doi:10.1093/jmcb/mjy025
Published online April 28, 2018

Article

A novel long non-coding RNA *Myolinc* regulates myogenesis through TDP-43 and Filip1

Giuseppe Militello[1,2,3,4], Mohammed Rabiul Hosen[1,2,3], Yuliya Ponomareva[1,2,3], Pascal Gellert[5], Tyler Weirick[1,2,3,4], David John[1,2,3], Sajedah Mahmoud Hindi[6], Kamel Mamchaoui[7], Vincent Mouly[7], Claudia Döring[8], Lidan Zhang[9], Miki Nakamura[9], Ashok Kumar[6], So-ichiro Fukada[9], Stefanie Dimmeler[1,2], and Shizuka Uchida[1,2,4,*]

[1] Institute of Cardiovascular Regeneration, Centre for Molecular Medicine, Goethe University Frankfurt, Frankfurt am Main 60590, Germany
[2] German Center for Cardiovascular Research, Partner side Rhein-Main, Frankfurt am Main 60590, Germany
[3] Department of Biosciences, Goethe University Frankfurt, Frankfurt am Main 60438, Germany
[4] Cardiovascular Innovation Institute, University of Louisville, Louisville, KY 40202, USA
[5] Breast Cancer Now Toby Robins Research Centre, Institute of Cancer Research, London SW3 6JB, UK
[6] Department of Anatomical Sciences and Neurobiology, University of Louisville School of Medicine, Louisville, KY 40202, USA
[7] Sorbonne Universités, UPMC Univ Paris 06, INSERM UMRS974, CNRS FRE3617, Center for Research in Myology, Paris 75013, France
[8] Dr. Senckenberg Institute of Pathology, Goethe University Frankfurt, Frankfurt am Main 60590, Germany
[9] Laboratory of Molecular and Cellular Physiology, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka 565-0871, Japan
* Correspondence to: Shizuka Uchida, E-mail: heart.lncrna@gmail.com

Edited by Zefeng Wang

Myogenesis is a complex process required for skeletal muscle formation during embryonic development and for regeneration and growth of myofibers in adults. Accumulating evidence suggests that long non-coding RNAs (lncRNAs) play key roles in regulating cell fate decision and function in various tissues. However, the role of lncRNAs in the regulation of myogenesis remains poorly understood. In this study, we identified a novel muscle-enriched lncRNA called '*Myolinc* (*AK142388*)', which we functionally characterized in the C2C12 myoblast cell line. *Myolinc* is predominately localized in the nucleus, and its levels increase upon induction of the differentiation. Knockdown of *Myolinc* impairs the expression of myogenic regulatory factors and formation of multi-nucleated myotubes in cultured myoblasts. *Myolinc* also regulates the expression of *Filip1* in a *cis*-manner. Similar to *Myolinc*, knockdown of *Filip1* inhibits myogenic differentiation. Furthermore, *Myolinc* binds to TAR DNA-binding protein 43 (TDP-43), a DNA/RNA-binding protein that regulates the expression of muscle genes (e.g. *Acta1* and *MyoD*). Knockdown of TDP-43 inhibits myogenic differentiation. We also show that *Myolinc*—TDP-43 interaction is essential for the binding of TDP-43 to the promoter regions of muscle marker genes. Finally, we show that silencing of *Myolinc* inhibits skeletal muscle regeneration in adult mice. Altogether, our study identifies a novel lncRNA that controls key regulatory networks of myogenesis.

Keywords: long non-coding RNA, skeletal muscle, myoblasts differentiation, transcriptional regulation

Introduction

Skeletal muscle is the most abundant tissue of the mammalian body that plays diverse roles, such as posture maintenance, respiration, blood circulation, locomotion, and whole-body metabolism (Biressi et al., 2013). While skeletal muscle is a terminally differentiated tissue, it retains regeneration capacity due to the presence of adult muscle stem cells called 'satellite cells'. In naïve conditions, satellite cells reside between the sarcolemma and the basal lamina in a state of relative metabolic inactivity. However, muscle injury and other perturbations (e.g. exercise) lead to the activation of these cells that is followed by several rounds of cell division, and eventually their differentiation into myoblasts/myocytes, which then either fuse with each other to form a new myofiber or fuse with injured myofibers to complete the repair process (Relaix and Zammit, 2012; Yin et al., 2013). Myogenesis, the process of skeletal muscle formation, is a highly coordinated process that involves the sequential expression of a number of transcription factors including *Paired box 3* (*Pax3*) and *Pax7* followed by the expression of myogenic regulatory factors (MRFs), such as *Myf5*, *MyoD*, *Myogenin*, and *MRF4* (Braun and

Received November 17, 2017. Revised February 26, 2018. Accepted March 29, 2018.

Gautel, 2011). In addition, a number of growth factors, non-coding RNAs, signaling pathways, and epigenetic regulators play critical roles in skeletal muscle formation both *in vivo* and *in vitro* (Bentzinger et al., 2012; Simionescu-Bankston and Kumar, 2016).

Most of the mammalian genome is transcribed in the form of RNA, but only a few percent of them encode for proteins (Lander et al., 2001; Uchida et al., 2012; Uchida and Dimmeler, 2015). Any non-protein-coding RNAs longer than 200 nucleotides (nt) are currently classified as 'long non-coding RNAs (lncRNAs)' whose functions remain largely unknown (Flynn and Chang, 2014; Uchida and Bolli, 2017). Given the diverse functions of protein-coding genes and their protein products, it is speculated that lncRNAs also possess various functions. Among them, the binding to other macromolecules (i.e. nucleic acids and proteins) is of great interest as they could be used as a molecular switch to activate or to inhibit biological processes. Up until now, there have been a number of screening studies being published regarding the detection and the characterization of lncRNAs in the skeletal muscle, including: *AK017368* (Liang et al., 2017), *Chronos* (Neppl et al., 2017), *lnc133b* (Jin et al., 2017), *Dum* (Wang et al., 2015), *linc-MD1* (Cesana et al., 2011), *Lnc-mg* (Zhu et al., 2017), *Linc-RAM* (Yu et al., 2017), *Linc-YY1* (Zhou et al., 2015), and *LncMyoD* (Gong et al., 2015). The levels of some lncRNAs are precisely regulated at specific stages of myogenesis. Some lncRNAs function as sponges for microRNAs, whereas others can regulate the activity of various proteins involved in epigenetic regulation and gene expression (Simionescu-Bankston and Kumar, 2016). For example, sponging activity of *linc-MD1* for *miR-133* is enhanced by the RNA-binding protein human antigen R (HuR), which is an important mechanism for maintaining myogenic cells in the early differentiation stage (Cesana et al., 2011; Legnini et al., 2014). Other lncRNAs, such as *MUNC* and *LncMyoD*, regulate myogenic differentiation through other mechanisms. *MUNC* positively regulates myogenesis through directly augmenting gene expression of *MyoD*, *Myogenin*, and *Myh3* from a heterologous promoter (Mueller et al., 2015) Similarly, *LncMyoD*, identified upstream of the *MyoD* gene, also promotes myogenesis potentially through facilitating withdrawal of myoblasts from the cell cycle through inhibiting translation of proliferation genes *N-RAS* and *c-Myc* (Gong et al., 2015). However, while the role of a few lncRNAs has been uncovered, it is possible that several other lncRNAs that are yet to be identified regulate myogenesis through distinct mechanisms (Butchart et al., 2016; Simionescu-Bankston and Kumar, 2016; Peng et al., 2017; Zhou et al., 2017).

In this study, we screened for long intergenic non-coding RNAs (lincRNAs) that are highly and specifically expressed in the skeletal muscle of mice and in cultured myoblasts after initiation of the differentiation program. Through this screening, we identify a lincRNA, which we named '*Myolinc*'. *Myolinc* expression is greatly increased during the myogenic differentiation. *Myolinc* also controls the expression of protein-coding gene *Filip1* in a *cis*-manner. Silencing of *Myolinc* or *Filip1* inhibits the differentiation of myoblasts into myotubes. Our results also demonstrate that *Myolinc* binds to the TAR DNA-binding protein (TDP-43), a transcriptional regulator, during the myogenesis. Finally, we show that *Myolinc*–TDP-43 interaction is essential for the expression of muscle-specific genes by direct binding of TDP-43 at their promoter regions. Altogether, our study has identified a novel molecular interaction that plays a key role during differentiation of myoblasts into myotubes.

Results

Screening of novel lincRNAs enriched in the skeletal muscle

In general, lncRNAs are more tissue-specifically expressed than protein-coding genes (Cabili et al., 2011; Derrien et al., 2012; Werber et al., 2014; Clark et al., 2015; Molyneaux et al., 2015). Given that many types of lncRNAs have been identified to date, we focused on lincRNAs, which are located in between protein-coding genes, as their expression patterns could be well separated from the nearby protein-coding genes (Rinn and Chang, 2012; Rinn, 2014; Goff and Rinn, 2015). To screen for skeletal muscle-enriched lincRNAs, mouse tissues were analyzed by microarrays. As a result, 16 lincRNAs were enriched in the skeletal muscle compared to nine other tissues (brain, embryo, kidney, liver, lung, ovary, spleen, testicle, and thymus) at the threshold values of 4-fold and $P < 0.05$ (Figure 1A). Since the microarray data were prepared from whole tissues, and the skeletal muscle contains several cell types, we further analyzed the expression patterns of lincRNAs using the microarray data of the murine myoblast cell line C2C12. During differentiation, only three lncRNAs were found to be upregulated on Day 2.5, Day 5, and Day 7 of differentiation compared to the undifferentiated cells. When the lists of differentially expressed (DE) genes are compared between these two datasets, only one lincRNA is identified in both datasets, which is, '*AK142388*'. A further confirmation experiment using mouse tissues showed high expression of *AK142388* in the skeletal muscle as well as in the brain and heart tissues (Figure 1B). Based on these and the following results, we named this lincRNA '*Myolinc*'.

Myolinc is located on chromosome 9 between two protein-coding genes, *Filip1* and *Tmem30a*, on the same genomic strand. To characterize *Myolinc*, rapid amplification of cDNA ends (RACE) experiment was conducted using RNA from C2C12 myoblasts. Results showed that *Myolinc* gene encodes two isoforms (Figure 1C). While the longer isoform (4431 nt) partially overlaps the 3′-end of *Filip1* transcript, the shorter one does not (3192 nt).

By definition, lncRNAs do not encode for proteins. However, recent studies show that some lncRNAs give rise to short peptides (Anderson et al., 2015; Nelson et al., 2016; Matsumoto et al., 2017). To examine the coding potential of *Myolinc*, we next performed an *in vitro* transcription/translation assay. However, no peptide synthesis was detected in both *Myolinc* isoforms (Figure 1D).

As some lncRNAs preferentially localize to the nucleus while others are in the cytoplasm, which suggest their possible functions, the subcellular localization of *Myolinc* was investigated.

