning in cell proliferation to genes functioning in skeletal muscle buildup and contraction. Our RNA-seq analysis revealed that many genes functioning in mRNA processing, ribonucleoprotein complex biogenesis, and ribosomes were downregulated during myogenic differentiation in C2C12 cells. This result suggests that myoblast differentiation is associated with delayed processing (i.e., capping, polyadenylation, and splicing) of newly synthesized mRNA (i.e., pre-mRNA) in the nucleus, delayed export of processed mRNA from the nucleus to the cytoplasm, and delayed translation of exported mRNA in the cytoplasm. These delays may be meant to reduce loss of mRNAs through membrane pores formed during myoblast fusion [55].

Our RNA-seq analysis also revealed that many genes involved in autophagy and related functional terms were upregulated during myoblast differentiation. These gene expression changes, as well as those involved in ribonucleoprotein complex biogenesis and mRNA processing mentioned above were not identified by earlier microarray-based gene expression analyses [52-54], perhaps because many of the genes involved in autophagy, ribonucleoprotein complex biogenesis, and mRNA processing were not included in the microarrays or in the functional analyses performed in these studies, or because microarray based gene expression analysis is less sensitive than RNA-seq in detecting changes in gene expression, in particular, changes in low-abundance transcripts [56-58].
The upregulation of genes functioning in autophagy during myoblast differentiation indicates that autophagy-mediated protein degradation increases during myoblast differentiation. Indeed, we confirmed increased autophagy in C2C12 cells during myogenic differentiation by detecting increased conversion of LC3-I to LC3-II in differentiating C2C12 myoblasts and by detecting increased accumulation of LC3-II in autophagy-blocked differentiating C2C12 myoblasts. The association of increased autophagy with myoblast differentiation suggests that autophagy might benefit myoblast differentiation. A positive role for autophagy in myoblast differentiation is supported by our results showing that blocking autophagy in C2C12 myoblasts impaired their expression of muscle-specific genes and their fusion into multinucleated myotubes. Our finding that autophagy plays a positive role in myoblast differentiation was consistent with earlier studies using different approaches or different myoblast models [24-29]. However, the conclusion that autophagy aids myoblast differentiation seems to contradict the observation that inducing autophagy with the mTOR inhibitor rapamycin inhibited myoblast differentiation in C2C12 cells [59]. One possible reason for this discrepancy is that the target of rapamycin, mTOR, controls not only autophagy but also other processes, such as overall protein synthesis, which is essential for myoblast differentiation [60, 61].

Autophagy is a process by which cells maintain hemostasis and survival by delivering organelles and proteins to lysosomes for degradation [62]. There is likely an increased generation of damaged, malformed, or unnecessary organelles and proteins during the differentiation and fusion of mononuclear myoblasts into multinucleated myotubes [63]. Therefore, increased autophagy may be one of the means used by myoblasts to clear these organelles and proteins and thereby maintain normal myogenic differentiation. Indeed, this role for autophagy in myoblast differentiation is supported by previous findings that upregulated autophagy is essential for mitochondrial degradation and biogenesis and protection against mitochondrial oxidative stress during myoblast differentiation [64, 65].
Myogenin is a master transcriptional regulator of myoblast differentiation [8]. Myogenin binds to and activates the transcription of muscle-specific genes as a heterodimer with E-proteins [66]. However, myogenin's DNA-binding and transactivating ability is inhibited by a group of proteins called the inhibitor of DNA binding (ID) proteins, which lack basic DNA binding domains [66]. The ubiquitin-proteasome system was previously found to degrade the ID proteins in differentiating myoblasts [67]. We speculate that the autophagy system might also degrade the ID proteins and thereby promote myoblast differentiation. Besides the ID proteins, proteins such as the cyclin-dependent kinases and transforming growth factor beta (TGF-β) also inhibit myoblast differentiation by keeping myoblasts in the proliferation phase [46, 68]. We speculate that the autophagy system could also target and degrade these proteins to promote myoblast differentiation.

CONCLUSIONS

A massive number of genes are upregulated or downregulated during myoblast differentiation. Genes upregulated during myoblast differentiation include those associated with skeletal muscle structure and contraction and autophagy. Genes downregulated during myoblast differentiation include those involved in the cell cycle, ribonucleoprotein complex biogenesis, and mRNA processing. Not only the expression of autophagic genes but autophagic activity is increased during myoblast differentiation. Increased autophagy is confirmed to benefit myoblast differentiation and fusion.
REFERENCE


30. Ge X, Zhang Y, Park S, Cong X, Gerrard DE, Jiang H: Stac3 inhibits myoblast...


Mukund K, Subramaniam S: Skeletal muscle: A review of molecular structure


68. Brennan TJ, Edmondson DG, Li L, Olson EN: Transforming growth factor
**Table 2.1** Two examples of functional terms associated with genes upregulated during myoblast differentiation.

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
<th>Gene Ratio</th>
<th>p-Value</th>
<th>GeneID</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0042692</td>
<td>muscle cell differentiation</td>
<td>186/5118</td>
<td>$1.28 \times 10^{-28}$</td>
<td>Pgm5, Myl2, Cacna1s, Lmod2, Tnnt3, Mybpc2, Myoz2, Mybpc1, Actc1, Casq1, Trim72, Actn2, Neb, Casq2, Myom3, Myh6, Cav3, Ankrd2, Csrp3, Mynp, Tcap, Rbm24, Klhl41, Tnnt1, Acta1, Myom1, Myom2, Fbxo40, Lmod3, Igf2, Igf1, Dmd, Myog, Alpk3, Ins2, Klhl40, Ryr1, Ankrd23, Tmod1, Myo18b, Tnu, Sypl2, Mef2c, Neu2, Smyd1, Synpo2l, Hopx, Nrap, Dysf, Fgf9, et al.</td>
</tr>
<tr>
<td>GO:0006914</td>
<td>autophagy</td>
<td>181/5118</td>
<td>$1.32 \times 10^{-27}$</td>
<td>Dcn, Fez1, Mapt, Trem2, Synpo2, Epm2a, Bmf, Htr2b, Casp1, Prkaa2, Trp53inp2, Ifng, Pink1, Lzts1, Map1lc3a, Pik3c2b, SrpX, Rnf152, Rasip1, Mtm1, Trl2, Irgm2, LiX1, Fbxw7, Usp13, Snapin, Adrb2, Fyc1, Trp53inp1, Stat3, Atp6v0a1, Nupr1, Tfeb, Dram2, Trib3, Nod1, A tg4a, Ctsd, Qsox1, Atg10, Zfyve1, Trim13, Atg13, Vps13d, Flen, Wipi1, Nod2, Fbxl2, Sirt2, Hspb8, et al.</td>
</tr>
</tbody>
</table>
Table 2.2 Two examples of functional terms associated with genes downregulated during myoblast differentiation.

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
<th>Gene Ratio</th>
<th>p-Value</th>
<th>GeneID</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0022613</td>
<td>ribonucleoprotein complex biogenesis</td>
<td>299/4657</td>
<td>$1.47 \times 10^{-107}$</td>
<td>Celf4, Suv39h1, Lyar, Exosc8, Nop56, Lsm2, Npm1, Lsm3, Dkc1, Nhp2, Fbl, Gar1, Gemin6, Snrpd1, Dis3, Ppan, Ruvbl2, Hsp90aa1, Nop2, Ncl, Ran, Noc21, Pa2g4, Exosc2, Xpo1, Nop58, Rc11, Snrpg, Npm3, Rrp15, Snrpe, Rbmx11, Mybbp1a, Ddx20, Ddx18, Rrp9, Ruvbl1, Ddx21, Rrs1, Snrpd3, Mrt04, Utp20, Rpsa, Eri1, Lsm6, Gemin5, Lsm4, Srsf1, Wdr43, Rps2, et al.</td>
</tr>
<tr>
<td>GO:0006397</td>
<td>mRNA processing</td>
<td>275/4657</td>
<td>$5.32 \times 10^{-83}$</td>
<td>Ccnb1, Celf4, Hmx2, Pthp1, Lsm2, Slbp, Npm1, Lsm3, Hnrnpa1, Lsm5, Snrpa1, Gemin6, Snrpd1, Ppl1, Rbmx2, Alyref, Pabpc1, Mbn13, Khdrbs3, Adarb1, Srsf7, Papolb, Ddx39, Snrpg, Zfp473, Srsf9, Dazap1, Lsm8, Tbrg4, Snnp40, Magohb, Snrpe, Rbmx11, Ddx20, Srt, U2af2, Lsm7, Snrpd3, Hnrrph1, Hnrrpm, Tra2a, Ttf2, Hnrrpa3, Snrpa, Rbm19, Cstf2, U2af1, Lsm6, Gemin5, Lsm4, et al.</td>
</tr>
</tbody>
</table>
Fig 2.1 Identification of differentially expressed genes during myoblast differentiation. mRNA transcriptomes in C2C12 cells before (n = 5) and six days after (n = 5) induction of myogenic differentiation were analyzed by RNA-seq. (A) The overall distribution of differentially expressed genes (Padj < 0.05). Genes upregulated and downregulated on day 6 of differentiation compared to before differentiation are indicated by red and blue dots, respectively. (B) Comparison of the fold changes in the expression of selected mRNAs determined by RNA-seq and RT-qPCR.
**Fig 2.2** Top 10 biological processes, cellular components, and molecular functions associated with genes upregulated (A) and downregulated (B) during myoblast differentiation. The X-axis indicates the functional terms, and the Y-axis indicates the number of genes linked to each functional term. In each functional category, terms are ranked from left to right according to their adjusted p-values, with the leftmost term having the smallest p-value.
Fig 2.3 Increased generation and accumulation of LC3-II in C2C12 cells during myogenic differentiation. Total protein lysates from C2C12 cells on day (D) 0 (i.e., before differentiation) and days 1, 2, and 4 of myogenic differentiation were analyzed by Western blotting. (A) Representative images of Western blots. (B) Average ratios of LC3-II to LC3-I protein. * and ** indicate p < 0.05 and <0.002, respectively (n = 6), based on Dunnett’s test.
Fig 2.4 Effects of chloroquine (CQ) on LC3-I and LC3-II protein expression in C2C12 cells during myogenic differentiation. C2C12 cells were cultured in differentiation medium in the presence (+) or absence (−) of 20 µM CQ. Total protein lysates from cells on days 0 (D0, before differentiation), 1 (D1), 2 (D2), and 4 (D4) of differentiation were analyzed by Western blotting, using antibodies specific to LC3 and β-Tubulin (as a loading control). (A) Representative images of Western blots. (B) Ratio of LC3-II to LC3-I protein. ** indicates p < 0.002 (n = 5), based on t-tests.
Fig 2.5 Effects of inhibiting autophagy with chloroquine on the morphological changes in C2C12 cells during myoblast differentiation. C2C12 cells were cultured in differentiation medium in the presence of 10 or 20 µM chloroquine. Control (ctrl) C2C12 cells were
cultured in differentiation medium in the absence of chloroquine. (A) Representative micrographs of C2C12 cells on days (D) 0, 1, 2, and 4 of differentiation. Arrows point to “dark rings” around the nuclei in CQ-treated cells. (B) Representative micrographs of C2C12 cells on day 4 of differentiation. Cells were stained with myosin heavy-chain (MHC) antibody (green) and DAPI (blue). (C) Quantification of fusion index and length and area of myotubes. Bars not sharing the same letter labels are different \( p < 0.05, n = 8 \), based on Tukey’s test.
Fig 2.6 Effects of inhibiting autophagy with chloroquine on the expression of selected muscle-specific genes in C2C12 cells during myogenic differentiation. C2C12 cells were cultured in differentiation medium in the presence of 10 or 20 µM chloroquine for 4 days. Control (ctrl) C2C12 cells were cultured in differentiation medium without chloroquine. (A) Relative mRNA expression levels. Bars not sharing the same letter label are different (p < 0.05, n = 6), based on Tukey’s test. (B) Representative Western blots of myogenin. (C) Comparison of myogenin protein expression levels in C2C12 cells between day (D) 0 and
days 1, 2, and 4 of differentiation. * and *** indicate p < 0.05 and 0.01, respectively, based on Dunnett’s test (n = 5). (D) Comparison of myogenin protein expression levels in C2C12 cells between chloroquine and control. *, **, and ns indicate p < 0.05, 0.002, and > 0.05, respectively, based on t-tests (n = 5). Data were log-transformed for normality prior to statistical analysis.
Identification of Enhancers and Transcription Factors Regulating the Myogenic Differentiation of Bovine Satellite Cells

ABSTRACT

**Background:** Satellite cells are the myogenic precursor cells in adult skeletal muscle. This study aimed to identify enhancers and transcription factors that regulate gene expression during the differentiation of bovine satellite cells into myotubes. **Results:** Chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) was performed to identify genomic regions where lysine 27 of H3 histone is acetylated (H3K27ac), i.e., active enhancers from bovine satellite cells before and during differentiation into myotubes. 19,027 and 47,669 H3K27ac-marked enhancers were consistently identified from two biological replicates of before- and during-differentiation bovine satellite cells. Of these enhancers, 5,882 were specific to before-differentiation, 35,723 to during-differentiation, and 13,199 common to before- and during-differentiation bovine satellite cells. Whereas most of the before- or during-differentiation-specific H3K27ac-marked enhancers were located distally from the transcription start site, the enhancers common to before- and during-differentiation were located both distally and proximally from the transcription start site. The three sets of H3K27ac-marked enhancers were associated with functionally different genes and enriched with different transcription factor binding sites. Specifically, many of the H3K27ac-marked enhancers specific to during-differentiation bovine satellite cells were associated with genes involved in muscle structure and development. They were enriched with binding sites for MyoD, AP-1, KLF, TEAD, and MEF2 transcription factors. A positive role was validated for TFAP4, a basic helix-loop-helix (bHLH) transcription factor in differentiating bovine satellite cells into myotubes by siRNA-induced knockdown
and plasmid-mediated overexpression. **Conclusions:** Tens of thousands of enhancers active in bovine satellite cells before or during differentiation have been identified. These enhancers contain binding sites for transcription factors whose role in satellite cell differentiation is well known and for transcription factors whose role in satellite cell differentiation is unknown. These enhancers and transcription factors are valuable for understanding the complex mechanism that mediates gene expression during satellite cell differentiation. Because satellite cell differentiation is a crucial step in skeletal muscle growth, the enhancers, the transcription factors, and their target genes identified in this study are also valuable for interpreting skeletal muscle trait-associated DNA sequences and mutations in cattle.

