Regulation of myogenic gene expression

Cristina Vicente-García, Juan Diego Hernández-Camacho, Jaime J. Carvajal

Abstract

Skeletal muscle development and regeneration is governed by the combined action of Myf5, MyoD, Mrf4 and MyoG, also known as the myogenic regulatory factors (MRFs). These transcription factors are expressed in a highly spatio-temporal restricted manner, ensuring the significant functional and metabolic diversity observed between the different muscle groups. In this review, we will discuss the multiple layers of regulation that contribute to the control of the exquisite expression patterns of the MRFs in particular, and of myogenic genes in general. We will highlight all major regulatory processes that play a role in myogenesis: from those that modulate chromatin status and transcription competence, such as DNA methylation, histone modification, chromatin remodeling, or non-coding RNAs, to those that control transcript and protein processing and modification, such as alternative splicing, polyadenylation, other mRNA modifications, or post-translational protein modifications. All these processes are exquisitely and tightly coordinated to ensure the proper activation, maintenance and termination of the myogenic process.

1. Introduction

Mammalian skeletal muscles arise from different mesodermal populations located in the prospective head, neck and caudal-most region of the embryo. In each of these locations, specific inductive signals trigger the myogenic cascade, which is coordinated by four bHLH transcription factors (TFs) known as myogenic regulatory factors, or MRFs [1]. Myf5, Mrf4 and MyoD function as determination factors in certain muscle precursors and very often, the absence of one of them is compensated by the activity of another member of the family. For example, MyoD largely rescues the delay in trunk muscle development observed in double Myf5: MRF4 knock-out mice [2-4], while some epaxial and all extraocular muscles (EOMs) are missing, indicating that MyoD cannot always compensate [5,6]. Myogenesis is completely abolished in triple Myf5: Mrf4: MyoD mutants because progenitors are never specified [5]. In contrast, MyoG, the fourth MRF, controls terminal differentiation, and in its absence, skeletal muscle is poorly developed and animals die at birth [6-8].

While it is clear that MRFs are fundamental for skeletal muscle development and thus, their expression and activity need to be tightly regulated, this review will focus on the multiple factors that regulate the expression and activity of myogenic genes in general. Emphasis will be placed on the main epigenetic processes: DNA methylation, histone modifications, chromatin remodeling, and regulation by non-coding RNAs. Post-transcriptional regulation processes will also be addressed.

2. DNA methylation

Mouse fibroblasts transdifferentiate into myocytes upon inhibition of DNA methylation [9,10]. Mounting evidence indicates that this epigenetic modification, the first to be discovered [11,12], plays a key role in myogenesis.

DNA methylation refers to the covalent addition of a methyl group into a cytosine base, leading to the formation of 5-methylcytosine (5mC). This occurs mainly in CpG dinucleotide pairs. In fact, it has been reported that 7% of cytosines within the sequence CCWG (W stands for A or T) are methylated in human skeletal muscle [13], although the role of non-CpG methylation in myogenesis remains largely unexplored. During demethylation, 5mCs are first converted into 5-hydroxymethyl cytosines (5hmCs), which undergo additional modifications to be eventually replaced by unmethylated cytosines [14]. Interestingly, 5hmCs may not be mere passive intermediates in the process, and it has been proposed that they would poise specific regions for full demethylation and rapid regulation of gene expression [15]. This might explain the unusual 5hmCs enrichment in Notch signaling genes [15], whose rapid induction upon muscle injury is required for proper...
Experimental Cell Research 419 (2022) 113299

2

C. Vicente-García et al.

repair [16]. In any case, skeletal muscle possesses considerable levels of 5hmCs with respect to other tissues -except brain [17]- and their functional role is just starting to be elucidated [18]. DNA methylation on adenine bases has recently been reported in mammals [19]. However, there is controversy as to whether this is a bona fide modification or a technical artifact [20]. In any case, this type of modification has not been reported in mammalian skeletal muscles, to the best of our knowledge.

Despite a common belief that DNA methylation is associated with gene repression, there is no clear link between methylation profile and expression outcome for a given gene, at least, in skeletal muscle [17,21,22]. The functional consequence depends greatly on the surrounding chromatin epigenetic signature [23], CpG density [24,25], TF availability [23], or the precise location of DNA methylation. Indeed, intragenic methylation is generally associated with alternative splicing regulation [17,26], alternative promoter usage [18,27] and activation of gene expression [28], while intergenic methylation on promoters and enhancers is involved in silencing [23,28,29]. In agreement, promoters and enhancers of muscle specific genes are frequently hypomethylated in muscle, presumably favoring transcription by allowing the binding of specific TFs [30].

Skeletal muscle displays specific DNA methylation signatures that in turn depend on fiber type [31] and gender [32], among other factors, and that are progressively remodeled throughout the individual life-span, the signature of mature adult skeletal muscle differing substantially from that of terminal myotubes in vitro [17,24], as well as of aged muscle [33,34]. In fact, muscle tissue methylome analyses can be used to predict chronological age [35]. Even satellite cells (SCs) possess specific profiles depending on their anatomical location [36] and age [37]. Modifying the niche (i. e, heterotopic transplantation of SCs) seems to affect methylation patterns at enhancers, but not promoters, while ageing leads to an increase in methylation heterogeneity in SCs, especially in CpG islands and regions associated with repressed genes, but also in other types of regulatory elements and promoters [37]. In fact, the increased heterogeneity in promoters could be responsible for the increased transcriptional variability observed in these cells, and ultimately, for their age-associated decline in function [37].