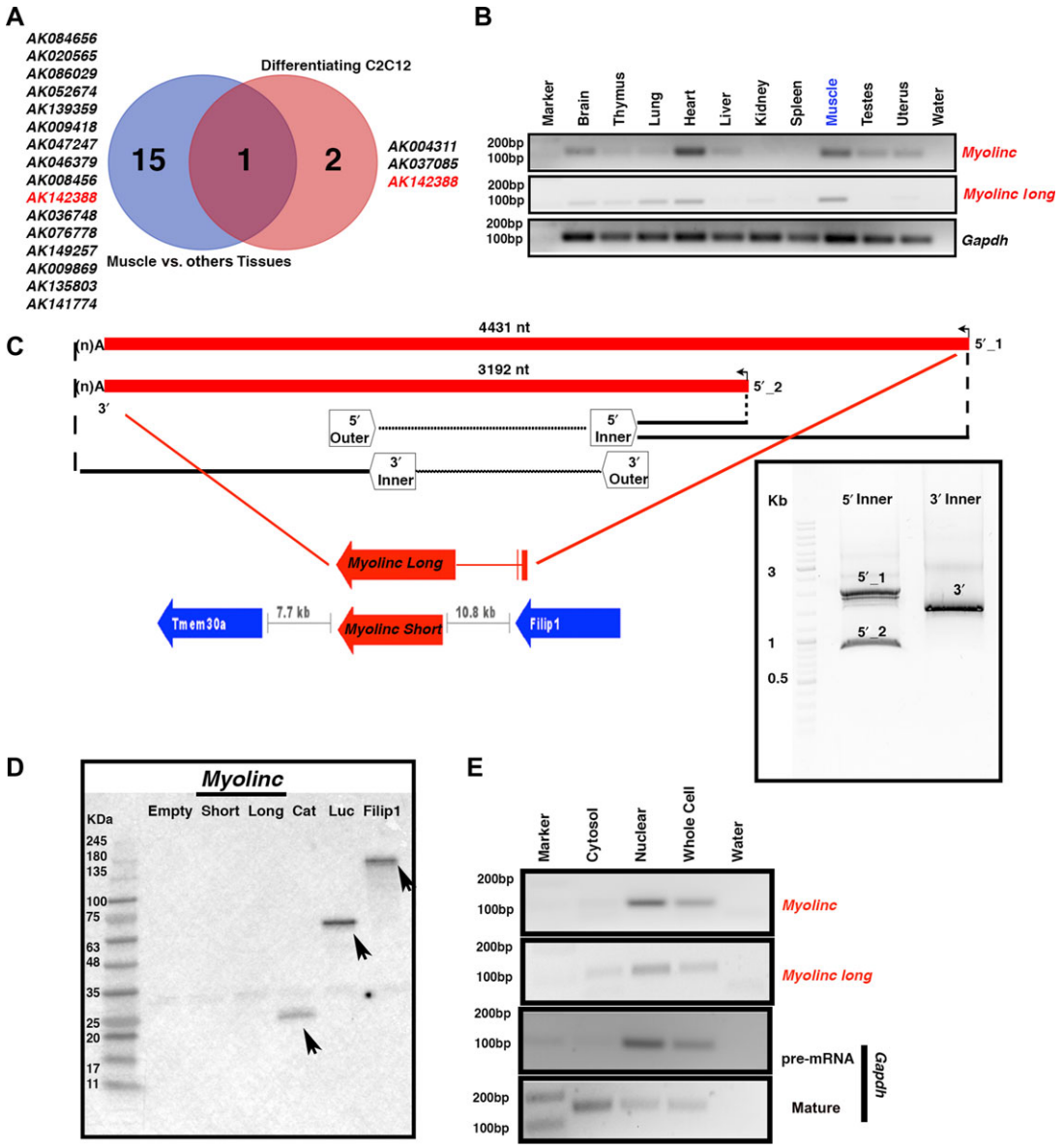

Figure 1 Screening for lincRNAs enriched in the skeletal muscle and in the differentiating C2C12 cells. (**A**) LincRNAs enriched in the skeletal muscle compared to nine other tissues and in the differentiating C2C12 cells. A Venn diagram indicates that only lincRNA *AK142388*, which we named to be '*Myolinc*', is highly enriched in both datasets. (**B**) RT-PCR result of 10 mouse tissues. While the primer pair '*Myolinc*' detects both isoforms, the longer isoform can be detected by using the primer pair targeting the sequence only present in the longer isoform but not the shorter one. Representative images from three independent experiments are shown. (**C**) Schematic representation of RACE experiment with mapping regions for the primer pairs used (upper image). Two transcription start sites (5′_1 and 5′_2) and one transcription termination site were identified. The genomic organization of *Myolinc* locus is shown in the lower image. (**D**) *In vitro* transcription/translation assay. The empty vector is used as a negative control. Compared to positive controls chloramphenicol acetyl transferase (CAT), Luciferase (Luc), and the nearby protein-coding gene *Filip1*, neither isoform of *Myolinc* encodes any peptides. (**E**) Subcellular localization. RT-PCR experiment was performed using RNA isolated from the differentiated C2C12 cells (4 days of differentiation). For controls, the pre-mRNA of *Gapdh* is shown for the nuclear fraction, while its mature form is shown for the cytosolic fraction. Representative images from three independent experiments are shown.

This analysis showed that *Myolinc* is localized in the nucleus of muscle cells (Figure 1E). Taken together, these results suggest that *Myolinc* is a nuclear localized lincRNA that is highly enriched in the skeletal muscle.

Myolinc regulates differentiation of myoblasts into myotubes

To understand the function of *Myolinc*, we first examined how the expression of *Myolinc* changes during myogenic differentiation. C2C12 myoblasts were incubated in differentiation medium (DM)

and the cells were collected at different time points to measure mRNA levels of *Myolinc* and other MRFs. The expression of *Myolinc* was hardly detectable in undifferentiated C2C12 myoblasts. However, the expression of *Myolinc* was drastically induced after 48 h of addition of DM and remained elevated at later time points. Moreover, the increase in mRNA levels of *Myolinc* correlates with the expression of *MyoD* and *Myogenin* in C2C12 myoblasts suggesting that *Myolinc* may have a role in the regulation of myogenic differentiation (Figure 2A).

To specifically investigate whether *Myolinc* plays a role in myogenic differentiation, we performed loss-of-function experiments using siRNAs. Given that the longer isoform showed the similar expression pattern to the *Myolinc* gene itself (Figure 2A), we employed siRNAs that target the shared sequences between both isoforms. C2C12 myoblasts were transfected with a scrambled siRNA or three *Myolinc* siRNA targeting different regions of *Myolinc* followed by incubation in DM. Remarkably, we found that siRNA-mediated knockdown of *Myolinc* drastically inhibited myotubes formation in C2C12 cultures (Figure 2B–D). To observe this effect at the molecular level, we measured mRNA levels of MRFs in control and *Myolinc*-knockdown C2C12 cultures after 48 h of addition of DM. Results showed that the mRNA levels of *Myf5*, an early myogenic factor, were increased, whereas mRNA levels of the late differentiation markers *MyoD* and *Myogenin* were repressed in *Myolinc*-knockdown cultures (Figure 2E). *Myomaker*, a muscle-specific transmembrane protein has been found to play a critical role in myoblast fusion during the myogenesis (Millay et al., 2013). Since knockdown of *Myolinc* inhibits myotube formation, we also investigated the effect of silencing of *Myolinc* on the expression of *Myomaker*. Interestingly, transcript levels of *Myomaker* were drastically reduced in C2C12 myoblast cultures upon knockdown of *Myolinc* (Figure 2E).

To further understand the role of *Myolinc* in the regulation of myogenesis at the molecular level, we next isolated total RNA from control and *Myolinc*-knockdown cultures incubated in DM for 48 h and performed a microarray analysis. When the thresholds of 1.4-fold and $P < 0.05$ were applied in *Myolinc*-silenced cells compared to the control, 444 (172 upregulated and 272 downregulated) DE genes were identified (Supplementary Table S1). Interestingly, gene ontology (GO) analysis for DE genes revealed a significant enrichment of GO terms related to muscle differentiation (Figure 2F). Furthermore, network analysis of DE genes suggested that *Myolinc* is involved in the regulation of myogenesis (Figure 2G). Altogether, these results suggest that *Myolinc* is an important regulator of myogenic differentiation.

Myolinc promotes myogenic differentiation through augmenting the expression of Filip1

Myolinc is located between two protein-coding genes, *Filip1* and *Tmem30a*, on the same genomic strand. It is now increasingly clear that lncRNAs regulate transcription of a number of genes in *cis*- or *trans*-manner. Thus, we investigated whether *Myolinc* regulates the expression of *Filip1* and *Tmem30a*. We first investigated how the levels of *Filip1* and *Tmem30a* change during the myogenic differentiation. Interestingly, the expression of *Filip1* was drastically increased after the addition of DM (Figure 3A). There was also a significant increase in the mRNA levels of *Tmem30a* in C2C12 cultures after the addition of DM (Figure 3B). We next investigated whether *Myolinc* plays any role in the expression of *Filip1* and *Tmem30a*. Intriguingly, knockdown of *Myolinc* caused a drastic reduction in the mRNA levels of *Filip1* suggesting a *cis*-regulation of *Filip1* by *Myolinc* (Figure 3C). By contrast, there was no significant effect of silencing of *Myolinc* on the expression of *Tmem30a* (Figure 3D).

Filip1 is a filament interacting protein that has been implicated in the regulation of cell motility (Nagano et al., 2002). However, the role of *Filip1* in myogenesis has not been investigated previously. Interestingly, knockdown of *Filip1* inhibited the formation of multi-nucleated myotubes in C2C12 cultures (Figure 3E and F). Moreover, knockdown of *Filip1* also reduced the mRNA levels of *MyoD* and *Myogenin* in C2C12 myoblasts after induction of the differentiation program (Figure 3G). Surprisingly, we found that silencing of *Filip1* reduced the expression of *Myolinc*, but not *Tmem30a*, in differentiating myoblasts suggesting that *Myolinc* and *Filip1* positively regulate each other's expression and both of them are involved in myogenic differentiation (Figure 3G).

Myolinc interacts with TDP-43 during myogenic differentiation

To further elucidate the mechanisms of action of *Myolinc*, we performed RNA pull-down experiment by labeling *Myolinc* with biotin and mixing it with cellular lysates of differentiated C2C12 cells. Compared to the negative control (i.e. antisense biotinylated *Myolinc*), a unique band was detected in the sense biotinylated *Myolinc* sample (Figure 4A). The band was analyzed by mass spectrometry in comparison to the similar region in the antisense *Myolinc* sample, which identified 10 proteins that are specifically detected in the sense *Myolinc* sample (Figure 4A). Of these, three potential binding partners (Hnrnpa0, Hnrnpd, and Tardbp) were tested individually by RNA pull-down followed by western blotting, which identified Tardbp (also known as 'TDP-43', which will be used hereafter) as a binding partner (Figure 4B and Supplementary Figure S1). To further confirm this binding, a reciprocal experiment was performed. We performed RNA immunoprecipitation (RIP) followed by RT-PCR. As shown in Figure 4C, *Myolinc* strongly binds to TDP-43, which confirms TDP-43 as a binding partner of *Myolinc*.