**Keywords:** Cattle, enhancer, skeletal muscle, transcription factor, TFAP4

*Part of this chapter has been published as Lyu, P., Settlage, R.E. & Jiang, H. Genome-wide identification of enhancers and transcription factors regulating the myogenic differentiation of bovine satellite cells. BMC Genomics 22, 901 (2021). https://doi.org/10.1186/s12864-021-08224-7*
INTRODUCTION

Skeletal muscle is the largest tissue in the body and plays an essential role in physiology [1, 2]. Skeletal muscle from meat-producing animals is a major food source for humans and animals. Adult skeletal muscle is composed mostly of muscle fibers [3]. A muscle fiber, also known as a myofiber, is a multinucleated muscle cell differentiated and fused from multiple mononuclear muscle cells called myoblasts. For most mammals, the total number of myofibers is determined prenatally [4]; thus, postnatal skeletal muscle growth results primarily from myofiber hypertrophy [4, 5]. Postnatal myofiber hypertrophy, however, requires additional nuclei [6, 7]. In postnatal animals, nuclei added to the existing myofibers are widely believed to come from satellite cells, mononuclear cells located near myofibers, and considered stem cells in adult skeletal muscle [8-10]. Satellite cells are normally quiescent but can be activated by muscle injury and nutritional and environmental changes [6, 7]. Once activated, satellite cells become and proliferate as myoblasts and then differentiate and fuse to generate new myotubes, the developing myofibers, or with existing myofibers to increase muscle fiber size.

Four transcription factors called myogenic regulatory factors (MRFs) play essential roles in myogenesis, forming muscle fibers from myoblasts or satellite cells [11, 12]. These MRFs include myogenic differentiation 1 (MYOD1, also known as MyoD and MYF3), myogenic factor 5 (MYF5), myogenin (MYOG, also known as MYF4), and myogenic factor 6 (MYF6, also known as MRF4 and herculin). All four MRFs are specifically or preferentially expressed in skeletal muscle [12]. All four MRFs are basic helix-loop-helix (bHLH) domain-containing transcription factors and regulate gene transcription by binding to the E-box sequence, CANNTG, where N is A, G, C, or T [13]. MYF5 and MYOD1 determine the myogenic lineage of stem cells redundantly [11, 12]. MYOG is essential for myoblast differentiation and fusion into myotubes [12, 14]. MYF6 was thought to play a similar role to MYOG in myoblast differentiation, but a more recent study indicated an
unexpected negative role of MYF6 in postnatal skeletal muscle growth [15].

Clearly, the differentiation and fusion of myoblasts or satellite cells into myotubes is a critical step in developing and growing skeletal muscle. The objective of this study was to further understand the regulation of gene expression during the differentiation of bovine satellite cells into myotubes. Cattle are agriculturally important animals, and a better understanding of gene regulation during satellite cell differentiation could lead to developing novel strategies to improve growth efficiency and meat quality in cattle. Genomic regions where lysine 27 of histone 3 protein is acetylated (H3K27ac) are considered active enhancers [16, 17]. We began this study by identifying genomic regions with H3K27ac modification in bovine satellite cells before and during induced differentiation and fusion into myotubes through chromatin immunoprecipitation coupled with deep sequencing (ChIP-seq).
MATERIALS AND METHODS

Isolation and culture of bovine satellite cells

Skeletal muscle was collected from Angus-crossbred steers slaughtered at the Virginia Tech Meat Center. Satellite cells were isolated through pronase digestion and differential centrifugation as described before [18, 19]. Satellite cells were cultured in the growth medium for about a week before being induced to differentiate and fuse into myotubes. Differentiation of bovine satellite cells into myotubes was induced by replacing the growth medium with the differentiation medium. The growth medium consisted of Dulbecco’s Modified Eagle Medium (DMEM), 10% fetal bovine serum (FBS) (R&D Systems, Minneapolis, MN), 2 mM L-glutamine, and 1% Antibiotic-Antimycotic (100×) (ABAM). The differentiation medium consisted of DMEM, 2% horse serum (R&D Systems), 2 mM L-glutamine, and 1% ABAM. All cell culture was performed at 37º C in a humidified, 5% CO₂ atmosphere. Unless otherwise indicated, all cell culture reagents were purchased from ThermoFisher Scientific (Waltham, MA). Satellite cells immediately before and two days after induction of differentiation were used in the following assays.

ChIP assay

Satellite cells were cross-linked in 1% formaldehyde for 10 minutes and then lysed in lysis buffer from the ChIP-IT kit (Active Motif, Carlsbad, CA). Cell nuclei were suspended in ChIP buffer from the ChIP-IT kit and then sheared on ice by 10 pulses of 20-second sonication using a sonic dismembrator Model 100 at setting 3 (ThermoFisher Scientific) to generate chromatin fragments of 200 to 1,000 bp. To identify the genomic regions associated with H3K27ac modification, chromatin fragments were incubated with an anti-histone H3K27ac antibody (ab4729, Abcam, Cambridge, MA) at 4 ºC overnight with gentle rocking. The H3K27ac antibody-chromatin complexes were separated from unbound chromatin fragments using protein G-Dynal beads (ThermoFisher Scientific). Chromatin
fragments immunoprecipitated by the H3K27ac antibody, and those before immunoprecipitation (i.e., input chromatin) were reverse cross-linked by incubating them at 65 °C for 4 hours. DNA was extracted and purified using spin columns from the ChIP-IT kit.

**ChIP-seq library construction and sequencing**

According to the supplier's instructions, ChIP-seq libraries were prepared using the NEBNext ChIP-Seq Library Prep Reagent Set for Illumina (New England BioLabs, Ipswich, MA). Briefly, ChIP-DNA or Input DNA was end-repaired using T4 DNA polymerase, Klenow DNA polymerase, and T4 polynucleotide kinase. End-repaired DNA was then added with 3’ dA overhangs using exonuclease minus Klenow DNA polymerase and dATP. The dA-tailed DNA fragments were ligated to the sequencing adaptor. DNA fragments of approximately 300 bp were selected from the adaptor-ligated DNA using AMPure XP Beads. The size-selected DNA fragments were amplified in 12 cycles of PCR using an index primer and a universal PCR primer. Each ChIP-seq or Input-seq library was assessed for quality on a Bioanalyzer before being sequenced. Four ChIP-seq libraries and two Input-seq libraries passed these initial quality assessments. They were single-end sequenced on an Illumina Hiseq 2500 at the Genomics Sequencing Center at Virginia Tech.

**ChIP-seq data analyses**

Sequences from ChIP-seq libraries were first trimmed to remove the adapters using Trimmomatic [20]. The trimmed reads were then mapped to the bovine genome assembly (ARS-UCD 1.2 BosTaur 9) using Hisat2 (2.2.0) [21]. The aligned reads were sorted and merged using SAMtools (1.9) [22]. Peak calling of the aligned reads was made using MACS3 (3.0.0a5) [23], where an H3K27ac-ChIP-seq library was compared to an Input-seq library (i.e., control) made from the same satellite cells, and where the q-value threshold was set as 0.05. The quality of peak enrichment in ChIP-seq reads was assessed
by [24, 25]. ChIP-seq peaks were visualized in the IGV browser (2.8.2) [26]. ChIP-seq peaks were annotated using ChIPseeker (1.22.1) [27]. Motif enrichment analyses were performed using HOMER (4.11.1) [28] or Meme-suite [29, 30] Gene ontology (GO) enrichment analysis was performed using the PANTHER Classification System [31-33].

**Small interfering RNA (siRNA)-mediated gene knockdown**

Bovine satellite cells in 12-well plates were transfected with 10 nM of siRNA targeting bovine TFAP4 mRNA using Lipofectamine RNAiMAX Reagent according to the supplier’s instructions (ThermoFisher Scientific). A universal negative control siRNA (MISSION siRNA Universal Negative Control #1, Millipore Sigma, Burlington, MA) was transfected as a negative control. The sense and antisense siRNA sequences targeting bovine TFAP4 mRNA were CACUCAGAAGGUGCCCUCUUUGCAA and UUGCAAAGAGGGCACCUUCUGAGUG, respectively. Twenty-four hours after the transfection, differentiation was induced by replacing the growth medium with the differentiation medium. Cells were cultured in the differentiation medium for four days. The differentiation degree of satellite cells was assessed by quantifying mRNA expression of markers of differentiated myoblasts, including CKM, MYH2, MYH3, and MYOG, as previously described [34].

**RNA extraction and reverse transcription-quantitative PCR (RT-qPCR)**

Total RNA was extracted from bovine satellite cells using the Direct-zol RNA Miniprep Kit (Zymo Research). Total RNA transcription into cDNA was repeated using ImProm-II reverse transcriptase and random primers according to the manufacturer’s instructions (Promega, Madison, WI). Quantitative PCR was performed using the SYBR Green chemistry as described previously [34]. The relative abundance of target mRNAs was calculated using the $2^{-\Delta\Delta Ct}$ method [35]. The Ct values for target mRNAs were normalized to the Ct values for HMBS, which was chosen as a reference gene because it was stably
expressed in different conditions [36].

**Immunocytochemistry and fusion index calculation**

Cells were first fixed by incubating them with 4% formaldehyde for 15 minutes at room temperature, followed by two rinses with PBS. Cells were permeabilized by incubation with 0.25% Triton X-100 for 10 minutes at room temperature. Cells were then incubated with 1% BSA and 0.05% Tween-20 in PBS at room temperature for 1 hour with gentle shaking to block nonspecific binding. Cells were incubated with the primary antibody for MYH proteins (NA-4, DSHB, Iowa City, IA, USA) at a 1:100 dilution in PBS overnight at 4°C. After two washes with PBS, cells were incubated with an anti-mouse IgG FITC secondary antibody (Sigma-Aldrich) at a 1:200 dilution for 1 hour at room temperature. Finally, cell nuclei were stained with 1 µg/mL of 4',6-diamidino-2-phenylindole (DAPI) for 10 minutes at room temperature. The stained cells were imaged using a fluorescence microscope. The number of total nuclei and the number of nuclei located in MYH-stained myotubes were counted from randomly selected areas of images using the ImageJ software. The fusion index was calculated as the number of nuclei within myotubes divided by the total nuclei.

**Total cellular protein extraction and western blot analysis**

Total cellular proteins were isolated by incubating cells with the RIPA buffer containing proteinase inhibitors on ice for ten minutes, followed by centrifugation at 15,000 g and 4 °C for 30 minutes. Protein concentration was measured using a BSA assay kit. For Western blot analysis, 20 µg of total protein was separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (6% stacking gel, 80V for 30 minutes; 10% separating gel, at 120V for 90 minutes, both at 4°C). After separation, the protein was transferred from the gel to a nitrocellulose membrane by electrophoresis (70 V, 4°C, 90 minutes). The membrane was first blocked with 5% non-fat milk in TBST at room
temperature for 1 hour and then incubated with the primary antibody overnight at 4°C. The membrane was then washed twice with TBST and subsequently incubated with the secondary antibody at room temperature for 1 hour. Two primary antibodies were used in this study. One was β-tubulin antibody (E7 from DSHB) used at 1:1000 dilution. The other one was TFAP4 antibody (A-8, sc-377042, Santa Cruz) used at 1:1000 dilution. The secondary antibody used in this study was IRDye 800CW goat anti-mouse IgG secondary antibody (926-32210, LI-COR Biosciences) at 1:15,000 dilution. Western blots were visualized using the LI-COR Odyssey Infrared Image System 9120, and fluorescence intensity was quantified using the Image Studio Lite software (LI-COR Biosciences).

**Statistical analysis**

ANOVA analyzed gene expression data. Two means were compared by t-test, and the Tukey test compared multiple means. All data are expressed as mean ± standard error.
RESULT

H3K27ac-marked enhancers in bovine satellite cells before and during differentiation

Four ChIP-seq libraries and two Input libraries constructed from bovine satellite cells immediately before and 2 days after induction of differentiation passed the quality control, and deep sequencing generated 23 to 40 million sequencing reads from these libraries (Table 3.1). Between 75% and 92% of these reads were uniquely mapped to the bovine genome, generating approximately 20 to 36 million uniquely mapped reads per library (Table 3.1).