Importantly, the skeletal muscle methylome is plastic, able to respond to stimuli such as exercise [38], acute hypertrophy induction [21] or, as in the case of SCs, the niche [36]. In this regard, it was recently shown that upon SCs transplant from EOMs into the tibialis anterior (TA) limb muscle, their methylome is partly remodeled into that of TA SCs, underscoring the influence of the microenvironment [36]. Interestingly, this remodeling does not occur at promoters, but at enhancers, with differentially methylated elements between TA and EOM SCs partially explaining the location-specific transcriptional profiles of these cells [36]. Methylome plasticity is also manifested throughout development. Indeed, during myogenic commitment and early differentiation, cells undergo a global increase in methylation in order to silence genes involved in embryogenesis or differentiation of non-muscle lineages, while muscle specific genes generally lose methylation and become activated [17,22]. Later in differentiation, some of these muscle specific genes are targeted for de novo methylation and eventually silenced [24]. In agreement, the expression of DNA methyltransferases and demethylases is dynamic during the whole process [39,40], and their activity is essential for differentiation to succeed (e.g. Refs. [40–42]). Thus, during differentiation, both methylation and demethylation events take place in parallel in different genes, as well as sequentially on the same genes. Interestingly, the expression profiles of some methyltransferases and demethylases vary among different muscle groups in the adult mouse [43] (http://msc.ledb.org/), which could contribute to their differential methylome signatures observed, at least, in humans [31].

Detailed analysis of the methylation status of several myogenic gene promoters, including the MRFs, revealed that demethylation is necessary, but not sufficient, for gene expression [40,44]. Instead, demethylation would permit the acquisition of a transcriptional poised state and gene expression would only start after the deposition of activating and/or the removal of repressive histone marks [40], as in the case of Myog induction described below. In parallel, some muscle enhancers are also remodeled during differentiation [22,40]. This is not surprising considering these analyses are typically conducted in vitro, and may not reflect the full situation in vivo. For instance, over twenty enhancers regulate the expression of Mrf4 and Myf5 in a precise spatio-temporal manner in the embryo [45]. Probably only those required in an in vitro setting would show dynamism [22].

A recent model posits that Pax7, the master regulator of trunk and limb muscle specification together with Pax3 [46], would function as a pioneer factor binding to closed chromatin states and leading to a first wave of DNA demethylation of the MRFs [40]. In a second wave, the MRFs would directly or indirectly recruit DNA methylases and histone modifying enzymes to other myogenic loci to carry on with differentiation [40]. MyoD would be key in this process since its pioneer activity, albeit limited [47,48], is widely recognized [49–51], and indeed responsible for the transdifferentiation phenomenon of fibroblasts into myoblasts upon DNA methylation inhibition [52,53].

3. Histone modification and chromatin remodeling

Most of what we know about histone modification and chromatin remodeling during myogenesis comes from in vitro studies using the C2C12 mouse myoblast cell line as a paradigm. These epigenetic processes refer to the modifications exerted on chromatin in order to create an open transcriptionally permisive state where regulatory factors can bind. This is achieved by histone-modifying enzymes and ATP-dependent complexes that are able to reposition nucleosomes [54]. Exchange of histone variants within nucleosomes is remarkably important for the regulation of myogenesis [55,56], and especially to maintain muscle lineage identity [57].

Epigenetic control of the transitions between quiescence, proliferation and differentiation involves the regulation of cell cycle regulators and muscle-specific genes (reviewed in Ref. [58]). MyoD is key in these transitions, and its expression needs to be tightly repressed in quiescence [59], because MyoD protein accumulation marks the onset of SC activation and the start of myogenesis [58]. In this process, MyoD re-wires the three-dimensional (3D) chromosome architecture, re-configuring genome-wide interactions between regulatory elements and target genes [60,61], while also regulating gene expression by direct binding to promoters and enhancers [50]. Importantly, not all intergenic MyoD-bound regions function as enhancers [50,62]. In some cases, MyoD would prime regulatory elements for later activation or act to suppress enhancer activity [50,63]. In this line, it has been proposed that MyoD can function as a repressor or activator depending on interacting cofactors [58]. In any case, MyoD binding is generally associated with histone acetylation [62], despite global deacetylation during differentiation [64]. This is not surprising considering its ability to recruit histone acetylases (HATs) [65–69]. Of note, histone acetylation contributes to transcriptional activation by reducing nucleosome:DNA interactions, leading to chromatin opening and TF binding. It was recently shown that Myf5 also induces local histone acetylation; however, it lacks MyoD tandem commitment within nucleosomes is remarkably important for the regulation of myogenesis [55,56], and especially to maintain muscle lineage identity [57].

Alongside HATs, MyoD is able to recruit the SWI/SNF chromatin remodeler complex at condensed loci [71,72]. This ability is encoded in the amino terminal cysteine-histidine-rich and the carboxy-terminal regions, which differ from other MRFs, especially MyoG, and probably account for their different activities [49,72–75].

For the purpose of illustrating the combinatorial action of epigenetic processes in myogenesis (Fig. 1) we will focus on their role in the activation of Myog, updating the work by Faralli & Dilworth [76]. The model described here comes from in vitro data. Thus, differences may
exist relative to embryonic development [63].