TDP-43 is a DNA/RNA-binding protein that has been linked to many important biological processes, such as splicing, mRNA stability, and regulation of transcription (reviewed in Lee et al., 2011). However, its role in myogenesis is unknown. Our analysis showed that the levels of TDP-43 are increased upon the addition of DM in C2C12 myoblasts (Figure 4E). By performing immunohistochemistry, we next investigated the cellular localization of TDP-43. Our analysis showed that TDP-43 is localized to the nuclei of myotubes in C2C12 cultures (Figure 4D). We next investigated the effect of knockdown of *TDP-43* on myogenic differentiation. Results showed that knockdown of *TDP-43* reduced the expressions of *MyoD* and *Myogenin*, whereas the expression of *Myf5* was increased (Figure 4F). This trend is similar to that of *Myolinc*-knockdown

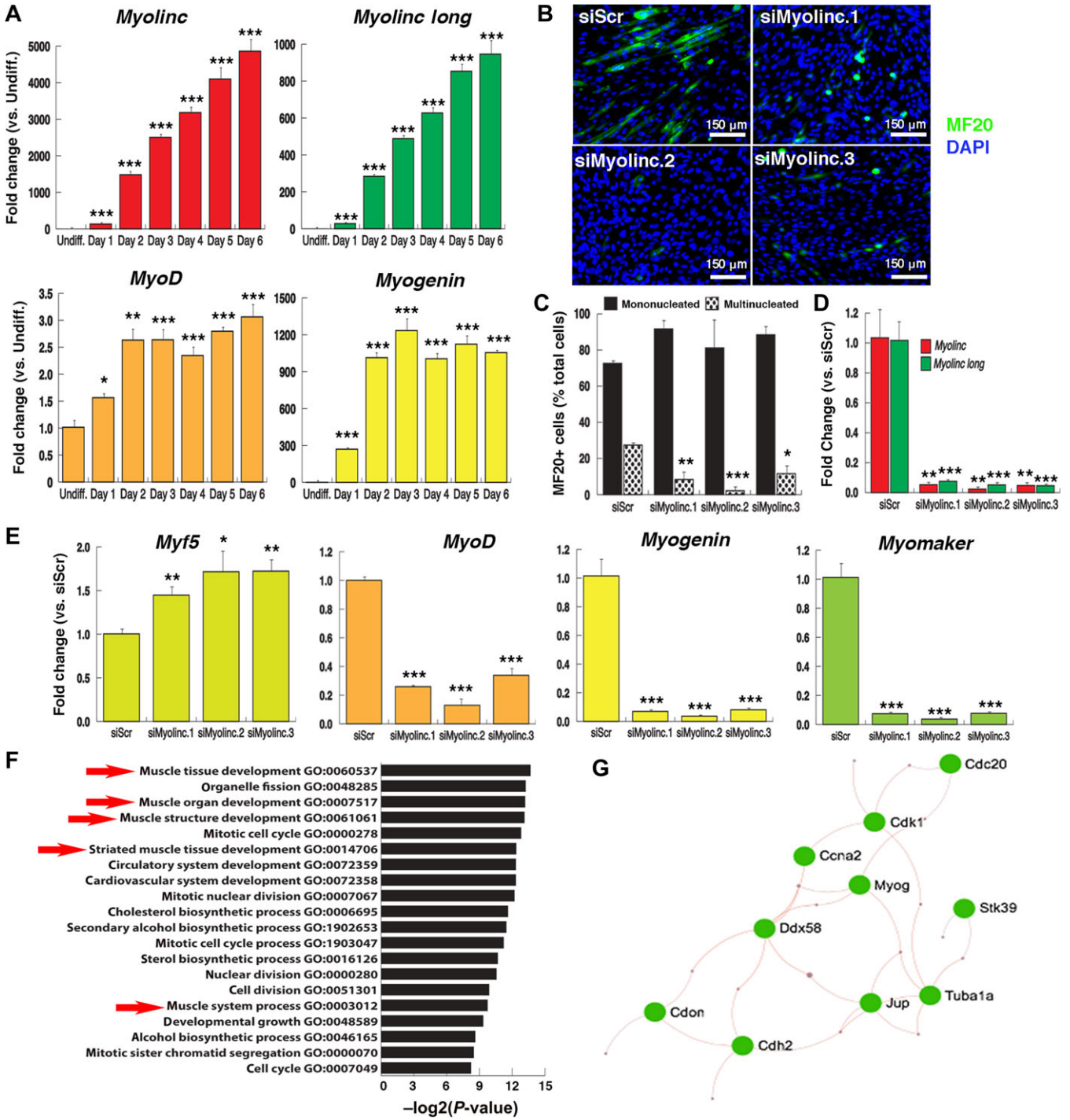

Figure 2 Functional characterization of *Myolinc* in C2C12 cells. (**A**) Expression analysis of *Myolinc, MyoD*, and *Myogenin* during the differentiation of C2C12 cells. Error bars represent standard error of the means (SEM). $n = 3$. (**B**) Representative images of myosin-positive differentiated myocytes after transfection of siRNAs against *Myolinc* (siMyolinc) or a scrambled siRNA as control (siScr). After 4 days of differentiation, C2C12 cells were stained with MF20 and counterstained with DAPI to visualize the nuclei. (**C**) Counting of mono- and multi-nucleated myosin-positive myocytes. Compared to siScr, the number of multi-nucleated myocytes decreased significantly after 4 days of differentiation in *Myolinc*-silenced cells. Error bars represent SEM. Ten random fields were selected for counting stained cells in each condition. (**D**) Silencing of *Myolinc*. Twenty-four hours after the transfection of siRNAs, C2C12 cells were differentiated for 2 days before collecting total RNA. Error bars represent SEM. $n = 3$. (**E**) Gene expression analysis of muscle regulatory factors upon silencing of *Myolinc*. Error bars represent SEM. $n = 3$. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.005$. (**F**) Top 20 enriched GO terms classified under the biological processes category. (**G**) Network analysis of downregulated genes. The names of downregulated genes are shown. The network shown is according to the myogenesis category of the REACTOME database (http://reactome.org/).

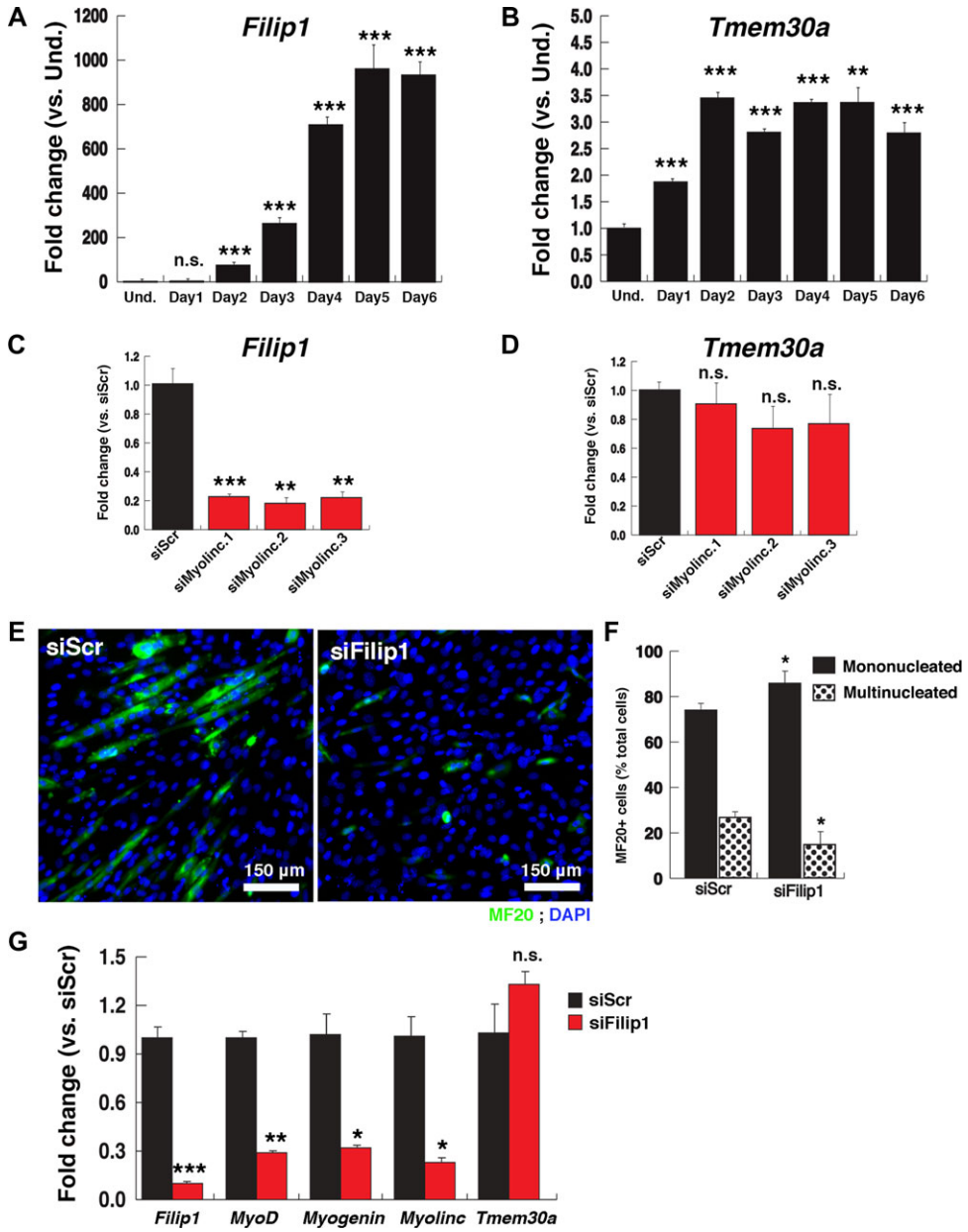

Figure 3 *Myolinc* regulates *in-cis Filip1*. (**A** and **B**) Expression analysis of *Filip1* (**A**) and *Tmem30a* (**B**) during the differentiation of C2C12 cells. Error bars represent SEM. $n = 3$. (**C** and **D**) Gene expression analysis of *Filip1* (**C**) and *Tmem30a* (**D**) upon silencing of *Myolinc* in C2C12 cells (2 days of differentiation). Error bars represent SEM. $n = 3$. (**E**) Representative images of myosin-positive differentiated myocytes after transfection of the siRNA against *Filip1* (siFilip1) or a scrambled siRNA as control (siScr). After 4 days of differentiation, C2C12 cells were stained with MF20 and counterstained with DAPI to visualize the nuclei. (**F**) Counting of mono- and multi-nucleated myosin-positive myocytes. Compared to siScr, the number of multi-nucleated myocytes decreased significantly after 4 days of differentiation in *Filip1*-knockdown cells. Error bars represent SEM. Ten random fields were selected for counting stained cells in each condition. (**G**) Gene expression analysis of *Filip1, MyoD, Myogenin, Myolinc*, and *Tmem30* upon silencing of *Filip1*. Error bars represent SEM. $n = 3$. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.005$. 'n.s.' represents 'not statistically significant'.

cells (Figure 2E). Interestingly, the expression of both *Myolinc* and *Filip1*, and to a lesser extent of *Tmem30a*, was downregulated upon the knockdown of *TDP-43* (Figure 4F). These results suggest that both *Myolinc* and *Filip1* genes are under the control of TDP-43 protein. To further confirm these results, *TDP-43 and Myolinc* were silenced in cultured primary myoblasts, which resulted in a clear impairment of the differentiation in terms of multi-nucleated myotubes formation (Figure 4G and H).

Next, *Myolinc* and *TDP-43* were overexpressed in C2C12 cells and differentiated for 3 days (Figure 5A and B). As evident from

Figure 5C, the percent distributions of mono- and multi-nucleated myocytes (MF20+ cells) after 3 days of differentiation are similar for the overexpression of *Myogenin*, *Myolinc*, and *TDP-43*. Indeed, compared to the control condition (empty vector), the overexpression of *Myolinc* and *TDP-43* caused an increase in the number of multi-nucleated myocytes. Interestingly, when *Myolinc* and *TDP-43* were co-overexpressed, the percent of multi-nucleated myocytes was even higher than overexpression of each gene alone. These data further suggest the involvement of such genes in the differentiation process, although the percent of multi-nucleated myocytes over mono-nucleated ones is much higher for the overexpression of *MyoD*, which is consistent with the literature that forced expression of *MyoD* drives myogenesis (Molkentin et al., 1995; Yang et al., 2009). Taken together, along with *Myolinc*, TDP-43 is important for the differentiation of myoblasts into myotubes.