Analyzing the uniquely mapped reads from each ChIP-seq library against those from the corresponding Input library using the MACS peak calling program identified more than 30,000 and 50,000 H3K27ac-marked genomic regions, i.e., active enhancers, from before-differentiation (BD) and during-differentiation (DD) bovine satellite cells, respectively. A phantompeakqualtools analysis indicated that the four ChIP-seq libraries had normalized strand cross-correlation (NSC) values between 1.16 and 1.18 and relative strand cross-correlation (RSC) values between 0.98 and 1.06, indicating strong enrichment of reads in peaks [24].

A Pearson correlation analysis revealed that the H3K27ac-marked enhancer regions identified from two biological replicates were highly correlated (Fig. 3.1A). A total of 19,027 H3K27ac-marked enhancers were consistently identified from two biological replicates of BD bovine satellite cells. In comparison, 47,669 H3K27ac-marked enhancers were consistently identified from two biological replicates of DD bovine satellite cells (Fig. 3.1B). A total of 5,882 H3K27ac-marked enhancers were found to be specific to BD bovine satellite cells, 35,723 H3K27ac-marked enhancers specific to DD bovine satellite cells, and 13,199 H3K27ac-marked enhancers common to both BD and DD bovine satellite cells.
Examples of H3K27ac-marked enhancers identified from BD and DD bovine satellite cells are shown in Table 3.2 and Fig. 3.2A. These enhancers were associated with the MYOG gene mentioned above, which is a transcription factor essential for myoblast differentiation, and the myosin heavy chain 3 (MYH3) gene, which, as indicated by its name, encodes a skeletal muscle-specific myosin heavy chain protein. As shown in Table 3.2 and Fig. 3.2A, MYOG-associated enhancers were marked with H3K27ac only in DD bovine satellite cells; MYH3-associated enhancers were marked with H3K27ac in both BD and DD bovine satellite cells, but more MYH3-associated enhancers were marked with H3K27ac in DD than in BD bovine satellite cells. As shown by Fig. 3.2B, the increased H3K27ac modification to MYOG- and MYH3-associated genomic regions in DD satellite cells was accompanied by increased expression of both genes in these cells.

**Genomic distribution of H3K27ac-marked enhancers in bovine satellite cells**

H3K27ac-marked enhancers specific to BD or DD bovine satellite cells had a different genomic distribution from the H3K27ac-marked enhancers common to both BD and DD bovine satellite cells. Whereas nearly 90% of H3K27ac-marked enhancers specific to BD or DD bovine satellite cells were located in the distal intergenic regions and introns, this percentage was only 60% for H3K27ac-marked enhancers common to BD and DD bovine satellite cells (Fig. 3.3A). Whereas approximately 6% and 17% of H3K27ac-marked enhancers specific to BD and DD bovine satellite cells, respectively, were located in the promoter regions, this percentage was almost 50% for H3K27ac-marked enhancers common to BD and DD bovine satellite cells (Fig. 3.3A). Whereas H3K27ac-marked enhancers specific to BD or DD bovine satellite cells were concentrated at 100,000 bp from the transcription start site (TSS), those common to BD and DD bovine satellite cells were concentrated at both 100 bp and 100,000 bp from the TSS (Figs. 3.3B and 3.3C).

**Functional terms enriched in genes associated with H3K27ac modification in bovine**
satellite cells

We performed gene ontology (GO) enrichment analyses on genes associated with H3K27ac modification in bovine satellite cells. Top biological processes and cellular components enriched in genes associated with H3K27ac modification, specifically in DD bovine satellite cells, were related to skeletal muscle structure, development, adaption, and cell cycle arrest (Tables 3.3 and 3.4). Interestingly, many genes involved in melanosome and phagocytic vesicle were also associated with H3K27ac modification in DD bovine satellite cells (Table 3.4). Most of the top molecular functions enriched in genes associated with H3K27ac modification, specifically in DD bovine satellite cells, were related to growth factor binding and serine and threonine kinase signaling (Table 3.5). Top biological processes, cellular components, and molecular functions enriched in genes associated with H3K27ac modification in BD bovine satellite cells included pentose metabolic process, postsynaptic membrane, transmembrane receptor protein tyrosine kinase activity, respectively, which were different from those enriched in genes associated with H3K27ac modification in DD satellite cells (Tables 3.3-3.5). Top biological processes, cellular components, and molecular functions enriched in genes associated with H3K27ac modification in BD and DD bovine satellite cells included those related to proteasome and autophagosome.

Transcription factor binding sites enriched in H3K27ac-marked enhancers in bovine satellite cells

Motif enrichment analyses of 35,373 enhancer regions marked with H3K27ac, specifically in DD bovine satellite cells, indicated enrichment of binding sites for many transcription factors. The top 30 motifs enriched in these enhancers included binding sites for the bHLH transcription factors MYF5, MYOG, TFAP4 (also known as AP4), MYOD1, TCF12, TCF21, ATOH1, and ASCL2, the basic leucine zipper (bZIP) transcription factors JUN,
FOS, FOSB, FOSL2 (also known as FRA2), FOSL1 (also known as FRA1), BATF, JUNB, BACH2, and ATF3, the Krüppel-like family (KLF) transcription factors KLF1, KLF5, and KLF14 (Table 3.6). Motif enrichment analyses of 5,882 enhancer regions marked with H3K27ac specifically in BD bovine satellite cells and 13,199 enhancers marked with H3K27ac in both BD and DD bovine satellite cells revealed enrichment of different sets of transcription factor binding sites. Top motifs enriched in enhancers marked with H3K27ac, specifically in BD bovine satellite cells, included binding sites for transcription factors ZIC3, TCP16, ASCL2, TFAP2C (also known as AP-2 gamma), ZIC2, MAX, and MYC. Top motifs enriched in enhancers marked with H3K27ac in both BD and DD bovine satellite cells included many members of the ETS family transcription factors such as ELK1, ELF1, ELK4, GABPA, ETV4, FLI1, ETV1, and ELF4.

**TFAP4 potential binding sites in differentiating bovine satellite cells**

As mentioned, the TFAP4 binding site was one of the top motifs enriched in H3K27ac-marked enhancers in differentiating bovine satellite cells. We located those TFAP4 binding sites using the FIMO tool of the MEME package [29]. Because of the lack of consensus sequence for bovine TFAP4 binding sites, two human TFAP4 motifs (MA1570.1 and MA0691.1 from the JARSPA database, Fig. 3.4A) were used to search the bovine enhancers for TFAP4 binding sites. This search indicated that 5,942 of 48,923 H3K27ac-marked enhancers in differentiating bovine satellite cells contained 6,763 potential TFAP4 binding sites. Table 7 shows examples of H3K27ac-marked enhancers containing TFAP4 binding sites and the specific locations of these binding sites. The 6,762 TFAP4 binding sites were associated with 3,157 genes. These genes included many genes involved in myogenesis (e.g., MYOD1, MYOG, MYH3, and CKM), genes related to cell cycle (e.g., CCND1 and CCND3), and genes related to energy metabolism (e.g., UCP1 and GSK3B). Fig. 3.4B shows two examples of TFAP4 binding sites associated with the MYOG and CKM genes. Using the SpaMo tool in the MEME package [30], we searched for motifs co-
existing with the TFAP4 motifs in H3K27ac-marked enhancers in differentiating bovine satellite cells. As shown in Fig 3.4C, Nr1H2 was found to be one of the motifs located next to the TFAP4 motifs. Other motifs co-existing with TFAP4 motifs included the binding sites for Foxn1, RELB, GLI3, EBF1, HINFP, and SOX18. These data indicated that Nr1H2, Foxn1, RELB, GLI3, EBF1, HINFP, and SOX18 may collaborate with TFAP4 in regulating gene expression in differentiating bovine satellite cells.

We next performed the GO enrichment analysis on TFAP4 binding sites-associated genes to see what biological processes might be regulated by TFAP4. Among the top GO terms enriched in these genes were anatomical structure morphogenesis, cellular response to an organic substance, tissue development, regulation of intracellular signal transduction, response to endogenous stimulus, muscle organ development, muscle tissue development, and striated muscle hypertrophy (Table 3.4).

**siRNA-induced TFAP4 knockdown delayed the differentiation of bovine satellite cells into myotubes**

The motif enrichment analyses of H3K27ac-marked enhancers in differentiating bovine satellite cells indicated that gene transcription during bovine satellite cell differentiation is controlled by additional transcription factors besides the widely known four MRFs. One of these additional transcription factors is TFAP4. To investigate the role of TFAP4 in myogenesis, we knocked down TFAP4 expression in bovine satellite cells using siRNA. Compared to bovine satellite cells transfected with a negative control siRNA, bovine satellite cells transfected with TFAP4 siRNA showed delayed formation of myotubes upon differentiation induction (Fig. 3.5A). Furthermore, myotubes formed from bovine satellite cells transfected with TFAP4 siRNA were smaller than those formed from control cells (Fig. 3.5A). Myotubes formed from cells transfected with TFAP4 siRNA also showed weaker MYH staining than myotubes formed from control cells (Fig. 3.5B). The fusion
index for the cells transfected with the TFAP4 siRNA was much lower than that for the control cells (P < 0.05, Fig. 3.5C). The expression levels of TFAP4 mRNA and protein in the cells transfected with TFAP4 siRNA were 50% less than those in the control cells, indicating effective TFAP4 knockdown by the transfected siRNA (Fig. 3.6). These data together suggested that siRNA-induced knockdown of TFAP4 inhibited the capacity of bovine satellite cells to differentiate into myotubes. In other words, these data indicated a positive role of TFAP4 in the differentiation of TFAP4 into myotubes.

**siRNA-induced TFAP4 knockdown reduced the expression of myogenic marker genes in differentiating bovine satellite cells**

To further determine the effect of TFAP4 knockdown on the differentiation of bovine satellite cells into myotubes, we quantified the expression levels of selected markers of myotubes, including MYH2, MYH3, MYOG, CKM, and MYMK. As shown in Fig. 3.6A, the mRNA levels of these five markers were significantly lower in cells transfected with TFAP4 siRNA than in control cells. We also determined the effect of TFAP4 knockdown on the expression of MYH, MYOG, and CKM genes at the protein level. As shown in Fig. 3.6C, the protein levels of these markers were 50% less in TFAP4-knocked down cells than in control cells. These expression data were consistent with the morphological data (Fig 3.5A), indicating that TFAP4 knockdown inhibited the differentiation of bovine satellite cells into myotubes.

**Effect of TFAP4 overexpression on differentiation of bovine satellite cells into myotubes**

We also determined the effect of TFAP4 overexpression on the differentiation of bovine satellite cells into myotubes. While no differences in morphology were observed between bovine satellite cells transfected with a TFAP4 expression plasmid and cells transfected with a negative control plasmid during 4-day differentiation (Fig. 3.7A), the mRNA levels
of myotube marker genes MYH3, CKM, and MYMK were higher (P < 0.05) in TFAP4-overexpressed cells than in control cells, and the mRNA levels of MYH2 and MYOG tended to be higher (0.05 < P < 0.1) in TFAP4-overexpressed cells than in control cells (Fig. 3.7B). Western blot analyses showed that MYH, CKM, and MYOG protein expression levels were higher or tended to be higher in TFAP4-overexpressed cells than in control cells (Fig. 3.8). These data indicated that TFAP4 overexpression had a moderate effect on the expression of myogenic genes in bovine satellite cells.
DISCUSSION

Histone modification affects gene transcription by altering the accessibility of chromatin and recruiting transcription factors and cofactors to chromatin [37]. Large-scale histone modification mapping has revealed that different types of enhancers are associated with different histone modifications. Active enhancers are associated with H3K27ac and histone 3 lysine 4 monomethylation (H3K4me1) modifications; primed or poised enhancers are marked with H3K4me1 but not H3K27ac; and silenced or repressed enhancers are often associated with H3K27me3 modification [38-41]. Based on these associations, ChIP-seq has been widely used to identify enhancers and other types of regulatory DNA regions in whole genomes [42-48]. This study identified 19,027 and 47,669 H3K27ac-marked enhancers in bovine satellite cells before and during differentiation and fusion into myotubes, respectively. Identifying these enhancers provides a valuable resource for understanding the mechanism that regulates gene expression during satellite cell differentiation, an essential step in skeletal muscle development and growth.

Compared to the consistent identification of 47,669 H3K27ac-marked active enhancers from two biological replicates of during-differentiation bovine satellite cells, only 19,027 H3K27ac-marked active enhancers were repeatedly identified from two samples of before-differentiation bovine satellite cells. This difference suggests much more active transcription factor binding to the genome, more active recruitment of histone acetylases, and hence much more active H3K27ac modification in bovine satellite cells during differentiation than before differentiation. There is a possibility that this difference was caused by biological variation, as indicated by the significant difference in the numbers of H3K27ac-marked enhancers identified from two samples of before-differentiation satellite cells. In our research on satellite cells, we have noticed that satellite cells from different animals differ in differentiation potential in culture, and this difference suggests animal-to-animal variation in gene expression and histone modification in satellite cells.
Enhancers can be located upstream or downstream of TSS, in introns or exons, and near or distantly from the promoters [49-51]. Some enhancers can be located in intergenic regions several hundred kilobases away from TSS and control gene transcription by forming DNA loops with the promoters [52-54]. Genomic distribution analyses showed that H3K27ac-marked enhancers in bovine satellite cells have similar genomic distribution to enhancers in other types of cells or species [49-51]. However, the genomic location of H3K27ac-marked enhancers in bovine satellite cells varies with the differentiation stage of these cells. Whereas most of the differentiation stage-dependent H3K27ac-marked enhancers in bovine satellite cells were located in the distal intergenic regions, most of the differentiation stage-independent H3K27ac-marked enhancers in bovine satellite cells were found in the promoter regions. This study also showed that H3K27ac-marked enhancers specific to during-differentiation bovine satellite cells were associated with muscle organization, adaptation, and development genes.