In quiescence, the Myog locus is maintained in a repressed chromatin state by hyper-methylation of its promoter, and the binding of the transcriptional repressor CIBZ [77]. Further, polycomb repressive complex 2 (PRC2) mediates the deposition of heterochromatin-associated H3K27me3 marks [64, 78], while HDAC1 keeps histones deacetylated [79]. YY1 and Snai1/2 could be involved in the recruitment of these proteins as in other loci [80, 81]. Indeed, Snai1/2 interact with histone deacetylases at GC-rich enhancers preceding the access of MyoD until differentiation starts, when Snai1/2 are repressed in an MRF-dependent manner [81]. Whether the same mechanism takes place at Myog remains undetermined. Additionally, Mef2D along with the histone chaperone HIRA, recruit other histone-modifying enzymes [82]. In parallel, homeodomain proteins
Pbx1 and Meis1a are constitutively bound at the promoter [83], with Pbx1 being originally recruited by the histone variant macroH2A1 [84]. The Pbx1/Meis dual functions as a pioneer complex attracting and stabilizing MyoD binding to noncanonical E-boxes [83,85]. Calcineurin (Cn), also bound to the promoter, would participate in MyoD stabilization [86]. After p38α-mediated phosphorylation [87], MyoD recruits KMT1A/Suv39h1, which deposits H3K9me2/3 marks at the promoter assisting in gene repression [88]. G9α might contribute to the deposition of these marks [89], while it is clear that it methylates and inactivates MyoD [89] and Me2D [90]. MyoD further recruits the SWI/SNF remodeling complex subunit BAF60c, but the complex is not yet assembled [91].

MyoD expression during differentiation requires the removal of repressive marks and the subsequent deposition of activating ones. The dynamics of these processes are not completely understood but the following events have been described. First, the chromatin remodeler Chd2 and the SRCAP complex mediate the incorporation of H3.3 and dynamics of these processes are not completely understood but the Chd2 and the SRCAP complex mediate the incorporation of H3.3 and MyoD further recruits the SWI/SNF remodeling complex subunit BAF60c, but the complex is not yet assembled [91].

MyoD expression during differentiation requires the removal of repressive marks and the subsequent deposition of activating ones. The dynamics of these processes are not completely understood but the following events have been described. First, the chromatin remodeler Chd2 and the SRCAP complex mediate the incorporation of H3.3 and dynamics of these processes are not completely understood but the MyoD further recruits the SWI/SNF remodeling complex subunit BAF60c, but the complex is not yet assembled [91].

4. Non-coding RNAs

The last decade was prolific in the discovery of long non-coding RNAs (lncRNAs) involved in skeletal muscle biology [Fig. 2] [122–126], many findings linked to high-throughput technologies [127–133]. lncRNAs are non-coding transcripts over 200 nucleotides, their function to date being correlated with their cellular distribution [123]. Nuclear lncRNAs can directly guide the epigenetic machinery to specific loci, like Dum, which recruits DNA methyltransferases to the nearby Dppa2 promoter eventually favoring myogenic differentiation [134]. Examples of lncRNAs interacting with chromatin silencing machinery abound, with SYSL and the ubiquitous Neat1 [136] recruiting PRC2 to myogenic loci to sustain proliferation, Pprr1b-lncRNA evicting PRC2 from the Myod1 locus to promote differentiation [137], or Lnc-Rewind recruiting G9α to the Wnt7b locus to allow myoblast expansion [138]. They can interact also with TFs and other proteins to regulate their activity, such as Linc-YY1, transcribed in the opposite direction to that of YY1. The interaction between Linc-YY1 and YY1 removes YY1/PRC2 repressive complexes from some myogenic loci, causing their activation. In parallel, Linc-YY1 can guide YY1 to STAT3-bound loci, silencing negative regulators of differentiation. Thus, in its absence, in vitro differentiation is delayed, and in vivo regeneration, suppressed [132]. Other examples include Myolinc, whose interaction with TDP-43 induces Myod1 expression at the onset of differentiation [139], or Myoparr, expressed from the Myog promoter, which acts as a scaffold for protein binding during Myog activation [140]. Further, Linc-RAM and SRA bind directly to MyoD, facilitating its interaction with the SWI/SNF remodeling complex [141], or with the p68/p72 helicases to activate transcription [142]. The case of lnc-MYH is particularly interesting. Located at the myosin Myh3-Myh13 cluster, it is expressed alongside these sarcomeric proteins. It binds and suppresses the activity of the INO80 chromatin remodeler complex, a complex that is fundamental for myoblast proliferation. Thus, coexpression of sarcomeric differentiation genes and the antiproliferative lnc-MYH ensures that both processes are synchronized [143]. Finally, it was reported that lncRNAs could regulate genome architecture: pCharme, an isoform of the muscle specific Charme lncRNA, interacts with MATRX3 and PTBP1 to bring myogenic loci into close nuclear proximity, forming 3D domains where transcription is facilitated [127,144,145].