Myolinc promotes proper binding of TDP-43 on the promoters of skeletal muscle genes

Given that both *Myolinc* and TDP-43 are preferentially localized to the nucleus of C2C12 cells (Figures 1E and 4D), we examined the role of TDP-43 as a transcriptional regulator by performing chromatin immunoprecipitation (ChIP) followed by next-generation sequencing (ChIP-seq) and compared it to the results of MyoD and Myogenin binding. On Day 2 after incubation of C2C12 cells in DM, the classification of bound regions by TDP-43 is more similar to that of RNA polymerase II ('Pol2') than MyoD and Myogenin when normalized to that of IgG (negative control) (Figure 6A and Supplementary Table S2). Moreover, TDP-43 showed preferential binding to exons and introns, which is consistent with the known function of TDP-43 in splicing (Baralle et al., 2013). In order to confirm ChIP-seq result of TDP-43, a biological replicate experiment was performed, which was sequenced twice to rule out the possibility of sequencing errors. In all three ChIP-seq experiments, a similar distribution of bound regions was obtained (Figure 6A).

Since our objective was to understand the role of TDP-43 as a transcriptional regulator, we next analyzed the bound regions of TDP-43 against known binding motifs. Interestingly, TDP-43 bound regions overlap to those of MyoD, Myogenin, and Myf5 in a statistically significant manner (Figure 6B and Supplementary Table S3). Next, genes that were identified to be bound by TDP-43, MyoD, and/or Myogenin were compared using identified regions categorized under the promoter-TSS ('transcription start site') (Figure 6C and Supplementary Table S4). Interestingly, TDP-43 shared the binding to promoters of genes enriched for MyoD and/or Myogenin-binding sites (209 (16.11%) out of 1267 genes bound by TDP-43), suggesting that the *Myolinc*–TDP-43 complex might play a role in the regulation of differentiation of C2C12 myoblasts.

We next performed the same set of ChIP-seq experiments after silencing of *Myolinc* in C2C12 cells on Day 2 after the addition of DM. A comparison of TDP-43-binding sites between normally differentiated and *Myolinc*-knockdown differentiated C2C12 cells showed 55 shared regions (Supplementary Table S5). In contrast, there are 1241 and 575 regions only found in normally differentiated and *Myolinc*-knockdown C2C12 cells, respectively. When genes that share binding regions for TDP-43 in their promoters were compared, there were 54 genes shared and 1162 and 519 genes only found in normally differentiated and *Myolinc*-knockdown C2C12 cells, respectively (Figure 6D), suggesting a reduced binding of TDP-43 to promoter regions in *Myolinc*-knockdown cells. Interestingly, when GO analysis was performed to those 1162 genes whose promoters are found to be bound by TDP-43 only in normally differentiated C2C12 cells but not in *Myolinc*-knockdown cells, GO terms related to transcriptional regulations were enriched (Figure 6E). To obtain a global view of signaling networks that were affected by silencing *Myolinc*, network analysis was performed and extracted for those genes involved in the 'regulation of transcription from RNA polymerase II promoter (GO:0006357)' (Figure 6F). From the network analysis, we identified MyoD to be highly connected and regulated upon silencing of *Myolinc*, which was supported by qRT-PCR results that indicated downregulation of *MyoD* upon silencing of *Myolinc* (Figure 2E). When peak intensities were compared to the input controls, it is clear that binding of TDP-43 to the promoter of *MyoD* is reduced upon silencing of *Myolinc* compared to the control condition (Figure 6G), which was further confirmed by ChIP-PCR experiment (Figure 6H). Furthermore, the reduction of TDP-43 binding to promoters of differentiation markers genes (e.g. *Acta1*, *Ccnd1*, *Tnnc1*, *Tnni1*) after silencing of *Myolinc* was also confirmed by ChIP-PCR experiments (Figure 6H). Interestingly, the binding of TDP-43 to the promoter of *Filip1* was also reduced upon silencing of *Myolinc* (Figure 6H), suggesting the *cis*-regulation and confirming our previously obtained gene expression data (Figure 3C). Taken together, our data suggest that *Myolinc* is important for the specific binding of TDP-43 on the promoters of genes that are important regulators of muscle differentiation.

Myolinc is involved in skeletal muscle regeneration in vivo

Our preceding findings suggest that *Myolinc* is an important regulator of myogenic differentiation *in vitro*. To understand the physiological significance of *Myolinc* in myogenesis, we next investigated the role of *Myolinc* in skeletal muscle regeneration in adult mice. Tibialis anterior (TA) muscle of mice was injured by intramuscular injection of snake venom cardiotoxin (CTX). On Day 2 and Day 3 post CTX injection, siRNA against *Myolinc* was injected into the left TA muscles of mice. In the same mouse, the siRNA against scramble sequence ('siScr') was injected in the right TA muscle as a control. Then, on the fourth and fifth day, mice were sacrificed (Figure 7A) for the quantification of knockdown. As shown in Figure 7B, statistically significant knockdown of *Myolinc* was achieved only in the regenerated muscle. The expression levels of *Myolinc* in satellite cells and myoblasts were negligible and below the detection limit (Figure 7C). Of note, the expression of *Myolinc, Filip1*, and *TDP-43* was upregulated during the muscle regeneration (Supplementary Figure S3A), further confirming that these genes are involved in the muscle differentiation process. For our further analysis, we used 5 days post injury as a collection time point. To understand the effect of knockdown of *Myolinc* on

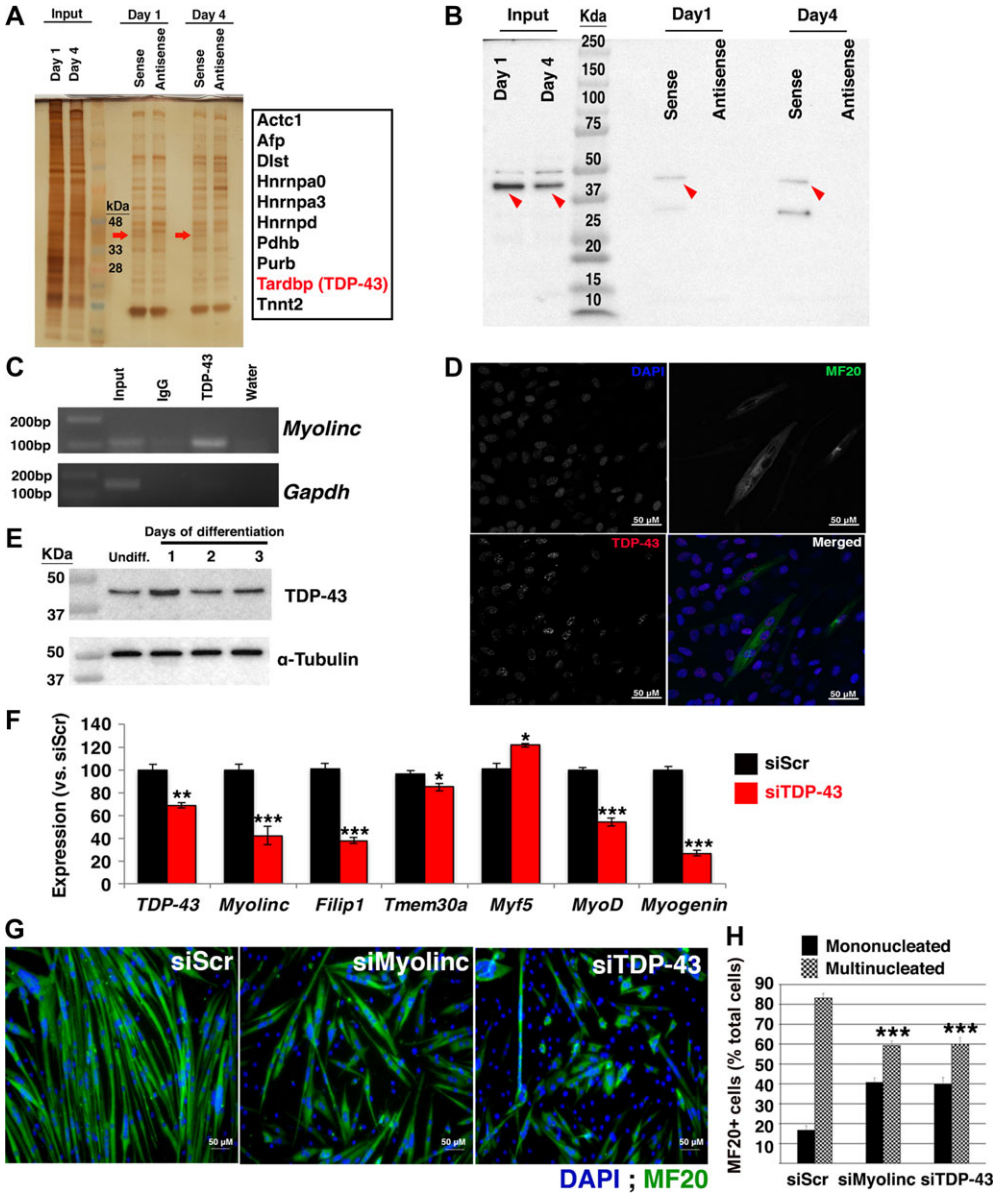

Figure 4 *Myolinc* interacts with TDP-43. (**A**) Silver staining after RNA pull-down. A unique band (indicated by red arrows) in the sense biotinylated *Myolinc* compared to the antisense biotinylated *Myolinc* was detected. The region that contains the unique band was cut as well as the corresponding region in the antisense sample. Bands were analyzed by mass spectrometry, which identified 10 proteins that are specific to the sense *Myolinc* (shown in the box). Representative images from three independent experiments are shown. (**B**) Western blotting after RNA pull-down using anti-TDP-43 antibody. The red arrowheads indicate the specific band for TDP-43 protein. Representative images from three independent experiments are shown. (**C**) RIP-PCR using anti-TDP-43 antibody. The strong binding of TDP-43 to *Myolinc* was observed in comparison to the negative control using anti-IgG antibody. Furthermore, the binding of TDP-43 to *Myolinc* is specific as it can be compared to the result of *Gapdh*, which shows the binding in the input but not in other samples. Representative images from three independent experiments are shown. (**D**) Immunofluorescence of TDP-43 in C2C12 cells on Day 2 of the differentiation. (**E**) Expression of TDP-43 protein during the differentiation of C2C12 cells. Anti-α-Tubulin antibody was used as a loading control for western blotting assay. (**F**) Gene expression analysis of muscle regulatory factors upon silencing of TDP-43. Error bars represent SEM. $n = 3$. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.005$. (**G**) Representative images of myosin-positive differentiated myocytes after transfection of siRNAs against *Myolinc* (siMyolinc), *Tardbp* (siTDP-43), or a scrambled siRNA as control (siScr). After 2 days of differentiation of primary myoblasts, the differentiated myocytes were stained with MF20 and counterstained with DAPI to visualize the nuclei. (**H**) Counting of mono- and multi-nucleated myosin-positive myocytes derived from primary myoblasts. Compared to siScr, the number of multi-nucleated myocytes decreased significantly after 2 days of differentiation in *Myolinc*- and *TDP-43*-knockdown cells. Six random fields were selected for counting stained cells in each condition. Error bars represent SEM. $^{***}P < 0.005$.