In contrast, H3K27ac-marked enhancers common to both before- and during-differentiation satellite cells were associated with genes involved in basic cellular functions and processes. These results suggest that distal enhancers are preferentially activated to increase the expression of genes determining the differentiation stage of satellite cells. In contrast, proximal enhancers or promoters are preferentially activated to increase the expression of genes maintaining the basic cellular function of satellite cells. This differentiation stage-dependent activation of distal and proximal enhancers in bovine satellite cells is apparently consistent with earlier findings that distal enhancers mediate the expression of cell type- or developmental stage-specific genes while core promoters and proximal enhancers are responsible for the expression of housekeeping genes [55, 56].

This study showed that genes associated with H3K27ac modification were expressed at greater levels than those without H3K27ac modification in bovine satellite cells, regardless of the differentiation stage of the cells. This result supports H3K27ac as a histone marker
for transcriptional activation [16]. This study also showed that many genes were associated with multiple H3K27ac-marked enhancers but that the numbers of H3K27ac-marked enhancers were not correlated with the expression levels of associated genes in bovine satellite cells. These results suggest that multiple H3K27ac-marked enhancers do not function in an additive manner to increase gene expression or that multiple H3K27ac-marked enhancers are functionally redundant in bovine satellite cells. Indeed, recent studies through the CRISPR-Cas9 approach demonstrate that not every enhancer is functionally important and that most enhancers provide only a supportive or backup role in regulating gene expression [57-59].

H3K27ac modification at enhancers results from transcription factor binding and subsequent recruitment of histone acetyltransferases such as p300 and CBP [60, 61]. This study identified many transcription factors that may bind to H3K27ac-marked enhancers in bovine satellite cells before or during differentiation. Among the transcription factors that are predicted to bind to H3K27ac-marked enhancers in during-differentiation bovine satellite cells are MYOG and MYOD1, which are widely known as the central transcriptional regulators of myoblast differentiation and regulate the expression of muscle-specific genes by binding to the motif called E-box [62]. Identifying the binding sites for MYOG and MYOD1 among the top motifs enriched in H3K27ac-marked enhancers in during-differentiation bovine satellite cells validates the quality of the enhancers identified in this study. This study shows that besides MYOG and MYOD1, many other transcription factors regulate gene expression during satellite cell differentiation. These other transcription factors include the AP-1 family of transcription factors (e.g., FOS), the KLF family of transcription factors (e.g., KLF1), the TEAD family of transcription factors (e.g., TEAD1), the MEF2 family of transcription factors (e.g., MEF2A), and TFAP4. Enrichment of binding sites for the AP-1 family of transcription factors in active enhancers in during-differentiation bovine satellite cells is intriguing because the member of the AP-1 family of transcription factors JUN is known to antagonize
the stimulatory effect of MYOD1 on myoblast differentiation [63]. However, overexpression of JUNB, a member of the AP-1 family of transcription factors closely related to JUN, increased hypertrophy and the muscle-specific gene MYH4 expression in C2C12 myoblasts [64]. Therefore, different members of the AP-1 family of transcription factors might have different effects on gene expression during myoblast or satellite cell differentiation. Besides the AP-1 family of transcription factors, the MEF2 family transcription factors MEF2A, MEF2C, and MEF2D [62, 65], the KLF family transcription factors KLF3 and KLF5 [66, 67]. The TEAD family transcription factors TEAD2, and TEAD4 [68-70] have also been shown to play a positive role in myoblast differentiation.

TFAP4 is a basic helix-loop-helix leucine-zipper transcription factor [71, 72], and its role in epithelial-mesenchymal translation and cancer has been extensively studied [73, 74]. Enrichment of the TFAP4 binding site in the enhancers active in differentiating bovine satellite cells suggests that TFAP4 might also play a role in myogenesis. This study found that many bovine myogenic genes contain potential TFAP4 binding sites and that siRNA-induced knockdown of TFAP4 in bovine satellite cells inhibited the differentiation of these cells into myotubes, including the expression of several myogenic genes containing putative TFAP4 binding sites in their enhancer regions. These results suggest that TFAP4 plays an essential role in myoblast differentiation by directly controlling the expression of myogenic genes. In this study, we found that TFAP4 overexpression, however, did not cause as significant an effect as TFAP4 knockdown on the differentiation of bovine satellite cells into myotubes. This difference suggests that the expression level of endogenous TFAP4 in bovine satellite cells may be sufficient for affecting gene expression and myoblast differentiation and that the activity rather than the expression level of TFAP4 determines its role in the differentiation of bovine satellite cells into myotubes.
CONCLUSION

In summary, we have identified thousands of active enhancers in before- or during-differentiation bovine satellite cells. These enhancers contain binding sites for many transcription factors, including MYOG and MYOD1, whose role in myoblast or satellite cell differentiation is widely known, and the AP-1 transcription factors, TFAP4, and many others, whose role in myoblast or satellite cell differentiation is less known or unknown. These enhancers and transcription factors should be valuable for elucidating the mechanism that mediate gene transcription during myoblast or satellite cell differentiation. Indeed, TFAP4 is confirmed to be an essential transcription factor for the differentiation of bovine satellite cells into myotubes. Because satellite cell differentiation is a crucial step of skeletal muscle development and growth, the enhancers, the transcription factors, and the genes targeted by these enhancers and transcription factors identified in this study should also be valuable for identifying and interpreting skeletal muscle trait-associated DNA sequences and mutations in cattle, which are agriculturally important animals.
REFERENCE


54. Dean A: **On a chromosome far, far away: LCRs and gene expression.** Trends Genet 2006, **22**(1):38-45.


### Table 3.1 Mapping summary of ChIP-seq libraries

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Library</th>
<th>Quality Filtered Reads</th>
<th>Mapped Reads</th>
<th>Mapping Rate (%)</th>
<th>Uniquely Mapped Reads</th>
<th>Uniquely Mapping Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Input</td>
<td>33,653,601</td>
<td>29,089,550</td>
<td>86.4</td>
<td>25,223,784</td>
<td>75.0</td>
</tr>
<tr>
<td>1</td>
<td>H3K27ac_BD</td>
<td>29,711,647</td>
<td>27,459,149</td>
<td>92.4</td>
<td>24,819,411</td>
<td>83.5</td>
</tr>
<tr>
<td>1</td>
<td>H3K27ac_DD</td>
<td>39,617,976</td>
<td>38,013,420</td>
<td>96.0</td>
<td>36,139,994</td>
<td>91.2</td>
</tr>
<tr>
<td>2</td>
<td>Input</td>
<td>24,480,296</td>
<td>21,751,323</td>
<td>88.9</td>
<td>19,052,773</td>
<td>77.8</td>
</tr>
<tr>
<td>2</td>
<td>H3K27ac_BD</td>
<td>23,063,317</td>
<td>21,580,412</td>
<td>93.6</td>
<td>19,676,362</td>
<td>85.3</td>
</tr>
<tr>
<td>2</td>
<td>H3K27ac_DD</td>
<td>31,550,183</td>
<td>30,421,776</td>
<td>96.4</td>
<td>29,009,219</td>
<td>92.0</td>
</tr>
</tbody>
</table>

1BD, before differentiation; DD, during differentiation.

### Table 3.2 Examples of H3K27ac-marked enhancers and associated genes in bovine satellite cells

<table>
<thead>
<tr>
<th>Chr</th>
<th>PeakStart</th>
<th>PeakEnd</th>
<th>GeneStart</th>
<th>GeneEnd</th>
<th>GeneSymbol</th>
<th>CellStage</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr16</td>
<td>782062</td>
<td>782253</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>798279</td>
<td>800338</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>800596</td>
<td>800870</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>801536</td>
<td>801833</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>804051</td>
<td>804227</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>804761</td>
<td>805609</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>806018</td>
<td>806188</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>806808</td>
<td>807049</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>830439</td>
<td>830694</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>834418</td>
<td>835077</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>836171</td>
<td>836762</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>Chromosome</td>
<td>Start Base 1</td>
<td>End Base 1</td>
<td>Start Base 2</td>
<td>End Base 2</td>
<td>Gene Symbol</td>
<td>Stage</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------</td>
<td>-------</td>
</tr>
<tr>
<td>chr16</td>
<td>837735</td>
<td>838159</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>838508</td>
<td>838769</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr16</td>
<td>839360</td>
<td>839542</td>
<td>797555</td>
<td>800395</td>
<td>MYOG</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29582805</td>
<td>29583077</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29594980</td>
<td>29595243</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29595616</td>
<td>29597652</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29598274</td>
<td>29598915</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29603029</td>
<td>29603187</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29612384</td>
<td>29612573</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29613638</td>
<td>29613770</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29614001</td>
<td>29614998</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29618302</td>
<td>29618478</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29624364</td>
<td>29624841</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29625585</td>
<td>29626580</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29631500</td>
<td>29631690</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29632712</td>
<td>29632903</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29634635</td>
<td>29634790</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29635013</td>
<td>29635603</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29635962</td>
<td>29636178</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29636376</td>
<td>29636759</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29638316</td>
<td>29638761</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29615976</td>
<td>29616342</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>BD&amp;DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29621065</td>
<td>29621208</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>BD&amp;DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29622186</td>
<td>29622458</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>BD&amp;DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29623347</td>
<td>29623592</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>BD&amp;DD</td>
</tr>
<tr>
<td>chr19</td>
<td>29641704</td>
<td>29641866</td>
<td>29601526</td>
<td>2962481</td>
<td>MYH3</td>
<td>BD&amp;DD</td>
</tr>
</tbody>
</table>

1Differentiation stage of cells from which H3K27ac-marked peaks were identified. BD, before differentiation; DD, during differentiation.

**Table 3.3** Top 10 GO biological processes enriched in genes associated with H3K27ac-
marked enhancers in differentiating bovine satellite cells

<table>
<thead>
<tr>
<th>GO biological process</th>
<th>FE$^1$</th>
<th>P-value</th>
<th>FDR$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>sarcomere organization (GO:0045214)</td>
<td>2.48</td>
<td>7.92E-04</td>
<td>2.10E-02</td>
</tr>
<tr>
<td>striated muscle cell development (GO:0055002)</td>
<td>2.38</td>
<td>5.82E-05</td>
<td>2.14E-03</td>
</tr>
<tr>
<td>myofibril assembly (GO:0030239)</td>
<td>2.37</td>
<td>7.88E-05</td>
<td>2.82E-03</td>
</tr>
<tr>
<td>regulation of muscle adaptation (GO:0043502)</td>
<td>2.12</td>
<td>9.64E-04</td>
<td>2.51E-02</td>
</tr>
<tr>
<td>negative regulation of G1/S transition of mitotic cell cycle (GO:2000134)</td>
<td>2.06</td>
<td>2.06E-03</td>
<td>4.82E-02</td>
</tr>
<tr>
<td>muscle cell development (GO:0055001)</td>
<td>2.03</td>
<td>1.78E-06</td>
<td>8.88E-05</td>
</tr>
<tr>
<td>skeletal muscle tissue development (GO:0007519)</td>
<td>1.94</td>
<td>4.96E-05</td>
<td>1.85E-03</td>
</tr>
<tr>
<td>skeletal muscle organ development (GO:0060538)</td>
<td>1.93</td>
<td>4.87E-05</td>
<td>1.82E-03</td>
</tr>
<tr>
<td>positive regulation of cellular amide metabolic process (GO:0034250)</td>
<td>1.92</td>
<td>1.24E-05</td>
<td>5.19E-04</td>
</tr>
<tr>
<td>negative regulation of mitotic cell cycle phase transition (GO:1901991)</td>
<td>1.92</td>
<td>1.08E-05</td>
<td>4.60E-04</td>
</tr>
</tbody>
</table>

$^1$Fold enrichment; $^2$False discovery rate
Table 3.4 Top 10 GO cellular components enriched in genes associated with H3K27ac-marked enhancers in differentiating bovine satellite cells

<table>
<thead>
<tr>
<th>GO cellular component</th>
<th>FE(^1)</th>
<th>P-value</th>
<th>FDR(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z disc (GO:0030018)</td>
<td>2.24</td>
<td>1.06E-06</td>
<td>2.96E-05</td>
</tr>
<tr>
<td>I band (GO:0031674)</td>
<td>2.10</td>
<td>2.90E-06</td>
<td>7.53E-05</td>
</tr>
<tr>
<td>myofibril (GO:0030016)</td>
<td>2.10</td>
<td>1.99E-09</td>
<td>8.02E-08</td>
</tr>
<tr>
<td>sarcomere (GO:0030017)</td>
<td>2.08</td>
<td>2.68E-08</td>
<td>8.83E-07</td>
</tr>
<tr>
<td>contractile fiber (GO:0043292)</td>
<td>2.08</td>
<td>1.73E-09</td>
<td>7.30E-08</td>
</tr>
<tr>
<td>nuclear matrix (GO:0016363)</td>
<td>2.00</td>
<td>3.76E-04</td>
<td>6.82E-03</td>
</tr>
<tr>
<td>nuclear periphery (GO:0034399)</td>
<td>1.93</td>
<td>2.53E-04</td>
<td>4.77E-03</td>
</tr>
<tr>
<td>actin filament bundle (GO:0032432)</td>
<td>1.93</td>
<td>1.54E-03</td>
<td>2.37E-02</td>
</tr>
<tr>
<td>phagocytic vesicle (GO:0045335)</td>
<td>1.92</td>
<td>3.86E-04</td>
<td>6.94E-03</td>
</tr>
<tr>
<td>melanosome (GO:0042470)</td>
<td>1.87</td>
<td>6.24E-04</td>
<td>1.08E-02</td>
</tr>
</tbody>
</table>