Many cytoplasmic lncRNAs block or enhance mRNA translation. For example, Inc-SMaRT binds to the G4 structure of Mtcy mRNA obstructing translation, and causing the downregulation of MLX transcriptional targets [146]. This inhibition occurs early during myogenesis presumably to control the onset of differentiation. In contrast, Inc-31 binds to and enhances the translation of ROCK1 mRNA, sustaining myoblast proliferation [127,147]. Cytoplasmic lncRNAs can regulate also target stability. Sirt1AS lncRNA [148] and OIP5-AS1 [149] protect Quiescent SCs epigenome reports need to be considered carefully as isolation methods can activate SCs [120]. Thus, these experiments often reflect the state of early-activated SCs rather than quiescent cells. It was initially thought that the level of the H3K27me3 marks in quiescence was low [121]. However, truly quiescent SCs would be rich in heterochromatin, which would be lost upon activation [59]. Early-activated SCs are also enriched in activating H3K4me3 marks, making their chromatin largely permissive to transcription [121]. Interestingly, promoters of genes associated with developmental processes different to myogenesis contain both marks, underscoring the potential of SCs to adopt non-myogenic fates [121]. Upon differentiation, H3K4me3 marks are maintained but the levels of H3K27me3 increase, even in genes marked with H3K4me3, which then become bivalent and repressed [121]. These findings mostly mirror what happens during differentiation of C2C12 cells, where global histone deacetylation is additionally observed [64]. Of note, SCs accumulate H3K27me3 with age, presumably a consequence rather than the cause of age-related permanent loss of transcriptional potential, and thus functionality, in these cells [121].
miRNAs controlling the MRFs and upstream regulators in the mouse. Table 1 and Fig. 2.

**Table 1**

<table>
<thead>
<tr>
<th>Target</th>
<th>miRNAs</th>
<th>Biological role</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pax3</td>
<td>miR-1, miR-27b, miR-206</td>
<td>Balance between proliferation and differentiation of SCs</td>
<td>[274,275]</td>
</tr>
<tr>
<td>Pax7</td>
<td>miR-1, miR-206, miR-431, miR-486</td>
<td>Balance between proliferation and differentiation of SCs</td>
<td>[276–278]</td>
</tr>
<tr>
<td>MyoD</td>
<td>miR-221, miR-222, miR-669, miR-669q</td>
<td>Inhibition of differentiation</td>
<td>[279]</td>
</tr>
<tr>
<td>Myf5</td>
<td>miR-20, miR-31</td>
<td>Prevention of skeletal muscle differentiation in the heart</td>
<td>[280]</td>
</tr>
<tr>
<td></td>
<td>miR-31, miR-106b</td>
<td>Inhibition of Myf5 transcript translation in the central nervous system</td>
<td>[281]</td>
</tr>
<tr>
<td></td>
<td>miR-374b</td>
<td>Maintenance of quiescence by preventing Myf5 translation in SCs</td>
<td>[282]</td>
</tr>
<tr>
<td></td>
<td>miR-186</td>
<td>Inhibition of differentiation</td>
<td>[283]</td>
</tr>
</tbody>
</table>

Sirt and Mef2c mRNAs, respectively, while m1/2-absRNAs [150] mediate target decay. Proteins can also be stabilized by IncRNAs, such as SAM, which prevents the ubiquitination and degradation of the Sustg1 protein, required for kinetochore assembly during myoblast proliferation [151]. Finally, cytoplasmic IncRNAs can neutralize the action of microRNAs (miRNAs). Dozens of miRNAs have a profound impact in muscle development and disease [152], either targeting myogenesis proper (e.g. the MRFs or upstream regulators, Table 1), or more general processes like cell division. Even muscle-specific miRNAs, the “myomiRs” have been described [153]. IncRNAs can act as competing endogenous RNAs (ceRNAs) for miRNAs, blocking their activity. Indeed, lnc-MD1 acts as a sponge for mir-133 and mir-135, and alleviates the repression of their targets, including the myogenic TFs MAML1 and MEF2C [153], as well as the RNA-binding protein HuR during the early stages of differentiation [154]. IncRNA AK017368 [155], lncMUMA [156], linc-smad7 [157], linc-yy1 [158,159], Inc-MGPF [160], MARI [161], lnc-231 [162] and Inc-ORA [163] belong to this category.

Both widely-expressed and muscle-specific IncRNAs have been found to play a role in myogenesis, muscle regeneration and disease [123]. For instance, lrm is highly expressed in the central nervous system, neuro-endocrine organs and muscle in the developing mouse embryo [164]. lrm promotes the assembly of the MyoD/Mef2D complex, inducing the expression of myogenic genes. This would explain why its absence inhibits differentiation in vitro and muscle regeneration in vivo [165]. Its role in non-muscle tissues is unknown.

The majority of IncRNAs involved in myogenesis are dynamically expressed [127,166], such as Malat1, highly expressed during myoblast proliferation, recruiting histone repressive complexes to MyoD targets. During differentiation, Malat1 is targeted for degradation by the pro-myogenic mir-181a. As a result, histone repressive complexes are destabilized, allowing MyoD to activate its targets [167]. Malat1 is also a sponge for mir-133, and by competing for this shared miRNA, it modulates the expression of SRF [168], also key in myogenesis [169].