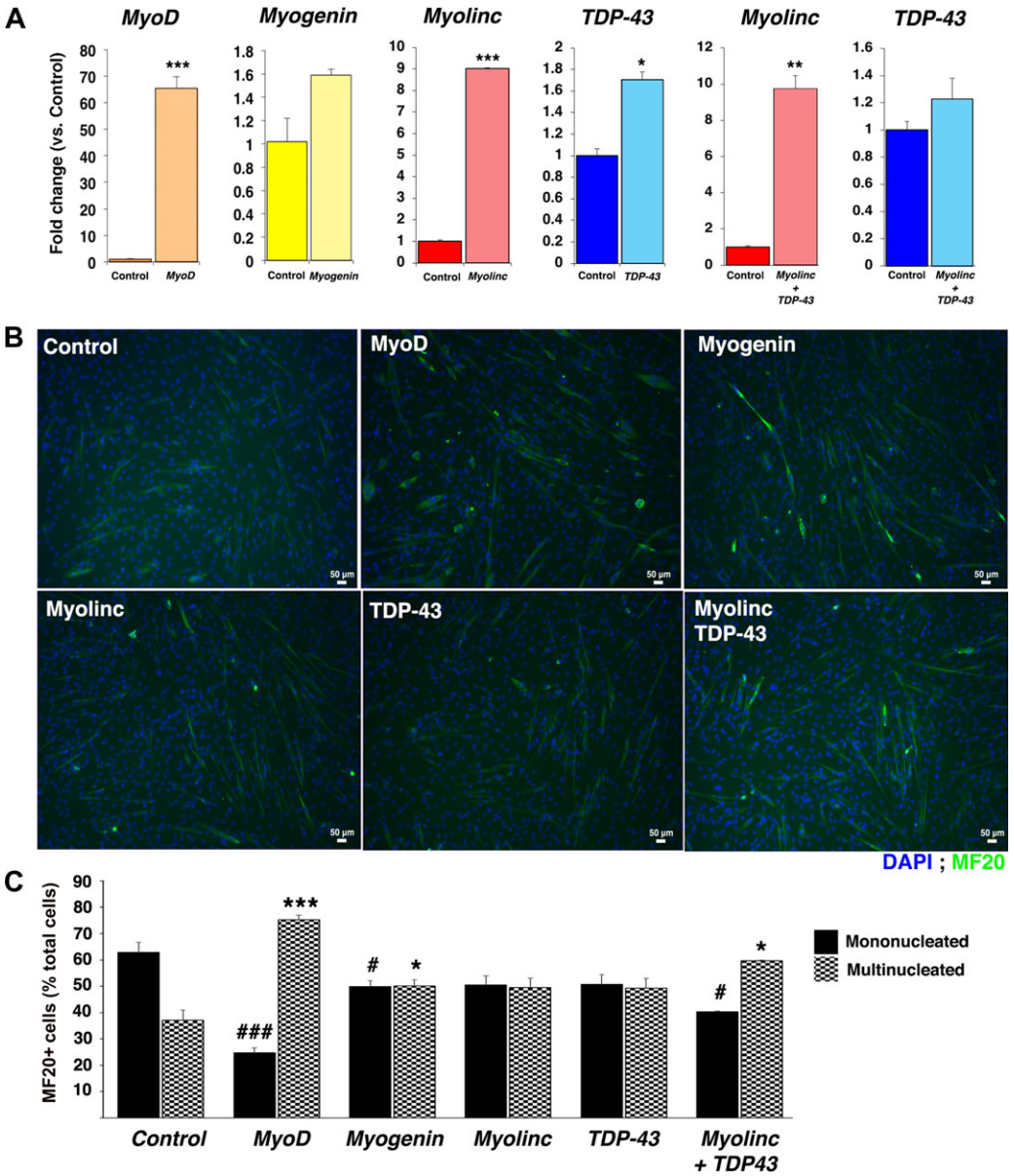

Figure 5 Overexpression of *Myolinc* and *TDP-43*. (**A**) qRT-PCR result of the target genes. The plasmids containing empty vector, *MyoD*, *Myogenin*, *Myolinc*, *TDP-43*, or *Myolinc* plus *TDP-43* were overexpressed in C2C12 cells and differentiated for 3 days. $n = 2$. (**B**) Representative images of differentiated myocytes upon overexpression of plasmids containing the target genes. (**C**) Counting of mono- and multi-nucleated myosin-positive myocytes. The statistics were calculated against the overexpression of empty vector (control). Three random fields were selected for counting stained cells in each condition. Error bars represent SEM. $^{*/\#}P < 0.05$, $^{**}P < 0.01$, $^{***/\#\#\#}P < 0.005$.

skeletal muscle regeneration, transverse sections of control and *Myolinc* siRNA-transfected TA muscles were generated and immunostained with an antibody against embryonic myosin heavy chain (eMyHC) followed by morphometric analysis (Figure 7D). Frequency distribution histograms representing the fiber cross-sectional area (CSA) showed that the percentage of eMyHC$^+$-myofibers with smaller CSA was increased, whereas those with higher CSA was reduced in TA muscle transfected with *Myolinc* siRNA compared to that transfected with scramble sequence siRNA (Figure 7E). Moreover, the average area filled with eMyHC$^+$ myofibers was significantly reduced in TA muscle transfected with *Myolinc* siRNA compared to that transfected with scramble sequence siRNA (Figure 7F) further suggesting that silencing of *Myolinc* inhibits skeletal muscle regeneration in response to injury. We also studied muscle regeneration by performing immunostaining with an antibody against Myogenin. This analysis showed that the number of Myogenin$^+$ cells was significantly reduced in the TA muscle injected with *Myolinc* siRNA compared to that injected with siScr control (Figure 7G and H). These results further suggest that *Myolinc* is involved in skeletal muscle regeneration *in vivo*.

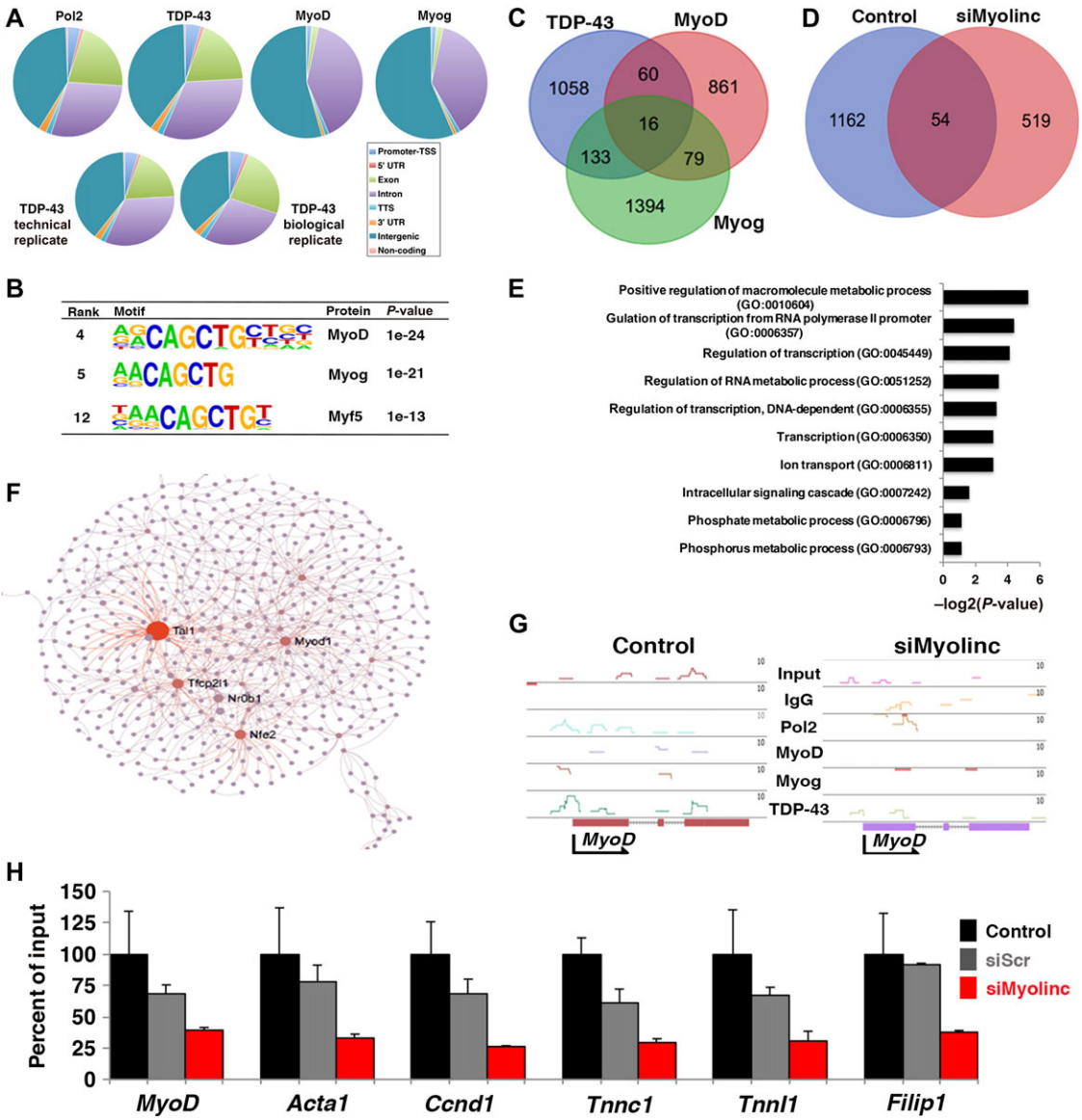

Figure 6 ChIP-seq analysis. (**A**) Distributions of identified peaks against the IgG control. The annotation information was provided via HOMER (Heinz et al., 2010). The identified peak profile of TDP-43 is more similar to that of Pol2 than those of MyoD and Myogenin. (**B**) Motifs of MRFs identified in TDP-43 ChIP-seq data. The whole list can be found in Supplementary Table S3. (**C**) Venn diagram showing numbers of genes whose promoter regions were bound by corresponding factors. The whole list can be found in Supplementary Table S4. (**D**) Venn diagram showing the number of genes whose promoter regions bound by TDP-43 in the normally differentiated ('Control') and *Myolinc*-knockdown ('siMyolinc') differentiated C2C12 cells. The whole list can be found in Supplementary Table S5. (**E**) GO analysis of genes found only in Control compared to siMyolinc. Top 10 most enriched GO terms are shown. (**F**) Network analysis of genes classified under 'regulation of transcription from RNA polymerase II promoter (GO:0006357)'. Names of highly connected hub genes are indicated. (**G**) Genomic view of the promoter of *MyoD* and its gene body (chr7:46376000–46379200). Binding of TDP-43 to the promoter of *MyoD* is reduced upon silencing of *Myolinc* compared with Control. (**H**) ChIP-qPCR targeting the promoters of selected genes. Error bars represent SEM. $n = 3$.