\(^1\)Fold enrichment; \(^2\)False discovery rate
Table 3.5 Top 10 GO molecular functions enriched in genes associated with H3K27ac-marked enhancers in differentiating bovine satellite cells

<table>
<thead>
<tr>
<th>GO molecular function</th>
<th>FE(^1)</th>
<th>P-value</th>
<th>FDR(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>growth factor binding (GO:0019838)</td>
<td>1.74</td>
<td>4.24E-04</td>
<td>2.61E-02</td>
</tr>
<tr>
<td>protein serine kinase activity (GO:0106310)</td>
<td>1.63</td>
<td>6.99E-05</td>
<td>6.20E-03</td>
</tr>
<tr>
<td>protein threonine kinase activity (GO:0106311)</td>
<td>1.62</td>
<td>9.06E-05</td>
<td>7.44E-03</td>
</tr>
<tr>
<td>isomerase activity (GO:0016853)</td>
<td>1.62</td>
<td>1.54E-04</td>
<td>1.16E-02</td>
</tr>
<tr>
<td>ubiquitin protein ligase binding (GO:0031625)</td>
<td>1.60</td>
<td>2.38E-05</td>
<td>2.24E-03</td>
</tr>
<tr>
<td>protein domain specific binding (GO:0019904)</td>
<td>1.59</td>
<td>2.48E-09</td>
<td>3.14E-07</td>
</tr>
<tr>
<td>ligase activity (GO:0016874)</td>
<td>1.59</td>
<td>3.47E-04</td>
<td>2.20E-02</td>
</tr>
<tr>
<td>identical protein binding (GO:0042802)</td>
<td>1.57</td>
<td>1.29E-26</td>
<td>4.39E-24</td>
</tr>
<tr>
<td>kinase binding (GO:0019900)</td>
<td>1.56</td>
<td>4.46E-10</td>
<td>6.38E-08</td>
</tr>
<tr>
<td>protein-macromolecule adaptor activity (GO:0030674)</td>
<td>1.56</td>
<td>1.68E-04</td>
<td>1.22E-02</td>
</tr>
</tbody>
</table>

\(^1\text{Fold enrichment; }^2\text{False discovery rate}
**Table 3.6** Top 30 motifs in differentiating bovine satellite cells

<table>
<thead>
<tr>
<th>Motif Sequence</th>
<th>Motif Name</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAGCTG</td>
<td>Myf5</td>
<td>1e-579</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>MyoG</td>
<td>1e-451</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>Ap4</td>
<td>1e-407</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>Tcf12</td>
<td>1e-395</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>E-box</td>
<td>1e-366</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>Tcf21</td>
<td>1e-332</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>HLH-1</td>
<td>1e-330</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>MyoD</td>
<td>1.00E-299</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Atoh1</td>
<td>1.00E-252</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Jun-AP1</td>
<td>1.00E-234</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>Ascl2</td>
<td>1.00E-212</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Fos</td>
<td>1.00E-203</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Fra2</td>
<td>1.00E-196</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Fra1</td>
<td>1.00E-190</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>AP-1</td>
<td>1.00E-186</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>BATF</td>
<td>1.00E-185</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>JunB</td>
<td>1.00E-184</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>bZIP52</td>
<td>1.00E-179</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Bach2</td>
<td>1.00E-179</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Atf3</td>
<td>1.00E-177</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>KLF5</td>
<td>1.00E-173</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>Sp5</td>
<td>1.00E-160</td>
</tr>
<tr>
<td>TGAATCA</td>
<td>Fosl2</td>
<td>1.00E-159</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>KLF1</td>
<td>1.00E-154</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>Ascl1</td>
<td>1.00E-149</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>KLF14</td>
<td>1.00E-148</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>VIP1</td>
<td>1.00E-144</td>
</tr>
<tr>
<td>CAGCTG</td>
<td>BHLHA15</td>
<td>1.00E-137</td>
</tr>
<tr>
<td></td>
<td>SUT1</td>
<td>1.00E-133</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>Elk1</td>
<td>1.00E-131</td>
</tr>
</tbody>
</table>
Table 3.7 Examples of TFAP4 binding sites (AP4BS)-containing enhancers in differentiating bovine satellite cells

<table>
<thead>
<tr>
<th>Chromatin</th>
<th>Enhancer_start</th>
<th>Enhancer_end</th>
<th>AP4BS_start</th>
<th>AP4BS_end</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr1</td>
<td>43,230,664</td>
<td>43,231,707</td>
<td>72</td>
<td>81</td>
</tr>
<tr>
<td>chr1</td>
<td>96,642,590</td>
<td>96,642,985</td>
<td>334</td>
<td>343</td>
</tr>
<tr>
<td>chr11</td>
<td>10,060,892</td>
<td>10,061,404</td>
<td>498</td>
<td>507</td>
</tr>
<tr>
<td>chr11</td>
<td>27,598,723</td>
<td>27,600,049</td>
<td>238</td>
<td>247</td>
</tr>
<tr>
<td>chr11</td>
<td>99,099,634</td>
<td>99,100,348</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>chr12</td>
<td>18,084,458</td>
<td>18,085,536</td>
<td>1027</td>
<td>1036</td>
</tr>
<tr>
<td>chr12</td>
<td>21,944,882</td>
<td>21,945,802</td>
<td>828</td>
<td>837</td>
</tr>
<tr>
<td>chr13</td>
<td>39,504,844</td>
<td>39,505,104</td>
<td>186</td>
<td>195</td>
</tr>
<tr>
<td>chr13</td>
<td>78,590,251</td>
<td>78,590,365</td>
<td>62</td>
<td>71</td>
</tr>
<tr>
<td>chr14</td>
<td>3,459,375</td>
<td>3,459,865</td>
<td>466</td>
<td>475</td>
</tr>
<tr>
<td>chr14</td>
<td>10,242,746</td>
<td>10,243,159</td>
<td>239</td>
<td>248</td>
</tr>
<tr>
<td>chr14</td>
<td>24,930,473</td>
<td>24,931,783</td>
<td>1081</td>
<td>1090</td>
</tr>
<tr>
<td>chr16</td>
<td>2,569,896</td>
<td>2,570,303</td>
<td>176</td>
<td>185</td>
</tr>
<tr>
<td>chr16</td>
<td>26,470,739</td>
<td>26,472,396</td>
<td>265</td>
<td>274</td>
</tr>
<tr>
<td>chr17</td>
<td>15,609,780</td>
<td>15,611,090</td>
<td>1086</td>
<td>1095</td>
</tr>
<tr>
<td>chr18</td>
<td>25,432,994</td>
<td>25,433,407</td>
<td>248</td>
<td>257</td>
</tr>
<tr>
<td>chr19</td>
<td>22,814,228</td>
<td>22,816,377</td>
<td>996</td>
<td>1005</td>
</tr>
<tr>
<td>chr19</td>
<td>22,814,228</td>
<td>22,816,377</td>
<td>1034</td>
<td>1043</td>
</tr>
<tr>
<td>chr19</td>
<td>45,660,965</td>
<td>45,661,932</td>
<td>336</td>
<td>345</td>
</tr>
<tr>
<td>chr19</td>
<td>53,471,169</td>
<td>53,471,678</td>
<td>248</td>
<td>257</td>
</tr>
<tr>
<td>chr2</td>
<td>79,816,927</td>
<td>79,818,389</td>
<td>213</td>
<td>222</td>
</tr>
<tr>
<td>chr20</td>
<td>56,202,547</td>
<td>56,202,762</td>
<td>49</td>
<td>58</td>
</tr>
<tr>
<td>chr22</td>
<td>646,108</td>
<td>646,698</td>
<td>169</td>
<td>178</td>
</tr>
<tr>
<td>chr22</td>
<td>13,324,716</td>
<td>13,324,922</td>
<td>175</td>
<td>184</td>
</tr>
<tr>
<td>chr22</td>
<td>55,644,181</td>
<td>55,645,326</td>
<td>76</td>
<td>85</td>
</tr>
<tr>
<td>chr23</td>
<td>10,540,538</td>
<td>10,540,695</td>
<td>144</td>
<td>153</td>
</tr>
<tr>
<td>chr23</td>
<td>17,703,775</td>
<td>17,704,362</td>
<td>464</td>
<td>473</td>
</tr>
<tr>
<td>chr</td>
<td>Start Base</td>
<td>End Base</td>
<td>Length</td>
<td>Average Depth</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
<td>-----------</td>
<td>--------</td>
<td>---------------</td>
</tr>
<tr>
<td>chr24</td>
<td>41,235,350</td>
<td>41,236,019</td>
<td>72</td>
<td>81</td>
</tr>
<tr>
<td>chr25</td>
<td>27,202,907</td>
<td>27,203,182</td>
<td>197</td>
<td>206</td>
</tr>
<tr>
<td>chr26</td>
<td>12,561,292</td>
<td>12,561,507</td>
<td>182</td>
<td>191</td>
</tr>
<tr>
<td>chr27</td>
<td>15,001,037</td>
<td>15,001,263</td>
<td>76</td>
<td>85</td>
</tr>
<tr>
<td>chr28</td>
<td>28,308,138</td>
<td>28,308,891</td>
<td>133</td>
<td>142</td>
</tr>
<tr>
<td>chr29</td>
<td>18,646,316</td>
<td>18,647,103</td>
<td>85</td>
<td>94</td>
</tr>
<tr>
<td>chr3</td>
<td>33,253,837</td>
<td>33,254,546</td>
<td>79</td>
<td>88</td>
</tr>
<tr>
<td>chr4</td>
<td>29,230,506</td>
<td>29,232,308</td>
<td>1321</td>
<td>1330</td>
</tr>
<tr>
<td>chr4</td>
<td>39,687,246</td>
<td>39,688,231</td>
<td>861</td>
<td>870</td>
</tr>
<tr>
<td>chr4</td>
<td>48,034,463</td>
<td>48,035,294</td>
<td>63</td>
<td>72</td>
</tr>
<tr>
<td>chr5</td>
<td>26,570,881</td>
<td>26,571,611</td>
<td>89</td>
<td>98</td>
</tr>
<tr>
<td>chr5</td>
<td>68,464,686</td>
<td>68,465,096</td>
<td>170</td>
<td>179</td>
</tr>
<tr>
<td>chr6</td>
<td>38,637,354</td>
<td>38,637,606</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>chr7</td>
<td>84,519,808</td>
<td>84,519,983</td>
<td>40</td>
<td>49</td>
</tr>
<tr>
<td>chr8</td>
<td>31,420,441</td>
<td>31,420,851</td>
<td>381</td>
<td>390</td>
</tr>
<tr>
<td>chr8</td>
<td>81,688,877</td>
<td>81,689,164</td>
<td>131</td>
<td>140</td>
</tr>
<tr>
<td>chr1</td>
<td>56,607,394</td>
<td>56,607,614</td>
<td>4</td>
<td>13</td>
</tr>
</tbody>
</table>
Table 3.8 Examples of biological processes enriched in genes associated with TFAP4 binding sites in differentiating bovine satellite cells

<table>
<thead>
<tr>
<th>Description</th>
<th>GeneRatio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>circulatory system development</td>
<td>67/813</td>
<td>4.22E-07</td>
</tr>
<tr>
<td>heart development</td>
<td>32/813</td>
<td>1.58E-06</td>
</tr>
<tr>
<td>tissue development</td>
<td>94/813</td>
<td>1.73E-06</td>
</tr>
<tr>
<td>regulation of animal organ morphogenesis</td>
<td>16/813</td>
<td>2.04E-06</td>
</tr>
<tr>
<td>muscle organ development</td>
<td>28/813</td>
<td>3.25E-06</td>
</tr>
<tr>
<td>striated muscle tissue development</td>
<td>28/813</td>
<td>3.25E-06</td>
</tr>
<tr>
<td>epithelial cell proliferation</td>
<td>34/813</td>
<td>5.10E-06</td>
</tr>
<tr>
<td>muscle tissue development</td>
<td>29/813</td>
<td>5.44E-06</td>
</tr>
<tr>
<td>regulation of muscle system process</td>
<td>23/813</td>
<td>7.18E-06</td>
</tr>
<tr>
<td>regulation of anatomical structure morphogenesis</td>
<td>56/813</td>
<td>1.30E-05</td>
</tr>
<tr>
<td>anatomical structure morphogenesis</td>
<td>114/813</td>
<td>2.65E-05</td>
</tr>
<tr>
<td>regulation of epithelial cell proliferation</td>
<td>26/813</td>
<td>3.55E-05</td>
</tr>
<tr>
<td>muscle adaptation</td>
<td>13/813</td>
<td>3.88E-05</td>
</tr>
<tr>
<td>epithelium development</td>
<td>56/813</td>
<td>4.42E-05</td>
</tr>
</tbody>
</table>
**FIGURE**

![Graphs and Venn diagrams showing the identification of H3K27ac-marked enhancers in before- and during-differentiation bovine satellite cells.](image)