Several IncRNAs are linked to MyoD (Fig. 2), either because it regulates their expression [132–134,141,160,170], or vice versa [129,137,139,156], they regulate its activity [141,142,165,167], or are transcribed from the Myod1 locus [129,133]. Indeed, MUNC (DRR) and CE RNA are transcribed from the distal and core enhancers of Myod1, respectively. Both would directly or indirectly open up chromatin, generating a permissive state for gene expression [171]. While CE RNA regulates the expression of Myod1 in cis [129], it is not clear whether MUNC also contributes [171–173], although it is involved in the regulation of other myogenic genes in trans, such as Myog, dependently or independently of Myod [129,171–173]. In contrast, IncMyoD, expressed 30 kb upstream Myod1, does not regulate its expression but sequesters IMP2, preventing the translation of proliferation genes [133].

At least two IncRNAs related with myogenesis are located at imprinted loci. Gt2/Meg3 acts as a scaffold molecule that guides PRC2 to the nearby Dlk1 gene, ensuring its silencing for appropriate skeletal muscle development [174,175]. More recently, Gt2/Meg3 has been proven to be essential for the proliferation and differentiation of myoblasts through the regulation of TGFβ signaling and the control of epithelial-mesenchymal transition during myogenesis [176]. Besides encoding mir-675-3p and mir-675-5p, known to promote muscle...
differenciation and regeneration [177], H19 plays various roles as a lncRNA in myogenesis: it sponges let-7 miRNAs [178], while also recruiting the KH-type splicing regulatory protein KSRP to lncRNA in myogenesis: it sponges DMD: Duchenne muscular dystrophy; MHC: myosin heavy chain.

egulin MLN, the most abundant micropeptide in adult skeletal muscle. It scripts, decreasing their stability [179]. In either case, differentiation and regeneration [177], and micropeptides, are well characterized. precocious differentiation. C. Vicente-García et al.

- different isoforms of Ca[2+]–ATPase SERCA inhibiting Ca2+ uptake into the SR, diminishing muscle contractility [180], while a specific isoform of SRA enzymes for SRAP, which prevents SRA-mediated activation of MyoD by sequestering SRA. This implies that differential splicing of SRA transcripts regulate the balance between coding and non-coding forms, with a direct impact on myogenesis [181]. More often, only micropeptides seem to be functional, as in the case of DWORF [182], SPAR [183], LEMP/MOXI/Mitoregulin/MPM [184-187], or Myomixer/Myomerger/Mionix [188-190].

CircRNAs constitute a peculiar class of RNA [191]: they are highly stable, lowly expressed and cell-type specific circular molecules mainly formed by a non-canonical splicing process called backsplicing [192]. Thousands of circRNAs have been identified in mammalian skeletal muscle (e.g. Ref. [193]), specially in livestock and poultry, while only a few circRNAs have been fully characterized in mouse and human, where they mainly function as sponges for miRNAs (Table 2). circRNAs can regulate the activity of certain proteins as well, and some of them can be translated [191].

<table>
<thead>
<tr>
<th>circRNA</th>
<th>Biological role</th>
<th>Mechanism</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>circARID1A</td>
<td>Inhibits proliferation, promotes differentiation and regeneration</td>
<td>Sponge for miR-6368</td>
<td>[285]</td>
</tr>
<tr>
<td>circBNC2</td>
<td>Promotes differentiation; downregulated in DMD</td>
<td>Probably competes with the production of the anti-myoync linear BNC2 mRNA form</td>
<td>[286]</td>
</tr>
<tr>
<td>circFgr2</td>
<td>Inhibits proliferation, promotes differentiation and regeneration</td>
<td>Sponge for miR-133, inducing the expression of Map3k20 thus activating the JNK/ MAPK pathway</td>
<td>[287]</td>
</tr>
<tr>
<td>circFUT10</td>
<td>Inhibits myoblast proliferation and differentiation</td>
<td>Sponge for miR-365a-3p, enhancing the expression of HOXA9</td>
<td>[288]</td>
</tr>
<tr>
<td>circHIPK3</td>
<td>Promotes differentiation</td>
<td>Sponge for miR-7, releasing TcR12 repression; also sponges miR-124-5p and miR-379-5p</td>
<td>[289,290]</td>
</tr>
<tr>
<td>circNfx</td>
<td>Promotes differentiation</td>
<td>Sponge for miR-204, alleviating the repression of its target MEF2C. Not fully experimentally validated</td>
<td>[291]</td>
</tr>
<tr>
<td>circQ6l</td>
<td>Promotes differentiation; downregulated in DMD</td>
<td>Unknown</td>
<td>[286]</td>
</tr>
<tr>
<td>circSamd4</td>
<td>Promotes differentiation</td>
<td>Sequester PUR proteins contributing to the derepression of MHC transcription</td>
<td>[292]</td>
</tr>
<tr>
<td>circZNF609</td>
<td>Promotes myoblast proliferation; involved in human cancers, including RMS; upregulated in DMD</td>
<td>Sponge for various miRNAs. Also translated.</td>
<td>[286,293]</td>
</tr>
</tbody>
</table>

5. Post-transcriptional myogenesis control by RNA-binding proteins (RBPs)

RBPs have been classically defined as proteins that bind RNA through a specific set of domains or combinations thereof, in order to control RNA metabolism at multiple levels. Recently, both concepts have been challenged [194,195]. Several reports have revealed a plethora of proteins lacking a clear RNA-binding domain, but that are indeed associated with RNA molecules, including metabolic and mitochondrial proteins [196,197]. These findings raise the possibility that, in some cases, RNA would be regulating RBPs rather than the opposite [194,195]. As mentioned above, lncRNAs clearly illustrate this concept, given that they can guide the activity of multiple proteins, including chromatin remodelers and more classical RBPs.