Discussion

Skeletal myogenesis is a tightly regulated process at both transcriptional and post-transcriptional levels. Accumulating evidence suggests that lncRNAs play a prominent role in the regulation of myogenesis (Simionescu-Bankston and Kumar, 2016). However, so far, only a few lncRNAs have been identified to play a role in myogenesis. In this study, we performed a genome-wide gene expression analysis that led to the identification of a novel muscle-enriched lncRNA *Myolinc*. Our systematic analysis revealed that *Myolinc* is important for the differentiation of myoblasts into myotubes. Furthermore, our experiments showed that knockdown of *Myolinc* inhibits skeletal muscle regeneration in

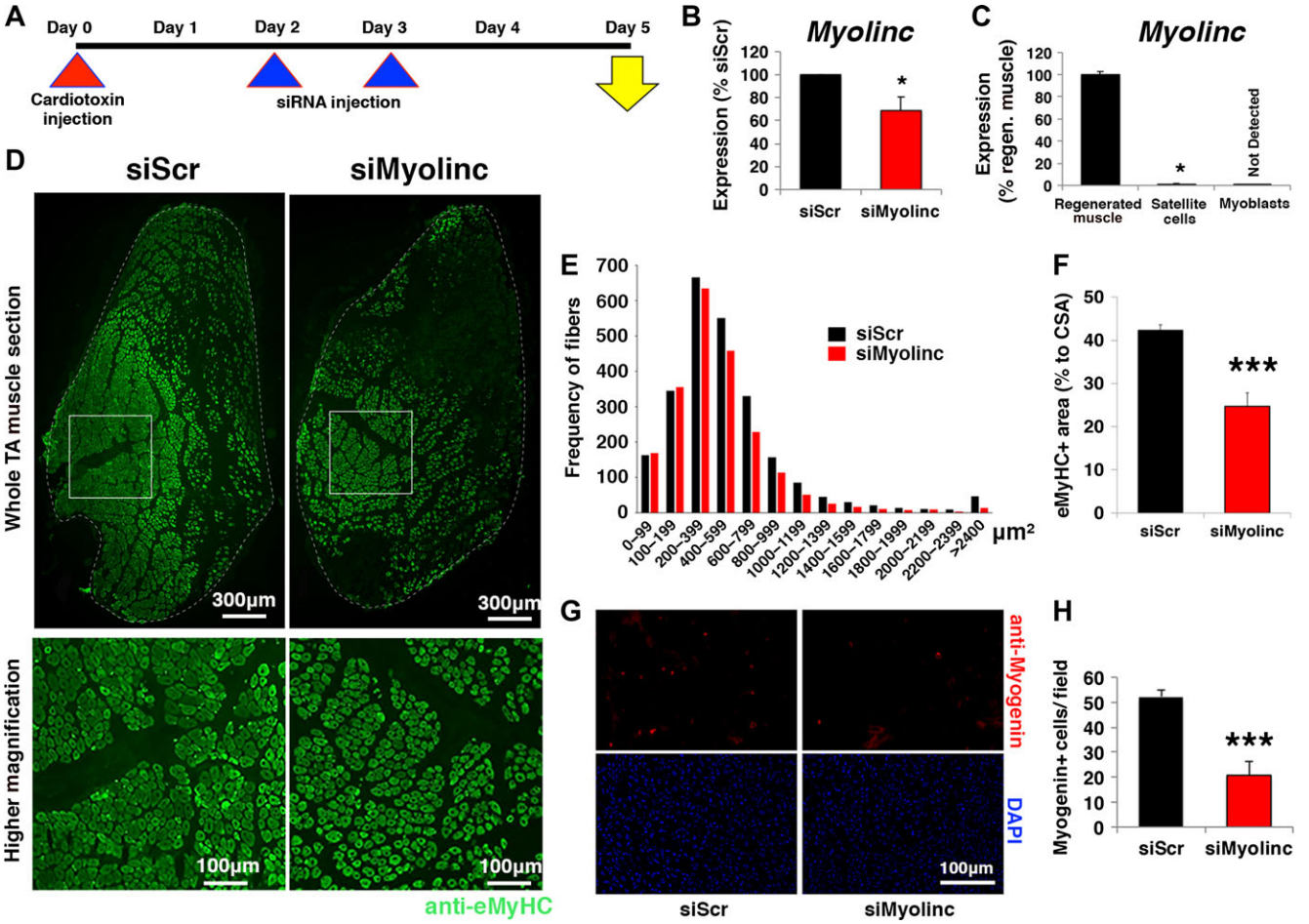

Figure 7 *In vivo* knockdown of *Myolinc*. (**A**) Scheme for injections. (**B**) Quantification of knockdown. Error bars represent SEM. $n = 6$ mice per group ($n = 2$ mice for Day 4 and $n = 4$ mice for Day 5 after CTX injection). (**C**) Quantification of *Myolinc* expression in regenerated muscle (Day 5 after CTX injection in siScr-treated mice; $n = 2$ mice), satellite cells (isolated from non-treated mice; $n = 2$ mice), and myoblasts (isolated from Day 2 after CTX injection; $n = 2$ mice). Error bars represent SEM. (**D**) Representative eMyHC staining images of siScr and siMyolinc-injected muscles. Compared to control, upon silencing of *Myolinc*, the regeneration of damaged muscles is impaired. Magnified images of white squared areas are shown below. $n = 4$ mice per group. (**E**) Frequency of eMyHC-positive myofibers per imaging field. The sizes of eMyHC-positive myofibers were quantified and separated based on their sizes as indicated. $n = 4$ mice per group. (**F**) eMyHC-positive area per CSA comparing between siScr and siMyolinc-injected muscles. Error bars represent SEM. $n = 4$ mice per group. (**G**) Representative images of sections stained with anti-Myogenin antibody (red) and counterstained with DAPI (blue). (**H**) Frequency of Myogenin-positive cells per imaging field. Error bars represent SEM. $n = 4$ mice per group. $^*P < 0.05$, $^{***}P < 0.005$.

adult mice. Interestingly, the expression of *Myolinc* was significant lower in *mdx* mice compared to the age-matched wild-type mice (Supplementary Figure S3B). Moreover, *Myolinc* expression was higher in slow-twitching soleus muscle compared to fast-twitching extensor digitorum longus muscle (Supplementary Figure S3C), which is in line with the previous findings showing that slow-twitch muscles have a better regenerative potential (Darr and Schultz, 1987; Kalhovde et al., 2005).

To understand the mechanism of *Myolinc* controlling the muscle differentiation, we performed loss-of-function experiment followed by molecular profiling by microarrays (Figure 2). Since a number of genes are differentially regulated upon silencing of *Myolinc* compared to the control (Figure 2F), bioinformatics analysis was performed to understand these changes. In particular, by taking the advantage of known information (e.g. GO terms, protein–protein interactions), network analysis was performed, which highlighted the most connected (i.e. hub genes) under the myogenesis category to be MyoD (Figure 6F). This trend was confirmed by further experiments, including qRT-PCR (Figure 2E) and ChIP-seq/PCR experiments (Figure 6). Mechanistically, our experiments showed that *Myolinc* interacts with a DNA/RNA-binding protein TDP-43, which binds to the promoter of *MyoD* and other muscle marker genes to control signaling pathways necessary for the differentiation of myoblasts into myofibers (Figure 6). Based on these results, we conclude that *Myolinc* functions as a molecular switch during the differentiation of proliferating myoblasts into myocytes and subsequently into myotubes. Indeed, silencing of *Myolinc* caused an upregulation of *Myf5* (Figure 2E), which

plays a pivotal role in the proliferation of myoblasts (Ustanina et al., 2007; Francetic and Li, 2011). On the other hand, silencing of *Myolinc* resulted in the downregulation of *MyoD* (Figure 2E), which is a master regulator of the muscle differentiation process (Rawls et al., 1998; Bergstrom et al., 2002; Blum et al., 2012). Furthermore, our findings indicate that *Myolinc* is an important regulator for the proper expression of MRFs at the onset the differentiation. Impairments at the early stages of differentiation inevitably and consequentially resulted in the perturbation of later stages of differentiation as we observed in terms of myotubes formation *in vitro* (Figure 2B–D), downregulation of late muscle markers (e.g. *Myogenin* and *Myomaker*) (Figure 2E), and *in vivo* muscle regeneration following injury (Figure 7).

It is now evident that lncRNAs can guide changes in gene expression either in a *cis*- (on neighboring genes) or *trans*- (distantly located genes) manner (Simionescu-Bankston and Kumar, 2016). *Myolinc* is flanked by two protein-coding genes *Filip1* and *Tmem30a*. In addition to *Myolinc*, we found that the expression of both *Filip1* and *Tmem30a* is increased during the myogenic differentiation (Figure 3A and B) suggesting that this gene locus is activated during the myogenesis. Interestingly, we found that *Myolinc* is needed for the expression of *Filip1*, but not *Tmem30a*. Moreover, silencing of *Filip1* also reduces the expression of *Myolinc* suggesting a co-operative interaction between them during the myogenesis. Published reports suggest that Filip1 interacts with Filamin A to control the start of neocortical cell migration from the ventricular zone (Nagano et al., 2002, 2004). While the role of Filip1 in myogenesis was not previously investigated, our results provide initial evidence that *Filip1* mediates the expression of MRFs and differentiation of myoblasts into myotubes (Figure 3E–G). However, the transcriptional mechanism by which the expression of *Myolinc* and *Filip1* is upregulated during myogenesis remains unknown. This is an area of future research.

Another mechanism by which lncRNAs can regulate cellular function is through binding to specific proteins and modulating their activity (Flynn and Chang, 2014; Uchida and Bolli, 2017). In this study, we have identified TDP-43 to be the molecular partner of *Myolinc*. In recent years, there have been an increased number of reports about TDP-43 as its aggregation in the cytoplasm and absence in the nucleus is one of the characteristics of most sporadic and familial forms of amyotrophic lateral sclerosis (Baralle et al., 2013). In myocytes, a cytoplasmic accumulation of TDP-43 was observed in various forms of myopathy patients (e.g. myotilinopathy and desminopathy) (Olive et al., 2009). Mechanistically, TDP-43 has been shown to interfere with the loading of mature *miR-1* and *miR-206* into the RNA-induced silencing complex in C2C12 cells (King et al., 2014), which indicates an interaction of TDP-43 with non-coding RNAs in the cytoplasm. In the case of lncRNAs, there are two lncRNAs shown to interact with TDP-43: *gadd7* (Liu et al., 2012) and *lncLSTR* (Li et al., 2015). Similar to *lncLSTR*, *Myolinc* is preferentially expressed in the nucleus (Figure 1E). Thus, the functionality of TDP-43 as a transcriptional regulator was investigated in this study. Up to now, there have been only three reports about the transcriptional activity of TDP-43 showing that TDP-43 suppresses the transcription of human immunodeficiency virus type 1 (HIV-1) by binding to the HIV-1 long terminal repeat (Ou et al., 1995) and negatively controls expressions of *Acrv1* (Lalmansingh et al., 2011) and *Cyp8b1* (Li et al., 2015). Here, we propose that TDP-43 also functions as a transcriptional activator in the context of myogenesis as our ChIP results show the binding of TDP-43 to promoters of muscle marker genes (e.g. *MyoD* and *Acta1*), which are upregulated upon TDP-43 binding (Figure 6H). However, our results are based on ChIP assay in which cells were fixed and crosslinked. Thus, native ChIP assay along with promoter–reporter assay coupled with point mutations on the TDP-43-binding sites must be performed to clearly define the requirement for TDP-43 binding in the activation of muscle marker genes. Furthermore, the binding of TDP-43 to the promoters of muscle marker genes was reduced when the interacting lncRNA *Myolinc* was silenced. Importantly, the molecular mechanism by which TDP-43 regulates the transcription remains elusive. Of note, TDP-43 has been shown to interact with a number of nuclear proteins including transcriptional regulators (Buratti et al., 2005; Freibaum et al., 2010). Thus, its function as transcriptional activator or repressor may depend on the presence of context-specific molecular cognates. Further studies are needed in order to shed light on the many functions of this important protein.