**Fig 3.1** Identification of H3K27ac-marked enhancers in before-differentiation (BD) and during-differentiation (DD) bovine satellite cells. (A) Pearson correlation analyses of ChIP-seq and Input libraries from two biological replicates. (B) Numbers of H3K27ac-marked enhancers were consistently identified in two experiments. (C) Numbers of H3K27ac-marked enhancers specific to BD or DD bovine satellite cells or common to both BD and DD bovine satellite cells.
Fig 3.2 Examples of H3K27ac-marked enhancers in bovine satellite cells. (A) IGV tracks showing H3K27ac-marked enhancers associated with the myogenin (MYOG) and myosin heavy chain 3 (MYH3) genes in before-differentiation (BD) and during-differentiation (DD) bovine satellite cells. (B) Relative expression levels of MYOG and MYH3 mRNAs in BD and DD bovine satellite cells. *P < 0.05. Gene expression data were retrieved from a previous study [19].
Fig 3.3 Genomic distribution and location of H3K27ac-marked enhancers in bovine satellite cells. (A) Percentages of H3K27ac-marked enhancers located in various genomic regions. BD, before differentiation; DD, during differentiation; UTR, untranslated region. (B) Distances of H3K27ac-marked enhancers from the transcription start sites (TSS). (C) Frequency of proximal H3K27ac-marked enhancers by distance from the TSS.
Fig 3.4 Identification of TFAP4 binding sites in H3K27ac-marked enhancers in differentiating bovine satellite cells. (A) Position weight matrices of two consensus human TFAP4 motifs used to search for TFAP4 binding sites in bovine enhancers (B) Examples of TFAP4 binding sites in H3K27ac-marked enhancers in differentiating bovine satellite cells. (C) Examples of enrichment of TFAP4 binding site and an adjacent TF binding site (in this case, Nr1H2) in H3K27ac-marked enhancers in differentiating bovine satellite cells.
**Fig 3.5** Effect of siRNA-mediated knockdown of TFAP4 on differentiation of bovine satellite cells. Bovine satellite cells were transfected with siRNA targeting TFAP4 mRNA or negative control siRNA and then induced to differentiate and fuse into myotubes. (A) Representative images of cells transfected with TFAP4 siRNA (siTFAP4) and cells transfected with control siRNA (CTRL) on days (D) 0, 1, 2, and 4 of differentiation. (B) Representative images of immunostained bovine satellite cells transfected with TFAP4 siRNA or control siRNA on day 4 of differentiation. MHC staining is shown in green color, and DAPI staining is shown in blue color. (C) Fusion index for bovine satellite cells transfected with TFAP4 siRNA or control siRNA on day 4 of differentiation. Satellite cells from 4 animals were used in these experiments. The fusion index was calculated by counting at least 1000 nuclei per animal. **P < 0.02 (n = 4).**
**Fig 3.6** Effect of TFAP4 knockdown on gene expression in bovine satellite cells. Cells from 3 or 4 animals were transfected with TFAP4 siRNA (siTFAP4) or control siRNA (CTRL) and then cultured in differentiation medium for 3 days. (A) qPCR quantification of TFAP4, MYH2, MYH3, MYOG, CKM, and MYMK mRNAs. *P < 0.05, ** P < 0.002, *** P < 0.001 (n = 4). (B) Western blot analysis of TFAP4, myosin heavy chain (MYC), CKM, and MYOG proteins. β-tubulin was detected as a loading control. (C) Quantification of intensity of western blot bands. *P < 0.05 (n = 3).
Fig 3.7 Effect of TFAP4 overexpression on differentiation of bovine satellite cells. Cells from 6 animals were transfected with a TFAP4 expression plasmid (oeTFAP4) or a control plasmid (CTRL) and induced to differentiate into myotubes. (A) Representative pictures of cells on days (D) 0, 1, 3 and 4 of differentiation. (B) mRNA expression levels of myotube markers MYH2, MYH3, MYOG, CKM, and MYMK as well as TFAP4 * P < 0.05, *** P < 0.001 (n = 6).
Fig 3.8 Effect of TFAP4 overexpression on protein expression of TFAP4, MYH, CKM, and MYOG in bovine satellite cells. Bovine satellite cells from 3 animals transfected with a TFAP4 expression plasmid (oeTFAP4) or a control plasmid (CTRL) were induced to differentiate into myotubes for 4 days. Protein expression was analyzed by Western blotting. (A) Representative images of Western blots. (B) Quantitative analysis of Western blots. * P < 0.05 (n = 3).
Chapter IV

Identification of Enhancers and Transcription Factors that Control Postnatal Skeletal Muscle Development and Growth in Cattle

ABSTRACT

While numerous studies have focused on muscle development and mass accumulation during adulthood, limited research has explored the molecular and cellular mechanisms driving the transition from neonatal to adult muscle. In this study, we first performed RNA-sequencing (RNA-seq) to compare gene expression profiles in skeletal muscle between neonatal calves and adult steers. Analyzing the RNA-seq data showed that 924 genes were downregulated and 1,021 upregulated (log2fold change >1, Padj < 0.05) from calf to steer muscle. Among genes downregulated in steer muscle were myosin heavy chain 3 (MYH3) and MYH8, and among genes upregulated in steer muscle were MYH7 and myoglobin (MB). Surprisingly, the expression levels of many of the so-called adult muscle genes, such as MYH1 and MYH2, were not different between calf and steer muscles. Gene ontology (GO) analysis showed that many of the genes downregulated in steer muscle are involved in the processes related to protein synthesis and glycolysis and that many of the genes upregulated in steer muscle function in blood vessel development and immune cell activation. We next conducted chromatin immunoprecipitation followed by sequencing (ChIP-seq) to identify genomic regions marked with H3K27ac (acetylation at lysine 27 on the H3 histone protein), i.e., active enhancers, from skeletal muscle of neonatal calves and adult steers. Analyzing the ChIP-seq data found 20,163 enhancers specifically active in calf muscle, 14,909 enhancers specifically active in steer muscle, and 27,002 enhancers active in both calf and steer muscle. Motif enrichment analysis revealed that calf-specific enhancers were enriched with binding sites for transcription factors including the Krüppel-
like (KLF) family and transcriptional enhanced associate domain (TEAD) family transcription factors, that steer-specific enhancers were enriched with binding sites for transcription factors including the Forkhead box class O (FOXO) family and SMAD family transcription factors 3, and that enhancers active in both calf and steer muscle were enriched with binding sites for transcription factors including the myogenic regulatory factor (MRF) family and myocyte enhancer factor 2 (MEF2) family transcription factors. These results shed light on the differences in gene expression and biological processes in skeletal muscle between newborn calves and adult steers. These results also shed light on the transcriptional mechanisms that control these differences, such as enhancers and transcription factors.

**Keywords** Calf, steer, skeletal muscle, transcription factor, enhancer
INTRODUCTION

Skeletal muscle development can be divided into embryonic, fetal, neonatal, and postnatal stages. In most mammals, myogenesis, the formation of muscle fibers or myofibers from myoblasts, occurs primarily during the embryonic and fetal stages. Myogenesis during the embryonic and fetal stages is controlled mainly at the transcriptional level by 4 Myogenic Regulatory Factors (MRFs), including myogenic factor 5 (MYF5), myogenic factor 6 (MYF6 or MRF4), myoblast determination protein 1 (MYOD1), and myogenin (MYOG) [1-4]. MRFs exert their transcriptional function by binding to the E-boxes as a heterodimer with E-proteins, and the consensus sequence of E-boxes is CANNTG, where N = any nucleotide [5]. Many myogenic genes possess E-boxes in their enhancer and promoter regions [6]. MYF5 and MYOD1 are important for early myoblast determination [2, 7, 8]. MYOG is required for myoblast terminal differentiation [9]. Apart from these 4 MRFs, the myocyte enhancer factor-2 family (MEF2) of transcription factors, consisting of MEF2A, B, C, and D [10], also play a crucial role in gene expression during myogenesis. Despite having unique binding sites [11], MEF2s exert their transcriptional function by interacting with MRFs [12]. In postnatal muscle development, MEF2s also govern muscle mass accumulation [13].

There are minimal changes in the number of myofibers during the postnatal stage [14]. Muscle mass gain during the postnatal stage mainly results from myofiber hypertrophy [14, 15]. One process contributing to postnatal muscle hypertrophy involves satellite cells, which are skeletal muscle stem cells in postnatal animals. Satellite cells play an essential role in muscle regeneration. In addition, satellite cells can fuse with existing myofibers to increase myofiber size [16]. Protein synthesis and degradation is another major process determining postnatal muscle mass gain. The mTOR signaling pathway plays a central role in controlling protein synthesis and degradation in skeletal muscle [17].
While considerable understanding has been gained about the transcription factors and regulatory DNA sequences that control gene expression in embryonic and fetal muscle, little is known about these factors and sequences that govern gene expression in postnatal skeletal muscle. Therefore, the aim of this study was to identify transcription factors and regulatory DNA sequences that control gene expression in the skeletal muscle of postnatal cattle. We achieved this goal by analyzing the transcriptomes and mapping the enhancers active in the skeletal muscle of both neonatal and adult cattle.
MATERIALS AND METHODS

Animals and sample collection

Longissimus muscle samples were collected from 5 finishing Angus-cross steers and five male neonatal (four one-day-old and one 2-day-old) Angus-cross calves at slaughter. Muscle samples were immediately frozen in liquid nitrogen and stored in a –80 ºC freezer.

RNA extraction

Muscle samples were homogenized in Tri-reagent (catalog# R2050-1-200, ZYMO RESEARCH, Irvine, California, USA) using a polytron homogenizer at high power output for 30 seconds followed by centrifugation at 5,000 ×g and room temperature for 6 min to remove debris. Total RNA was then extracted from the homogenate using the Direct-zol RNA Microprep Kit (R2062, ZYMO RESEARCH). RNA concentration and purity were determined by Nanodrop One (840274200, ThermoScientific, Waltham, MA, USA). RNA integrity was measured by Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA samples with RNA integrity numbers (RIN) > 7 were used for RNA-seq.

RNA-seq library construction, sequencing, and data analysis

RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (E7760S, New England BioLabs, Ipswich, MA, USA). All libraries were paired-end sequenced on an Illumina sequencing system at Novogene Corporation (Sacramento, CA, USA). Raw fastq files of RNA-seq were filtered to remove low-quality sequences. Clean reads were mapped to the bovine reference genome BosTaurus9 using HISAT2 (2.0.5) [18]. The gene expression level was calculated as the number of fragments per kilobase of transcript sequence per millions of base pairs sequenced (FPKM) using featureCounts (v1.5.0-p3)[19]. Differentially expressed genes (DEGs) were detected using
DESeq2 [20]. The gene ontology analysis (GO analysis) was performed using DEGs under the cutoff (p-value (Padj) < 0.05 and log2fold change > 1) by the R package, clusterProfiler (4.8.1) [21, 22]. The pipeline is shown in Fig. 4.1 A and B.

**RT-qPCR and data analysis**

One microgram of total RNA was denatured by incubation with random primers at 70 ºC for 10 min, followed by cooling on ice for 5 min. Reverse transcription of total RNA was performed by incubating 5 µL RNA–random primer mix with 4 µL 5 × reverse transcription buffer, 4.8 µL MgCl2, 1 µL dNTP, 1 µL ImProm-II reverse transcriptase, 0.5 µL RNasin Ribonuclease Inhibitor, and 3.7 µL RNase free H2O at 42 ºC for 90 min, at 70 ºC for 10 min, and then at 4 ºC for 5 min. Quantitative PCR was performed in duplicate using Fast SYBR Green Master Mix on a 7500 Fast Real-Time PCR system (Applied Biosystems). Relative gene expression levels were calculated using the ∆∆Ct method [42], using the HMBS gene as a reference gene.

**Chromatin immunoprecipitation (ChIP) assay and ChIP-seq library construction and sequencing**

Muscle tissue was immersed in 1% formaldehyde (F8775, Sigma-Aldrich, St. Louis, MO, USA) and continuously cut into small pieces at room temperature for 15 minutes. Cross-linking was quenched by adding 0.125 M glycine (ThermoFisher Scientific, Hampton, NH, USA) and incubating for 10 minutes at room temperature. To isolate nuclei, muscle tissue was incubated with pre-cold hypotonic buffer N (Hepes 10mM, MgCl2 2mM, KCl 25mM) on ice for 30 minutes and then homogenized first with a pestle A and then pestle B in a Dounce homogenizer. After adding 0.2M sucrose, the lysates were centrifuged at 1000 g and 4ºC for 10 min. The nuclei were resuspended in buffer N (Hepes 10mM, MgCl2 2mM, KCl 25mM, sucrose 250mM) and centrifuged again. Finally, the nuclei were resuspended in a sonication buffer from the iDeal ChIP-seq kit for histones (C01010051, Diagenode...
Liège, Liege, Belgium). Chromatin fragmentation was performed on ice using a sonicator at power 3, with 20 seconds on and 40 seconds off for 5 cycles. Chromatin immunoprecipitation was carried out using 2 μg of H3K27ac antibody (ab1791, Abcam, Waltham, Boston, USA) per ChIP and the iDeal ChIP-seq kit according to the manufacturer’s instructions. The ChIP-seq library was prepared using the NEBNext ChIP-Seq Library Prep Reagent Set for Illumina (New England BioLabs). Briefly, the ends of DNA fragments from the ChIP assay were repaired using T4 DNA polymerase, Klenow DNA polymerase, and T4 polynucleotide kinase. The resulting end-repaired DNA fragments were subjected to 3′ dA overhang addition using exonuclease minus Klenow DNA polymerase and dATP. The dA-tailed DNA fragments were ligated to sequencing adaptors. Subsequently, DNA fragments of approximately 300 bp were selectively enriched from the adaptor-ligated DNA using AMPure XP Beads. The size-selected DNA fragments were then amplified with 12 cycles of PCR, employing an index primer and a universal PCR primer. ChIP-seq libraries were paired-end sequenced on an Illumina sequencing system at Novogene Corporation (Sacramento, CA, USA).