Many RBPs play key roles in myogenesis and are associated with muscular [198-200] or neuromuscular [200,201] diseases, even if most of them are ubiquitously expressed [198]. Exceptions include Rbm24, Rp3l3 and some isoforms of Fxr1 [202], which are specific for skeletal muscle and, to a lower extent, for the heart as well [203]. Rbm24, partially under the control of MyoD [204], is the best characterized muscle-specific RBP. In vitro, it promotes myoblast differentiation through stabilization of Myog transcripts [205], and is required also for somitic differentiation in vivo by a yet unknown mechanism [204], and for myoblast fusion and differentiation during regeneration in vivo, presumably by regulating the alternative splicing of transcripts involved in this process, such as Mef2d, Naca, Rock2 and Lrf1p1 [206]. In line with the fact that not only RBPs with classical features are able to bind RNA, Myf5 has recently been recognized to interact with hundreds of mRNAs [207]. Indeed, Myf5 binds to cyclin D1 transcript enhancing its translation, which would be one of the multiple ways in which Myf5 promotes myogenesis [207].

Of note, the expression of many RBPs is dynamic during myogenesis, underscoring their differential requirement in the various steps of the process [196]. For example, lq2b2p2 expression is low in quiescent SCs, increases strongly upon activation, and eventually disappears during terminal differentiation [208]. This pattern of expression is in agreement with the role of lq2b2p2 in promoting proliferation by enhancing the translation of c-myc, Sp1 and Ifg1r transcripts [208]. In contrast, Fxr1 activity increases during differentiation [196], when it is absolutely essential [209].

Thanks to their ability to translocate between nucleus and cytoplasm, RBPs mediate in all steps of RNA biology: alternative splicing, transcript stability/degradation, translational regulation, nuclear export, localization, and editing [195,198,203]. Since skeletal muscle is among the tissues with the highest levels of alternative splicing (AS) [210], the most heavily studied RBPs in the field are those involved in this process [198].

5.1. Alternative splicing

RBPs involved in AS include members of the CELF (CUGBP1 or CELF1, CELF2), Muscleblind-like (MBNL) and Ribfox families, Quaking (Qk), PTB (also called PTBP1 or hnRNPI) and SRSF10. Each of these families recognizes specific transcript motifs that favor exon skipping or inclusion [199,211]. Importantly, many splicing events are regulated by the combined action of several RBPs, which can act cooperatively (e.g. Ref. [212]) or antagonistically (e.g. Ref. [213]).

AS expands the repertoire of isoforms that can be translated from a single transcript [211], and multiple AS events have been described in myogenic TFs [214]. One of the most remarkable cases is Me2d2, which can function as a transcriptional repressor or as an activator depending on the exon composition [215]. AS also regulates transcript stability through nonsense-mediated decay (NMD), which ensures that transcripts containing premature STOP codons within the coding sequence are detected and degraded. Curiously, Rbm4 regulates AS-coupled NMD of PTB [213]. Considering that both RBPs antagonistically regulate
mature AS events by competing for the same RNA-binding motif, the fact that one modulates the expression of the other has a profound physiological impact during myoblast differentiation [213].

5.2. Alternative polyadenylation (APA)

Following transcription, the newly produced mRNA molecules are capped and polyadenylated, processes required for nuclear export, translation and protection from degradation by exonucleases. Transcripts may contain several polyadenylation sites, enabling the production of transcripts of various lengths by APA.

Considering that 3′-UTRs often harbor binding sites for RBPs, miRNAs and other mRNAs, shorter transcripts escape from regulation [216]. For example, APA contributes to control Pax3 expression in quiescence: in the limbs, U1 snRNA interacts with Pax3 transcripts and serves as scaffold for snRNPs U1, which induces the expression of a long isoform susceptible to inhibition by miR-206 [217]. In other tissues with lower expression of U1 snRNA, such as the diaphragm, the composition of Pax3 transcripts is shifted towards the shorter isoform [217]. Other RBPs involved in APA include PABPN1 [218, 219] and MBLN [220].

5.3. mRNA modifications

Among the many chemical modifications identified in RNA molecules, methylation of adenosine bases at the N6 position (m6A) is the most abundant in the organism, being crucial during myogenesis [221]. Different RBPs are in charge of ‘writing’ these modifications at specific RNA motifs, “reading”, and “eraser” them [222, 223]. Multiple lines of evidence highlight the important role of the METTL3 methyltransferase as the “writer” in skeletal muscle. For example, in the adult, METTL3 is required for muscle size maintenance and hypertrophic responses [224]. During myogenesis, this protein is involved in the biogenesis of muscle-specific miRNAs [225], is essential for SC biology [226], and stabilizes MyoD1 transcripts during myoblast proliferation [227]. Not surprisingly, its knockdown impairs later differentiation [228]. In contrast to expectations, depletion of the FTO demethylase “eraser” also inhibits differentiation [228]. The m6A methylome is dynamic during myogenesis [229]. This presumably constitutes a way to regulate the process through the action of the different RBP “readers”, which can lead to transcript degradation, stabilization and translation, splicing or transport [222, 223]. All in all, the impact of m6A modifications in myogenesis is clear and deserves further investigation.