There are some limitations to the current study. First, the *in vitro* experiments were performed with C2C12 cells, which may not reflect the physiological condition. To this shortfall, we

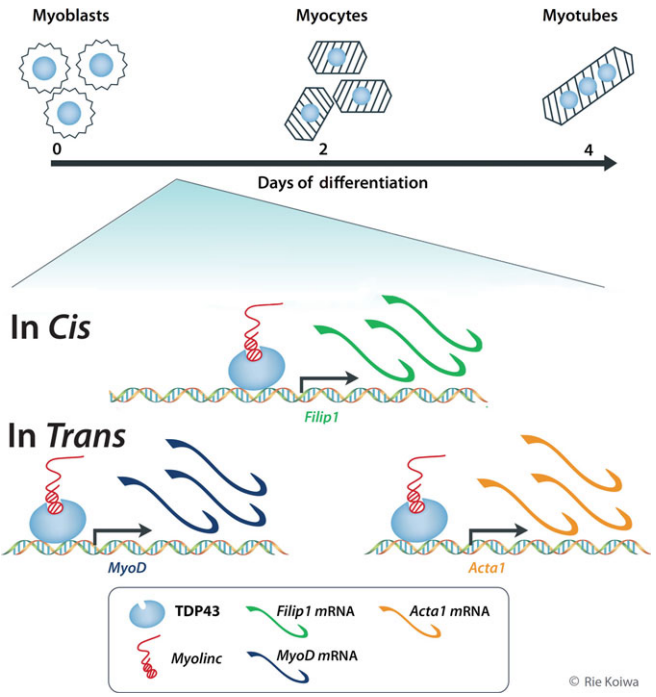

Figure 8 A model proposed for *Myolinc*. *Myolinc* recruits TDP-43 to the promoters of *Filip1* and muscle marker genes (e.g. *MyoD* and *Acta1*) to regulate myogenic regulatory networks for the differentiation of myoblasts into myocytes and the fusion of myocytes to myotubes.

silenced *Myolinc* and *TDP-43* in cultured primary myoblasts, which showed the same phenotype as C2C12 cells (Figure 4G and H). Moreover, the muscle regeneration experiment upon CTX injection was performed (Figure 7), which showed delayed regeneration due to silencing of *Myolinc*. However, longer follow-up study has not been performed, which is necessary to understand the prolonged effects of silencing of *Myolinc*. Of note, during the muscle regeneration, the expression of *Myolinc* is significantly elevated (Supplementary Figure S3A), which will be difficult to suppress the expression of *Myolinc* efficiently. To this shortfall, it is necessary to generate genetic ablation of *Myolinc in vitro* and *in vivo* using classical gene targeting or CRISPR/Cas9 system.

Based on our results, we propose a model for the mechanisms of action of *Myolinc* during myogenesis (Figure 8). At the onset of differentiation, *Myolinc* recruits TDP-43 to the promoter of *Filip1* to regulate its transcription in a *cis*-manner. At the same time, the complex between *Myolinc* and TDP-43 binds to promoters of muscle marker genes (e.g. *MyoD* and *Acta1*) to regulate myogenic regulatory networks for the differentiation of myoblasts into myocytes and the fusion of myocytes to myotubes. Altogether, our study has identified a novel lncRNA and molecular interactions involved in myogenic differentiation.

Materials and methods

Cell lines

Murine myoblast cell line C2C12 cells were cultured in growth medium consisting of DMEM with high glucose and pyruvate (Thermo Fisher Scientific, #41966-029) supplemented with 20% FBS (Thermo Fisher Scientific, #10270106), antibiotics (100 units of penicillin and 100 μg of streptomycin per ml, Sigma-Aldrich, #P4333) at 37°C in a humidified atmosphere containing 5% CO_2. To induce the differentiation process, ∼80% confluent C2C12 cells were cultured in DM consisting of DMEM with low glucose and pyruvate (Thermo Fisher Scientific, #31885-023) supplemented with 2% horse serum (Sigma-Aldrich, #H1270) and antibiotics (100 units of penicillin and 100 μg of streptomycin per ml, Sigma-Aldrich, #P4333).

In vivo analyses of Myolinc

The procedures for experimental animals were approved by the Experimental Animal Care and Use Committee at the Osaka University. The complex of siRNA and Invivofectamine (Thermo Fisher Scientific, #IVF3001) was prepared according to the manufacturer's protocol. Briefly, a siRNA was mixed with Invivofectamine to obtain a final concentration of 2.5 mg/ml siRNA duplex. After incubating at 50°C for 30 min, the lipid complex was diluted with 14 ml of PBS (pH = 7.4). The diluted lipid complex was centrifuged in Amicon Ultra-15 centrifugal device with Ultracel-50 membrane (Millipore) until the total volume reached 500 μl. The mice (7 to 8-week-old male C57BL/6) were injured by the injection of cardiotoxin to TA muscles using our established protocol (Ogawa et al., 2015). On Day 2 and Day 3 post injury, 30 μl of Invivofectamine–siScr complex and the same amount of the complex with siMyolinc.2 were injected into the right and left TA muscles, respectively. A total of 75 μg of siRNA was injected into each muscle. Five days after the cardiotoxin injection, the mice were sacrificed to extract muscles, which were snap-frozen in liquid nitrogen-cooled isopentane (Wako Pure Chemical Industries).

For the expression analysis of *Myolinc*, cryosections (14–16 μm) from siScr- or siMyolinc-treated muscle were collected in a tube including 375 μl TRIzol LS (Thermo Fisher Scientific, #10296-10). After the addition of 125 μl of PBS, the tube was incubated at room temperature for 5 min. After adding 100 μl chloroform, the sample was vortexed for 1 min, and left at room temperature for 5 min, and centrifuged at 12000× *g* at 4°C for 15 min. RNeasy Micro Kit (Qiagen, #74004) was used for further RNA purification according to the manufacturer's instruction. To quantify the efficiency of knockdown, siMyolinc-injected muscle was normalized to siScr-injected muscle of the same mouse.

To examine the effect of knockdown of *Myolinc*, transverse cryosections (6 μm) of the muscle were fixed with cooled-acetone for 10 min. In order to block endogenous mouse IgG, an M.O.M. kit (Vector Laboratories) was used. The reaction of primary antibodies (anti-eMyHC (Developmental Studies Hybridoma Bank, #F1.652); and anti-Myogenin, Clone F5D (Dako, #IR067)) was carried out at 4°C overnight and stained with the appropriate secondary antibody. The signals were recorded photographically using a BZ-X700 fluorescence microscope (Keyence). The quantitative analyses were performed using BZ-H3 system (Keyence).

Mononuclear cells from uninjured or CTX-injected limb muscles were prepared using 0.2% collagenase type II (Worthington Biochemical Corp.) as previously described (Uezumi et al., 2006). To prepare myogenic cells from CTX-injected muscle, Lympholyte (Cedarlane Laboratories) was used to remove debris according to manufacturer's instructions. The mononuclear cells were stained with FITC-conjugated anti-CD31, CD45, phycoerythrin-conjugated anti-Sca-1, and biotinylated-SM/C-2.6 (Fukada et al., 2004) antibodies. Cells were then incubated with streptavidin-labeled allophycocyanin (BD Biosciences) on ice for 30 min and resuspended in PBS containing 2% FCS and 2 μg/ml PI. SM/C-2.6^{+}Sca-1^{-}CD31^{-}CD45^{-} fractions were collected using a FACS Aria II^{TM} flow cytometer (BD Immunocytometry Systems). Data were collected using FACSDiva^{TM} software (BD Biosciences).

The following procedures were conducted in strict accordance with the institutional guidelines and were approved by the IACUC and Institutional Biosafety Committee of the University of Louisville (IACUC no. 16663). For the comparison of *Myolinc* expression between fast- and slow-twitch muscles, RNAs were isolated from extensor digitorum longus and soleus muscles, respectively, from 10-week-old C57BL/6 mice. For the comparison of *Myolinc* expression between wild-type (C57BL/6) and *mdx* mice, RNAs were extracted from tibialis anterior muscle from 12-week-old mice.

Microarray experiments and data analysis

Total RNA was prepared with TRIzol Reagent (Thermo Fisher Scientific, #15596026). The quality of extracted RNA was assessed by the Agilent Bioanalyzer (Agilent Technologies). GeneChip®Mouse Gene 1.0 ST Arrays (Affymetrix) were utilized according to the manufacturer's protocol and scanned. The CEL

files were analyzed through the updated version of noncoder web interface (http://noncoder.mpi-bn.mpg.de) (Gellert et al., 2013) using the pipeline set up for Gene Array Analyzer (GAA) web interface (http://gaa.mpi-bn.mpg.de) (Gellert et al., 2012). After the normalization by Robust Multi-array Average (RMA) (Irizarry et al., 2003) and the application of moderate *t*-statistics via the limma package (Ritchie et al., 2015), Transcript Cluster IDs that match that do not match to a gene or to multiple genes were discarded. Then, a standard deviation was calculated across samples. For a gene that matches to multiple Transcript Cluster IDs, the Transcript Cluster ID with the highest standard deviation was kept for further analysis.

For the screening of *Myolinc*, sample GeneChip Exon 1.0 ST arrays (exon arrays) provided by Affymetrix were utilized as in our previous studies (Gellert et al., 2009, 2013).

GO and KEGG pathway analyses were performed using DAVID (https://david.ncifcrf.gov/home.jsp) (Huang da et al., 2009) and LSKB software (http://www.lskb.w-fusionus.com) (World Fusion. inc.). Network analysis was conducted via NetworkAnalyst (http://www.networkanalyst.ca) (Xia et al., 2015).

Primary myoblasts

Primary myoblasts were isolated from hind limb muscles of 7-week-old C57BL/6 mice. Briefly, mice were euthanized and the hind limb muscles were isolated. The excess connective tissues and fat were cleaned in sterile PBS. Muscle tissue was then minced into a coarse slurry and enzymatically digested at 37°C for 1 h by adding 400 IU/ml collagenase II (Worthington). The digested slurry was spun, pelleted, and triturated multiple times, and then sequentially passed through a 70-µm and then 30-µm cell strainer (BD Falcon). The filtrate was spun at 1000× *g* and suspended in myoblast growth medium (MGM; Ham's F-10 medium with 20% FBS supplemented with 10 ng/ml of basic fibroblast growth factor). Cells were first re-fed after 3 days of initial plating followed by pre-plating for 15–30 min for the first few passages to select for a pure myoblast population. Pure myoblasts were then expanded through undergoing serial divisions.

For *Myolinc* and *TDP-43* gene silencing experiments in primary myoblasts, scrambled siRNA, *Myolinc* siRNA, or *TDP-43* siRNA oligonucleotides were electroporated into primary myoblasts (1500 V, 10 ms duration, 3 pulses) using the Neon transfection system following manufacturer's protocol (Invitrogen). After 24 h of transfection, the cells were incubated in DM (2% horse serum in DMEM) for 48 h to allow for myotube formation. Cells were then fixed with 3.7% formaldehyde in PBS for 10 min at room temperature and permeabilized with 0.3% Triton X-100 in PBS for 7 min. Cells were then blocked with 2% BSA in PBS and incubated with mouse anti-MyHC (MF-20) at room temperature for 2 h followed by goat anti-mouse Alexa Fluor 568 at room temperature for 1 h. Nuclei were visualized through counterstaining with DAPI for 3 min. Stained cells were imaged and analyzed using a fluorescent inverted microscope (Nikon Eclipse TE 2000-U), a digital camera (Digital Sight DS-Fi1), and Elements BR 3.00 software (Nikon).