**ChIP-seq data analysis**

Raw ChIP-seq sequencing data were processed using Trimmomatic (0.39) [23] and SAMtools (1.9) [24] to remove the sequencing adaptor and low-quality sequences. ChIP-seq sequences were mapped to the reference bovine genome BosTaurus9 using HISAT2 (2.2.0) [18]. Quality control of uniquely mapped ChIP-seq data was performed using phantompeakqualtools [25, 26]. Uniquely mapped ChIP-seq sequences were used for peak calling, performed using Macs3 (3.0.0b1) [27]. Differential ChIP-seq peaks were detected using bdgdiff in Macs3. Peak annotation was done using the annotatePeak function of the ChIPseeker package (1.36.0) [28]. Motif enrichment analysis of ChIP-seq peaks was performed using Homer [29]. The ChIP-seq data analysis pipeline is shown in Fig. 4.1 C to H.
RESULTS

Gene expression profiles in calf and steer muscle

RNA-seq of mRNAs in the skeletal muscle of five newborn calves and five adult steers detected 18,819 genes in these samples. Based on a principal component analysis (PCA) (Fig. 4.2 A), the RNA-seq data from 5 calves clustered as a group while those from 5 steers clustered as the other group, indicating significant differences in transcriptome between calf and steer muscle. Applying thresholds of the adjusted p-value (Padj) < 0.05 and the absolute value of log2fold change > 1, 1,945 DEGs were identified between calf and steer muscle. Among these DEGs, 924 were down-regulated, and 1,021 were up-regulated from calf to steer, indicating significant changes in muscle transcriptome from newborn calf to adult steer. The volcano plot in Fig. 4.2 B illustrates genes up-regulated and down-regulated from newborn calf to adult steer. Examples of genes down-regulated in muscle from calf to steer were myosin heavy chain 3 (MYH3), myosin heavy chain 8 (MYH8), myoblast determination protein 1 (MYOD1), and insulin-like growth factor 2 (IGF2). Examples of genes up-regulated from calf to steer were integrin beta 5 (ITGB5), myoglobin (MB), and myosin light chain kinase 4 (MYLK4). Based on the PCA, there was more variation between calf muscle transcriptomes than between steer muscle transcriptomes (Fig. 4.2 A). This suggests that gene expression in muscle is more dynamic in newborn calves than in adult steers. To validate the RNA-seq data, we measured the expression levels of ten muscle-specific genes, including MB, MYH1, MYH2, MYH3, MYH7, MYH8, MYOD1, MYOG, MYL1, and CKM using RT-qPCR. As shown in Fig. 4.2 C, fold changes (calf/steer) in the expression of these genes detected by RT-qPCR were highly correlated (R²=0.93) with those detected by RNA-seq.

Functions enriched in genes differentially expressed between calf and steer muscle

To understand the functions of genes differentially expressed between calf and steer muscle,
we conducted Gene Ontology (GO) analyses on these genes. The top 20 GO terms (ranked by the number of genes contained) enriched in genes upregulated in calf muscle include terms related to protein synthesis (translation, RNA processing, mRNA splicing, etc.) and carbohydrate metabolism (carbohydrate derivative metabolic process, monosaccharide metabolic process, glycolytic process, pyruvate metabolic process, etc.) (Fig. 4.3). The top 20 GO terms enriched in genes upregulated in steer muscle include terms related to the muscle development process (muscle system process, skeletal system development, muscle adaptation, etc.), vessel development process (circulatory system development, blood vessel development, vasculature development, angiogenesis, blood vessel morphogenesis, etc.), and immune activation (immune response, adaptive immune response, response to chemokine, etc.) (Fig. 4.3).

**Active enhancers in calf and steer muscle**

Active enhancers are genomic regions actively regulating gene expression and are marked with H3K27ac [30, 31]. We identified these genomic regions in the skeletal muscle of 3 newborn calves and 3 adult steers by ChIP-seq. We constructed two libraries from each sample, one using DNA immunoprecipitated with an H3K27ac antibody and the other with input DNA. At least 10 million uniquely mapped reads were generated from each library, and the average unique mapping rate was approximately 70% (Table 4.1). We assessed the quality of these ChIP-seq sequences with phantompeakqualtools [26]. As shown in Fig. 4.4 A, the ChIP-seq data from steer samples S2 and S3 and calf samples C2, C3, and C4 exhibited NSC >1.1 and RSC >0.8 values, indicating that they met the quality standards for ChIP-seq data [25]. Therefore, the ChIP-seq data from steer sample S1 did not meet these standards and were excluded from further analysis. Using Macs3 20,163, H3K27ac peaks were uniquely identified from calf muscle; 14,909 H3K27ac peaks were uniquely identified from steer muscle; and 27,002 H3K27ac peaks were identified from both calf and steer muscle (Fig. 4.4 B).
Genomic distribution of enhancers active in calf muscle or steer muscle or both

More than 90% of active enhancers identified from calf and steer muscle were located in the promoters, introns, or intergenic regions of the genome, while less than 10% of them were in the 5’UTR, 3’UTR, or exons (Fig. 4.5 A). In terms of distance from the transcription start site (TSS), calf muscle- or steer muscle-specific active enhancers showed a unimodal distribution, peaking at ~100 kb from the TSS. In contrast, active enhancers shared by both calf and steer muscle showed a bimodal distribution, peaking at ~1,000 bp and ~100 kb from the TSS (Fig. 4.5 B). Within 3 kb from the TSS, active enhancers unique to steer muscle or common to both steer and calf muscle concentrated at ~500 bp upstream and downstream from the TSS, whereas those unique to calf muscle concentrated at the TSS (Fig. 4.5 C and D).

Genes associated with enhancers active in calf muscle or steer muscle or both

A gene nearest to an enhancer is considered to have the greatest chance of being regulated by the enhancer [32]. Annotating the enhancers identified above showed that 6,838 genes were associated with calf muscle-specific active enhancers, that 4,680 genes were associated with steer muscle-specific active enhancers, and that 8,451 genes were associated with enhancers active in both calf and steer muscle. Fig. 4.6 A-C shows examples of these genes and the associated active enhancers. As demonstrated by this figure, MHY3 (Fig. 4.6 A), MB (Fig. 4.6 B), and MYH7 (Fig. 4.6 C) were each associated with multiple enhancers that were active in calf or steer muscle or in both. However, the MYH3 gene was associated with more active enhancers in calf than steer muscle, and the MB and MYH7 genes were associated with more active enhancers in steer than calf muscle.

Motifs enriched in enhancers active in calf or steer muscle or both

Typical enhancers are enriched with transcription factor binding sites or motifs. We next
identified transcription factor binding sites enriched in enhancers active in calf muscle, steer muscle, or both. Using p-value < 0.001 as the cutoff, 462, 333, and 474 motifs were enriched in enhancers active in calf muscle, steer muscle, or both calf and steer muscle, respectively. Tables 4.2, 4.3, and 4.4 show the top 10 (ranked by p-value) of these motifs. The top 10 motifs enriched in calf muscle-specific enhancers included the binding sites for the ETS family of transcription factors (ELF1, ELK1, and ELK4), and transcription factors MEF2A, THAP11, and NRF1, indicating that these are the major transcription factors regulating gene transcription in calf muscle. The top 10 motifs enriched in steer muscle-specific enhancers included four MEF2 family transcription factors (MEF2A, MEF2B, MEF2C, and MEF2D), two Forkhead family members (FOXP1 and FOXO3), TFAP4, and GABPA, suggesting these are the major transcription factors regulating gene expression in steer muscle. Top motifs enriched in enhancers active in both calf and steer muscle included the 4 MEF2 family transcription factors and transcription factors ELK1, ELK4, ELF1, and SP5, indicating that these transcription factors are the major transcription factors regulating gene expression in both calf and steer muscle.

**Transcription factors driving differential gene expression between calf and steer muscle**

We associated the genes differentially expressed with the enhancers differentially activated between calf and steer muscle to identify the transcription factors controlling differential gene expression between calf and steer muscle. As shown by Fig. 4.7, 1,247 of 1,886 genes upregulated in calf muscle were associated with 9,657 enhancers differentially activated in calf muscle; 569 of 1,021 genes upregulated in steer muscle were associated with 9,256 enhancers differentially activated in steer muscle.

We next searched for motifs enriched in enhancers that were both differentially activated and associated with genes upregulated in calf or steer muscle. Using p-value < 0.001 as
the cutoff, 363 and 206 motifs were enriched in enhancers associated with genes upregulated in calf and steer muscle, respectively. As shown in Table 4.5, the top 10 (ranked by p-value) motifs enriched in enhancers associated with genes upregulated in calf muscle were binding sites for the ETS family transcription factors ELF1, ELK1, and ELK4, the MEF2 family transcription factors MEF2A and MEF2B, and transcription factors SIX2 and BMYB. These results suggest that these are the major transcription factors upregulating gene expression in calf muscle. The top 10 motifs enriched in enhancers associated with genes upregulated in steer muscle included binding sites for the MEF2 family transcription factors MEF2D, MEF2C, MEF2A, and MEF2B, binding sites for MYOG, MYF5, and TFAP4, and binding sites for FOXO1 and FOXP1 (Table 4.6). These results suggest that these transcription factors are the major transcription factors that upregulate gene expression in steer muscle.
DISCUSSION

The physical differences between newborn calf and adult steer muscle are evident as the latter looks redder and feels firmer. Biochemically, the neonatal muscle is thought to contain MYH3 and MYH8 as the main myosin heavy chain isoforms [33, 34]. In contrast, the adult muscle is believed to express MYH1, MYH2, and MYH7 as the main myosin heavy chain isoforms [35]. Surprisingly, the RNA-seq data from this study showed that the mRNA levels of MYH1, MYH2, and many other so-called adult muscle genes (e.g., CKM, TNNT1, TNNC2, MYL1, etc.) are not much different between neonatal calf and adult steer muscle. This finding suggests that from the standpoint of expression of muscle-specific genes, the neonatal calf muscle is more mature than traditionally thought. However, the RNA-seq data from this study showed significant differences in gene expression between calf and steer muscle. More than 900 genes were expressed at higher levels in calf muscle than in steer muscle, and more than 1000 genes were expressed at higher levels in steer muscle than in calf muscle. Among the skeletal muscle-specific MYH genes (i.e., MYH1, 2, 3, 4, 7, and 8), MYH3 and MYH8 were downregulated from calf to steer muscle, while MYH7 was upregulated from calf to steer muscle. Downregulation of MYH3 and MYH8 from the neonatal to adult stage is consistent with characterizing these two genes as embryonic and perinatal MYHs, respectively [36]. Because MYH7 is the primary MYH isoform associated with slow oxidative fibers, increased expression of MYH7 in adult steer muscle suggests that adult steer muscle is more oxidative than neonatal calf muscle [37].

Functional analyses of genes upregulated in calf muscle showed that many of them are involved in protein synthesis-related processes, including peptide biosynthesis, translation, and regulation of RNA splicing. This result suggests that the rate of protein synthesis is higher in neonatal calf muscle than in adult steer muscle, which was also reported before [38, 39]. Based on the functional analyses, genes involved in the glycolytic and pyruvate metabolic processes are upregulated in calf muscle compared to steer muscle. This result
suggests that compared to adult steers, neonatal calves rely more on glycolysis for energy production in muscle.

Functional analysis of genes upregulated in steer muscle indicated that genes functioning in the muscle development process, muscle system process, and muscle adaptation are upregulated in adult steer muscle. This result suggests muscle development continues in adulthood, perhaps in response to changed nutrition, physical activity, or environment. Functional analyses of genes upregulated in adult steer muscle also indicated that many genes involved in processes related to blood vessel development are upregulated in adult steer muscle compared to their expression in the neonatal calf muscle. This result suggests that adult steer muscle is more angiogenic and contains more blood vessels than neonatal calf muscle, which is reflected by the fact that adult steer muscle looks much redder than neonatal calf muscle. Being more angiogenic and containing more blood vessels are also consistent with the fact that adult muscle contains more slow, oxidative fibers that require more blood supply than young muscle [40]. Functional analyses of genes differentially expressed between calf and steer muscle revealed that genes involved in immune response-related processes are upregulated in adult steer muscle. Increased immune responses in adult cattle muscle could be induced by the high rate of muscle regeneration in adult animals because macrophages and other immune cells are needed to remove aging or injured muscle cells [41, 42]. Increased production of cytokines such as IL-6 by immune cells in adult muscle might directly aid muscle regeneration [43].

This study identified tens of thousands of enhancers active in the neonatal calf or adult steer muscle. Motif analyses of these enhancers revealed that they are enriched with binding sites for hundreds of transcription factors, suggesting that various transcription factors participate in gene regulation in cattle muscle. The motif analyses indicated that the MEF2 family transcription factors are the principal transcription factors that control gene expression in both calf and steer muscle. The motif analyses also indicated that the KLF
family transcription factors, KLF1, KLF3, KLF14, and KLF5, also control gene expression in both calf and steer muscle. Although KLF3 and KLF5 have been reported to have positive effects on skeletal muscle differentiation [44, 45], the KLF family transcription factors are better known for their roles in controlling gene expression in smooth muscle [46, 47] and blood vessel development [48]. It remains to be determined whether these transcription factors target skeletal muscle-specific genes or smooth muscle-specific genes in skeletal muscle, which contains vascular smooth muscle in addition to skeletal muscle fibers.