5.4. Other processes

Some RBPs regulate mRNA stability or translation efficiency by specifically targeting muscle-restricted transcripts, or by regulating more general RNAs also involved in muscle biology. That is the case of AUFI, also known as hnrNPD, which targets many cell cycle regulators, early response genes, inflammatory mediators, cytokines, G protein-coupled receptors and oncoproteins [230]. In muscle, AUFI promotes the degradation of several mRNAs through binding to AU-rich elements, or AREs, in their 3′-UTRs. AUFI targets include the cell cycle regulator cyclin D1, the MyoD and Myf2 inhibitor Twist1, and the SH3 pathway inhibitor Rgs5 [231]. Further, AUFI binds to the 3′-UTR of Myf2c, increasing translation without affecting stability [232]. All these events result in myoblast cell cycle exit, promoting differentiation and maturation [231, 232]. Similarly to AUFI, Lin-28 promotes differentiation by increasing IGF-2 growth factor translation [233]. This is a particularly intriguing case considering that Lin-28 is a known marker for “stemness” and thus, it is absent from most adult tissues, except cardiac and skeletal muscles [224]. In parallel, the ubiquitous splicing regulator CUGBP1 controls myoblast proliferation and differentiation by increasing the translation of Myf2a, p21 and cyclin D1 transcripts [235, 236]. CUGBP1 is also a potent mRNA destabilizing factor, causing the decay of transcripts containing GU-rich elements (GREs) in their 3′-UTRs, including MyoD1 [237].

As in the case of AS, RBPs function cooperatively, antagonistically or sequentially, making the specific molecular activity and eventual physiological impact of each RBP context-dependent. For example, during proliferation, Myog and p21 transcripts are targeted for degradation by KSRP [238]. Upon differentiation, KSRP is released and these and other transcripts, such as Myod1 and HMGB1, are stabilized by the ubiquitously expressed HuR RBP [239–241]. The sequential activity of KSRP and HuR ensure efficient differentiation. In parallel, both KSRP and HuR cooperate to degrade nucleophosmin during differentiation [242], while Pitx2-HuR complexes promote myoblast proliferation through cyclin D1 stabilization [243].

RBPs add a new layer in the control of the life cycle of SCs. The absence of AUFI impairs self-renewal due to decay failure of specific transcripts, such as matrix metalloprotease Mmp9, involved in the degradation of the niche extracellular matrix components [244]. MyoD marks the onset of SC activation, being transcribed in quiescence, while translation is strictly controlled to prevent premature activation. Indeed, Staufen1 inhibits Myod1 translation [245], while Zip36 (a.k.a. Tristetraprolin or TTP), leads to ARE-mediated transcript degradation [246]. Both RBPs are essential for primed quiescent state maintenance. Furthermore, Zip36 (Brf1) and Zip36 (Brf2) act redundantly regulating SC fate and function, although their specific targets remain to be established [247]. In parallel, Myf5 is also expressed during quiescence, but transcripts are sequestered in ribonucleoprotein (RNP) granules together with fragile X mental retardation protein FMRP, among other RBPs, and miR-31 to prevent premature translation. Dephosphorylation of FMRP leads to RNP dissociation and SC activation [248]. The strategy of sequestration into RNP granules has been observed also for sarcomeric transcripts to presumably regulate sarcomere assembly during development and regeneration [249].

6. Post-translational modifications

Post-translational modifications (PTMs) mainly control protein interactions with other molecules, protein stability or their subcellular localization. Among the different types of PTMs involved in myogenesis, phosphorylation stands out as of critical importance [250]. For example, the phosphorylation increase in the heterogeneous nuclear ribonucleoprotein K (hnRNPK), leads to cytoplasmic retention, impacting upon myoblast proliferation and differentiation [251], while the phosphorylation status and cleavage of HuR also influence its cellular localization [252]. In parallel, phosphorylation of KSRP [238], Zip36 [246] and CUGBP1 [236] regulate their RNA-binding capacity. Beyond RBPs, many other proteins associated with myogenesis are targeted for phosphorylation, such as the alpha subunit of the eukaryotic initiation factor 2 (eIF2α). Its phosphorylated form, p-eIF2α, plays a central role in SC quiescence, self-renewal and expansion, while its dephosphorylation marks the transition to SC activation and differentiation through differential translation of selected mRNAs by each eIF2α molecular form [253, 254].

Multiple signaling pathways converge on the MRFs (Fig. 3), especially through phosphorylation, which, in general, inhibits myogenesis. Indeed, protein kinase C and/or A (PKC/A) phosphorylate Myog [255, 256], MyoD and Mhr4 [257], negatively affecting their DNA binding capacity in vitro. However, only in the case of MyoD on Thr115 [258] and MyoG on Thr87 [256], this negative effect is reproduced in vivo. Similarly, p38 MAP-kinase phosphorylation represses Mhr4 activity, presumably to silence myogenic genes at the terminal stages of differentiation [259]. This kinase targets multiple sites on MyoD as well. It phosphorylates Ser199 and Ser200, inducing the assembly of a repressive transcriptional complex through the recruitment of histone modifying enzymes (see above) [87]. However, the role of Ser5 phosphorylation is unclear [260]. Cyclin-dependent kinases cdk1 and cdk2 can also target Ser5 and Ser200 in vitro, but only Ser200 in vivo, compromising MyoD activity and stability [261]; cdk1 is able to
phosphorylate MyoG, although the functional consequences remain unknown [262]. In contrast, other kinases have a positive impact on muscle development, including cdk9 [263] and the MAPK signaling activator MOS [264], both of which target MyoD. Of note, casein kinase II (CKII) overexpression promotes differentiation by increasing the transcriptional activity of Mrf4, MyoD [265] and Myf5 [266], which are substrates for this kinase in vitro. However, only in the case of Myf5, direct phosphorylation by CKII is determinant for the positive effects CKII exerts on myogenesis [266].