The motif analyses revealed the likely enhancers and transcription factors that interact to drive and maintain the transcriptome in adult steer muscle. These transcription factors include the Forkhead family transcription factors, which have traditionally been associated with muscle atrophy and autophagy [49, 50]. However, the Forkhead transcription factors also promote angiogenesis [51, 52]. This suggests that the Forkhead family transcription factors might have dual functions in adult steer muscle: on the one hand, they control gene expression involved in skeletal muscle development, and on the other hand, they regulate gene expression related to angiogenesis. Another interesting revelation by the motif analyses was that 80% of enhancers active in steer muscle contain the binding site for SMAD3, which is an effector of the TGF-β signaling pathway [53]. SMAD3 and the TGF-β signaling pathway are known as negative regulators of skeletal muscle development and growth [54-56], but they also promote smooth muscle development within blood vessels [57]. Thus, SMAD3 might be another transcription factor with dual functions in adult steer muscle.
REFERENCE


47. King KE, Iyemere VP, Weissberg PL, Shanahan CM: **Kruppel-like factor 4**


### Table 4.1 Mapping summary of ChIP-seq data

<table>
<thead>
<tr>
<th>Library ID</th>
<th>ChIP antibody</th>
<th>Total align</th>
<th>Mapping in Paired</th>
<th>Mapping Rate</th>
<th>Unique Mapping</th>
<th>Unique Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>bC2IP</td>
<td>H3K27ac</td>
<td>74634838</td>
<td>70967695</td>
<td>95.09%</td>
<td>56350762</td>
<td>75.50%</td>
</tr>
<tr>
<td>bC2IN</td>
<td>Input</td>
<td>70785696</td>
<td>69326310</td>
<td>97.94%</td>
<td>53759772</td>
<td>75.95%</td>
</tr>
<tr>
<td>bC3IP</td>
<td>H3K27ac</td>
<td>91002585</td>
<td>89137870</td>
<td>97.95%</td>
<td>78469686</td>
<td>86.23%</td>
</tr>
<tr>
<td>bC3IN</td>
<td>Input</td>
<td>60009312</td>
<td>58737829</td>
<td>97.88%</td>
<td>46235274</td>
<td>77.05%</td>
</tr>
<tr>
<td>bC4IP</td>
<td>H3K27ac</td>
<td>76190229</td>
<td>74451468</td>
<td>97.72%</td>
<td>63409290</td>
<td>83.22%</td>
</tr>
<tr>
<td>bC4IN</td>
<td>Input</td>
<td>68916878</td>
<td>67484943</td>
<td>97.92%</td>
<td>52863594</td>
<td>76.71%</td>
</tr>
<tr>
<td>bS1IP</td>
<td>H3K27ac</td>
<td>29630126</td>
<td>28190421</td>
<td>95.14%</td>
<td>19452942</td>
<td>65.65%</td>
</tr>
<tr>
<td>bS1IN</td>
<td>Input</td>
<td>19943410</td>
<td>18867670</td>
<td>94.61%</td>
<td>10748624</td>
<td>53.90%</td>
</tr>
<tr>
<td>bS2IP</td>
<td>H3K27ac</td>
<td>22423098</td>
<td>21141996</td>
<td>94.29%</td>
<td>15386996</td>
<td>68.62%</td>
</tr>
<tr>
<td>bS2IN</td>
<td>Input</td>
<td>90077221</td>
<td>85444241</td>
<td>94.86%</td>
<td>56340014</td>
<td>62.55%</td>
</tr>
<tr>
<td>bS3IP</td>
<td>H3K27ac</td>
<td>150461472</td>
<td>145764403</td>
<td>96.88%</td>
<td>117746282</td>
<td>78.26%</td>
</tr>
<tr>
<td>bS3IN</td>
<td>Input</td>
<td>123583006</td>
<td>119989365</td>
<td>97.09%</td>
<td>86844098</td>
<td>70.27%</td>
</tr>
</tbody>
</table>
Table 4.2 Top 10 motifs enriched in calf muscle-specific enhancers

<table>
<thead>
<tr>
<th>Sequence</th>
<th>motif</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCCGAAGT</td>
<td>ETS</td>
<td>1.00E-188</td>
</tr>
<tr>
<td>ACTAACCTTCCCAGAATGC</td>
<td>Ronin</td>
<td>1.00E-187</td>
</tr>
<tr>
<td>ACTTCCGGT</td>
<td>Elk1</td>
<td>1.00E-168</td>
</tr>
<tr>
<td>CTCGCATGCGGC</td>
<td>NRF1</td>
<td>1.00E-154</td>
</tr>
<tr>
<td>CAAGATGGCCGGC</td>
<td>YY1</td>
<td>1.00E-152</td>
</tr>
<tr>
<td>ACCCGAAGT</td>
<td>ELF1</td>
<td>1.00E-145</td>
</tr>
<tr>
<td>ACTTCCGGT</td>
<td>Elk4</td>
<td>1.00E-145</td>
</tr>
<tr>
<td>ACTAACCTTCCCAGAATGC</td>
<td>GFY-Staf</td>
<td>1.00E-137</td>
</tr>
<tr>
<td>CTAATAATAG</td>
<td>Mef2a</td>
<td>1.00E-126</td>
</tr>
<tr>
<td>CGCGCATGCGC</td>
<td>NRF</td>
<td>1.00E-124</td>
</tr>
</tbody>
</table>
Table 4.3 Top 10 motifs enriched in steer muscle-specific enhancers

<table>
<thead>
<tr>
<th>Sequence</th>
<th>motif</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Sequence Image" /></td>
<td>Mef2d</td>
<td>1.00E-91</td>
</tr>
<tr>
<td><img src="image2" alt="Sequence Image" /></td>
<td>Mef2c</td>
<td>1.00E-88</td>
</tr>
<tr>
<td><img src="image3" alt="Sequence Image" /></td>
<td>Mef2b</td>
<td>1.00E-84</td>
</tr>
<tr>
<td><img src="image4" alt="Sequence Image" /></td>
<td>Mef2a</td>
<td>1.00E-83</td>
</tr>
<tr>
<td><img src="image5" alt="Sequence Image" /></td>
<td>bHLH18</td>
<td>1.00E-64</td>
</tr>
<tr>
<td><img src="image6" alt="Sequence Image" /></td>
<td>GABPA</td>
<td>1.00E-63</td>
</tr>
<tr>
<td><img src="image7" alt="Sequence Image" /></td>
<td>Ets1-distal</td>
<td>1.00E-59</td>
</tr>
<tr>
<td><img src="image8" alt="Sequence Image" /></td>
<td>FOXP1</td>
<td>1.00E-58</td>
</tr>
<tr>
<td><img src="image9" alt="Sequence Image" /></td>
<td>Ap4</td>
<td>1.00E-58</td>
</tr>
<tr>
<td><img src="image10" alt="Sequence Image" /></td>
<td>Foxo3</td>
<td>1.00E-55</td>
</tr>
</tbody>
</table>
Table 4.4 Top 10 motifs enriched in enhancers shared by both calf and steer muscle

<table>
<thead>
<tr>
<th>Sequence</th>
<th>motif</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTAATTGGAC</td>
<td>Mef2d</td>
<td>1.00E-178</td>
</tr>
<tr>
<td>GAATTCCGCTT</td>
<td>Elk1</td>
<td>1.00E-174</td>
</tr>
<tr>
<td>GAATTCCGCTT</td>
<td>Elk4</td>
<td>1.00E-166</td>
</tr>
<tr>
<td>CTAATAAG</td>
<td>Mef2a</td>
<td>1.00E-158</td>
</tr>
<tr>
<td>CTAATTGGAC</td>
<td>Mef2b</td>
<td>1.00E-156</td>
</tr>
<tr>
<td>ACCGGAAGT</td>
<td>ELF1</td>
<td>1.00E-151</td>
</tr>
<tr>
<td>CTAATAAG</td>
<td>Mef2c</td>
<td>1.00E-150</td>
</tr>
<tr>
<td>AGGGCCGGAC</td>
<td>Sp5</td>
<td>1.00E-147</td>
</tr>
<tr>
<td>ACCGGAAGT</td>
<td>ETS</td>
<td>1.00E-134</td>
</tr>
<tr>
<td>ACCGGAAGT</td>
<td>GABPA</td>
<td>1.00E-125</td>
</tr>
</tbody>
</table>
**Table 4.5** Top 10 motifs enriched in enhancers associated with genes upregulated in calf muscle

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Motif</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="sequence1.png" alt="Sequence" /></td>
<td>Elk1</td>
<td>1.00E-49</td>
</tr>
<tr>
<td><img src="sequence2.png" alt="Sequence" /></td>
<td>Mef2a</td>
<td>1.00E-46</td>
</tr>
<tr>
<td><img src="sequence3.png" alt="Sequence" /></td>
<td>Elk4</td>
<td>1.00E-43</td>
</tr>
<tr>
<td><img src="sequence4.png" alt="Sequence" /></td>
<td>Mef2b</td>
<td>1.00E-38</td>
</tr>
<tr>
<td><img src="sequence5.png" alt="Sequence" /></td>
<td>ETS</td>
<td>1.00E-37</td>
</tr>
<tr>
<td><img src="sequence6.png" alt="Sequence" /></td>
<td>NRF1</td>
<td>1.00E-35</td>
</tr>
<tr>
<td><img src="sequence7.png" alt="Sequence" /></td>
<td>Mef2d</td>
<td>1.00E-35</td>
</tr>
<tr>
<td><img src="sequence8.png" alt="Sequence" /></td>
<td>Six2</td>
<td>1.00E-32</td>
</tr>
<tr>
<td><img src="sequence9.png" alt="Sequence" /></td>
<td>ELF1</td>
<td>1.00E-32</td>
</tr>
<tr>
<td><img src="sequence10.png" alt="Sequence" /></td>
<td>BMYB</td>
<td>1.00E-29</td>
</tr>
</tbody>
</table>
**Table 4.6** Top 10 motifs enriched in enhancers associated with genes upregulated in steer muscle

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Motif</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTAATTTTGGC</td>
<td>Mef2b</td>
<td>1.00E-35</td>
</tr>
<tr>
<td>CTAATAATTAG</td>
<td>Mef2c</td>
<td>1.00E-31</td>
</tr>
<tr>
<td>CTAATTTTAGG</td>
<td>Mef2d</td>
<td>1.00E-27</td>
</tr>
<tr>
<td>CTAATAATAG</td>
<td>Mef2a</td>
<td>1.00E-24</td>
</tr>
<tr>
<td>TAACAGCTGT</td>
<td>Myf5</td>
<td>1.00E-15</td>
</tr>
<tr>
<td>CTAGCTGTCA</td>
<td>bZIP52</td>
<td>1.00E-15</td>
</tr>
<tr>
<td>CATTACAT</td>
<td>FOXK2</td>
<td>1.00E-12</td>
</tr>
<tr>
<td>CCCCCGCC</td>
<td>SUT1</td>
<td>1.00E-12</td>
</tr>
<tr>
<td>TAAATCGC</td>
<td>BMYB</td>
<td>1.00E-12</td>
</tr>
<tr>
<td>TAAATCGC</td>
<td>Pitx1</td>
<td>1.00E-12</td>
</tr>
</tbody>
</table>
Fig 4.1 Flowchart of experimental design and data analysis
Fig 4.2 RNA-seq analysis of transcriptomes in calf and steer muscle. (A) Principal component analysis (PCA) of RNA-seq data from 5 calf (blue dots) and 5 steer (red dots) muscle samples. (B) Volcano plot of RNA-seq data. Criteria for genes upregulated (UP) and downregulated (DOWN) in calf muscle versus steer muscle were $P_{adj} < 0.05$ and $\mid \log_{2}\text{fold change} \mid > 1$. NO means genes that were not differentially expressed between calf and steer muscle. (C) Comparison of expression of selected genes quantified by RT-qPCR and RNA-seq.
Fig 4.3 Top 20 biological processes enriched in genes upregulated in calf or steer muscle.
Fig 4.4 Identification of H3K27ac-marked enhancers in calf and steer muscle by ChIP-seq. (A) Quality control analysis of ChIP-seq data from 3 calves and 3 steers by the phantompeakqualtools. (B) Numbers of H3K27ac-enriched genomic regions (“peaks”) uniquely identified from calf muscle (calf), uniquely identified from steer muscle (steer), and identified from both calf and steer muscle (common).
**Fig 4.5** Genomic distribution of calf muscle-specific, steer muscle-specific, and common H3K27ac-marked enhancers. (A) Percentages of enhancers located in various genomic regions. (B) Frequency of enhancers by distance from the transcription start site (TSS). (C) and (D) Distribution of enhancers within 3000 bp from the TSS.
**Fig 4.6** Examples of H3K27ac-marked enhancers ("peaks") identified from calf or steer muscle or both calf and steer muscle. (A) Enhancers associated with the myosin heavy chain 3 (MYH3) gene. (B) Enhancers associated with the myosin heavy chain 7 (MYH7) gene. (C) Enhancers associated with the myoglobin (MB) gene. Peaks specific to calf muscle, specific to steer muscle, and common to both calf and steer muscle are indicated in blue, red, and gray, respectively.