Other PTMs beyond phosphorylation have been identified on the MRFs. As exemplified at the Myog locus (see above), G9a-mediated methylation of MyoD targets the protein for degradation, and blocks the activation by pCAF, preventing premature differentiation [89,267]. When appropriate stimuli arrive, pCAF acetylates MyoD, increasing its DNA binding ability, triggering differentiation [100]. MyoD degradation is additionally elicited by atrogin-1/MAFbx ubiquitination on Lys133 [268], although phosphorylation by the MAPK kinase MEK1 on Tyr156 nearby confers protection [269]. The rare N-terminal ubiquitination mechanism can also take place [270], possibly mediated by HUWE1, at least in vitro [271].

Finally, diverse stimuli induce PTMs on the MRFs. For example, c-Abl phosphorylates MyoD in response to DNA damage to suppress differentiation [272], while muscle activity exerts the opposite effect by inducing calcium/calmodulin-dependent protein kinase II (CaMKII)-mediated phosphorylation of MyoG [273].

7. Concluding remarks

Myogenesis is a highly complex process that needs to be tightly regulated during development and regeneration, to ensure that the newly formed skeletal muscle is fully functional and integrated with the rest of the organs. Chromatin status, 3D organization of the genome, availability and PTMs on TFs and cofactors, alternative splicing and regulation of the stability of transcripts, regulation of translation and transcript localization (Fig. 4) are important factors to consider when studying how myogenesis proceeds. While often only one piece of the puzzle is analyzed at a time, far reaching conclusions should be drawn considering the whole picture. Further, skeletal muscle is highly heterogeneous, with different muscle groups exhibiting different contractile, metabolic and regenerative properties. During development, the formation of these muscle groups depends on the origin of the progenitors, with distinctive inductive signals triggering the process in different regions of the embryo, adding yet another layer of complexity in the study of myogenesis regulation [1]. Thus, because of its intricate nature, most of our knowledge of myogenesis comes from in vitro models using established muscle-derived cell lines, such as C2C12 cells, or primary cultures, which constitute a fairly simple approach. Cell culture experiments serve to address questions that would be very difficult–even impossible–to tackle in a more physiological setting, but the information obtained may not reflect reality due to, paradoxically, their extreme simplicity. For example, in vitro studies usually disregard biochemical or mechanical signals myoblasts receive by surrounding tissues. Therefore, it will be very important to turn our efforts towards the development of tools that enable the study of myogenesis and skeletal muscle biology using a more integral perspective. In this sense, while still under an in vitro setting, the use of the so-called somitoids, somite-like organoid structures derived from induced pluripotent stem cells (iPSCs), constitutes a very promising tool for the study of the earliest stages of muscle development [294]. Furthermore, if these
organoids are derived from human samples suffering from any muscle-related disease, they could be used to shed more light into the pathophysiology of that disease from a personalized perspective, taking into account factors often neglected such as the influence of gender or age. Remarkably, in recent years the application of single-cell technologies have opened new and exciting avenues in the skeletal muscle field, unravelling new resident cellular and myonuclear populations, and showing how they respond to diverse stimuli [295]; their combination with organoids, including the adaptation of OMICs approaches still not available for single-cell research, should open new and exciting avenues. Further developments, including microchamber and 3D cultures, specially 3D-bioprinting approaches, need to be further developed for the study of more complex interactions driving myogenesis.

**Author statement**

Cristina Vicente-García, Juan Diego Hernández-Camacho and Jaime J. Carvajal wrote the review and generated the figures and tables.
showed in the review.

Declaration of competing interest

None.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was supported by the Association Française contre les Myopathies (AFM) through fellowship grant #22450 to C.V.-G., by the Spanish Ministry of Science and Innovation to J.J.C.

References


[36] K. Tsumagari, C. Baribault, J. Terragni, K.E. Varley, J. Gertz, S. Pradhan, Myopathies (AFM) through fellowship grant #22450 to C.V.-G., by the Spanish Ministry of Education, Culture and Sports through fellowship #22447, and PID2020-117058GB-I00 from the Spanish Ministry of Science and Innovation to J.J.C.


A role for histone deacetylase HDAC1 in modulating the transcriptional pathway targets SWI-SNF chromatin-remodeling complex to muscle-specific loci, specific gene expression by binding target promoters and facilitating the execution of the muscle program, Mol. Cell. 4 (1999) 725.

J. Wang, L. Kedes, Acetylation of MyoD directed by PCAF is necessary for the incorporation of MyoD-BAF60c into Brg1-based SWI/SNF chromatin-remodelling complex, EMBO J. 31 (2012) 2994.


S. Sweta, T. Dudnakova, S. Sudheer, A.H. Baker, R. Bhushan, Importance of long non-coding RNAs (lncRNAs) in myogenesis: a miR-31 overlapping lncRNA


C. Vicente-García et al.
Experimental Cell Research 419 (2022) 113299