Accession numbers

The sequences of *Myolinc* isoforms have been deposited in the GenBank: longer (MG243346) and shorter (MG243347) isoforms. The microarray and ChIP-seq data associated with this publication have been deposited in the Gene Expression Omnibus under the accession number: GSE69530.

Statistics

Data are presented as mean ± SEM unless otherwise indicated. To calculate a *P*-value, *F*-test was applied to determine an equality of two variances followed by two-tailed Student's *t*-test with the appropriate assumption for variances (either equal or unequal variances) via Microsoft Excel.

Supplementary material

Supplementary material is available at *Journal of Molecular Cell Biology* online.

Acknowledgements

We thank Wenjun Jin for his excellent technical assistance, Prof. Carlo Gaetano for his Bioruptor Plus machine, Dr Francesco Spallotta for help with confocal microscope, and Rie Koiwa for the image of proposed model of *Myolinc*. The authors acknowledge the platform for immortalization of human cells of the Myology Institute in Paris, as well as the London MRC Neuromuscular Research Centre and the Biomedical Research Centre of Great Ormond Street Hospital for Children in London.

Funding

This work was supported by the LOEWE Center for Cell and Gene Therapy (State of Hessen; to S.U. and S.D.), the DFG (SFB834 to S.U. and S.D., UC 67/2-1 to S.U.), the German Center for Cardiovascular Research (DZHK; to S.U. and S.D.), V.V. Cooke Foundation (Kentucky, USA; to S.U.), grant from the University of Louisville School of Medicine (to S.U.), University of Louisville 21st Century University Initiative on Big Data in Medicine (to S.U.), and the startup funding from the Mansbach Family, the Gheens Foundation, and other generous supporters at the University of Louisville (to S.U.).

Conflict of interest: none declared.

Author contributions: S.U. and G.M. conceptualized the study. S.U., P.G., D.J., and T.W. performed computational analysis. G.M., M.R.H., Y.P., and S.M.H. performed *in vitro* experiments. A.K., K.M., and V.M. intellectually contributed to the development of the project. C.D. performed microarray experiments. M.N., L.Z., and S.F. performed *in vivo* experiments. S.D. and A.K. provided general guidance. S.U. supervised the project. The manuscript was prepared by G.M., A.K., S.D., and S.U. with inputs from co-authors. All authors read and approved the final manuscript.

References

Anderson, D.M., Anderson, K.M., Chang, C.L., et al. (2015). A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell *160*, 595–606.

Baralle, M., Buratti, E., and Baralle, F.E. (2013). The role of TDP-43 in the pathogenesis of ALS and FTLD. Biochem. Soc. Trans. *41*, 1536–1540.

Bentzinger, C.F., Wang, Y.X., and Rudnicki, M.A. (2012). Building muscle: molecular regulation of myogenesis. Cold Spring Harb. Perspect. Biol. *4*, pii: a008342.

Bergstrom, D.A., Penn, B.H., Strand, A., et al. (2002). Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. Mol. Cell *9*, 587–600.

Biressi, S., Bjornson, C.R., Carlig, P.M., et al. (2013). Myf5 expression during fetal myogenesis defines the developmental progenitors of adult satellite cells. Dev. Biol. *379*, 195–207.

Blum, R., Vethantham, V., Bowman, C., et al. (2012). Genome-wide identification of enhancers in skeletal muscle: the role of MyoD1. Genes Dev. *26*, 2763–2779.

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

Buratti, E., Brindisi, A., Giombi, M., et al. (2005). TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through Its C-terminal tail. J. Biol. Chem. *280*, 37572–37584.

Butchart, L.C., Fox, A., Shavlakadze, T., et al. (2016). The long and short of non-coding RNAs during post-natal growth and differentiation of skeletal muscles: focus on lncRNA and miRNAs. Differentiation *92*, 237–248.

Cabili, M.N., Trapnell, C., Goff, L., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. *25*, 1915–1927.

Cesana, M., Cacchiarelli, D., Legnini, I., et al. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell *147*, 358–369.

Clark, M.B., Mercer, T.R., Bussotti, G., et al. (2015). Quantitative gene profiling of long noncoding RNAs with targeted RNA sequencing. Nat. Methods *12*, 339–342.

Darr, K.C., and Schultz, E. (1987). Exercise-induced satellite cell activation in growing and mature skeletal muscle. J. Appl. Physiol. (1985) *63*, 1816–1821.

Derrien, T., Johnson, R., Bussotti, G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. *22*, 1775–1789.

Flynn, R.A., and Chang, H.Y. (2014). Long noncoding RNAs in cell-fate programming and reprogramming. Cell Stem Cell *14*, 752–761.

Francetic, T., and Li, Q. (2011). Skeletal myogenesis and Myf5 activation. Transcription *2*, 109–114.

Freibaum, B.D., Chitta, R.K., High, A.A., et al. (2010). Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. J. Proteome Res. *9*, 1104–1120.

Fukada, S., Higuchi, S., Segawa, M., et al. (2004). Purification and cell-surface marker characterization of quiescent satellite cells from murine skeletal muscle by a novel monoclonal antibody. Exp. Cell Res. *296*, 245–255.

Gellert, P., Ponomareva, Y., Braun, T., et al. (2013). Noncoder: a web interface for exon array-based detection of long non-coding RNAs. Nucleic Acids Res. *41*, e20.

Gellert, P., Teranishi, M., Jenniches, K., et al. (2012). Gene Array Analyzer: alternative usage of gene arrays to study alternative splicing events. Nucleic Acids Res. *40*, 2414–2425.

Gellert, P., Uchida, S., and Braun, T. (2009). Exon Array Analyzer: a web interface for Affymetrix exon array analysis. Bioinformatics *25*, 3323–3324.

Goff, L.A., and Rinn, J.L. (2015). Linking RNA biology to lncRNAs. Genome Res. *25*, 1456–1465.

Gong, C., Li, Z., Ramanujan, K., et al. (2015). A long non-coding RNA, LncMyoD, regulates skeletal muscle differentiation by blocking IMP2-mediated mRNA translation. Dev. Cell *34*, 181–191.

Heinz, S., Benner, C., Spann, N., et al. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Irizarry, R.A., Hobbs, B., Collin, F., et al. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249–264.

Jin, C.F., Li, Y., Ding, X.B., et al. (2017). lnc133b, a novel, long non-coding RNA, regulates bovine skeletal muscle satellite cell proliferation and differentiation by mediating miR-133b. Gene *630*, 35–43.

Kalhovde, J.M., Jerkovic, R., Sefland, I., et al. (2005). 'Fast' and 'slow' muscle fibres in hindlimb muscles of adult rats regenerate from intrinsically different satellite cells. J. Physiol. *562*, 847–857.

King, I.N., Yartseva, V., Salas, D., et al. (2014). The RNA-binding protein TDP-43 selectively disrupts microRNA-1/206 incorporation into the RNA-induced silencing complex. J. Biol. Chem. *289*, 14263–14271.

Lalmansingh, A.S., Urekar, C.J., and Reddi, P.P. (2011). TDP-43 is a transcriptional repressor: the testis-specific mouse acrv1 gene is a TDP-43 target in vivo. J. Biol. Chem. *286*, 10970–10982.

Lander, E.S., Linton, L.M., Birren, B., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Lee, E.B., Lee, V.M., and Trojanowski, J.Q. (2011). Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. Nat. Rev. Neurosci. *13*, 38–50.

Legnini, I., Morlando, M., Mangiavacchi, A., et al. (2014). A feedforward regulatory loop between HuR and the long noncoding RNA linc-MD1 controls early phases of myogenesis. Mol. Cell *53*, 506–514.

Li, P., Ruan, X., Yang, L., et al. (2015). A liver-enriched long non-coding RNA, lncLSTR, regulates systemic lipid metabolism in mice. Cell Metab. *21*, 455–467.

Liang, T., Zhou, B., Shi, L., et al. (2017). lncRNA AK017368 promotes proliferation and suppresses differentiation of myoblasts in skeletal muscle development by attenuating the function of miR-30c. FASEB J. *32*, 377–389.

Liu, X., Li, D., Zhang, W., et al. (2012). Long non-coding RNA gadd7 interacts with TDP-43 and regulates Cdk6 mRNA decay. EMBO J. *31*, 4415–4427.

Matsumoto, A., Pasut, A., Matsumoto, M., et al. (2017). mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature *541*, 228–232.

Millay, D.P., O'Rourke, J.R., Sutherland, L.B., et al. (2013). Myomaker is a membrane activator of myoblast fusion and muscle formation. Nature *499*, 301–305.

Molkentin, J.D., Black, B.L., Martin, J.F., et al. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. Cell *83*, 1125–1136.

Molyneaux, B.J., Goff, L.A., Brettler, A.C., et al. (2015). DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. Neuron *85*, 275–288.

Mueller, A.C., Cichewicz, M.A., Dey, B.K., et al. (2015). MUNC, a long noncoding RNA that facilitates the function of MyoD in skeletal myogenesis. Mol. Cell. Biol. *35*, 498–513.

Nagano, T., Morikubo, S., and Sato, M. (2004). Filamin A and FILIP (Filamin A-Interacting Protein) regulate cell polarity and motility in neocortical subventricular and intermediate zones during radial migration. J. Neurosci. *24*, 9648–9657.

Nagano, T., Yoneda, T., Hatanaka, Y., et al. (2002). Filamin A-interacting protein (FILIP) regulates cortical cell migration out of the ventricular zone. Nat. Cell Biol. *4*, 495–501.

Nelson, B.R., Makarewich, C.A., Anderson, D.M., et al. (2016). A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science *351*, 271–275.

Neppl, R.L., Wu, C.L., and Walsh, K. (2017). lncRNA Chronos is an aging-induced inhibitor of muscle hypertrophy. J Cell Biol. *216*, 3497–3507.

Ogawa, R., Ma, Y., Yamaguchi, M., et al. (2015). Doublecortin marks a new population of transiently amplifying muscle progenitor cells and is required for myofiber maturation during skeletal muscle regeneration. Development *142*, 51–61.

Olive, M., Janue, A., Moreno, D., et al. (2009). TAR DNA-Binding protein 43 accumulation in protein aggregate myopathies. J. Neuropathol. Exp. Neurol. *68*, 262–273.

Ou, S.H., Wu, F., Harrich, D., et al. (1995). Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. J. Virol. *69*, 3584–3596.

Peng, X., Sun, K., Zhou, J., et al. (2017). Bioinformatics for novel long intergenic noncoding RNA (lincRNA) identification in skeletal muscle cells. Methods Mol. Biol. *1556*, 355–362.

Rawls, A., Valdez, M.R., Zhang, W., et al. (1998). Overlapping functions of the myogenic bHLH genes MRF4 and MyoD revealed in double mutant mice. Development *125*, 2349–2358.

Relaix, F., and Zammit, P.S. (2012). Satellite cells are essential for skeletal muscle regeneration: the cell on the edge returns centre stage. Development *139*, 2845–2856.

Rinn, J.L. (2014). lncRNAs: linking RNA to chromatin. Cold Spring Harb. Perspect Biol. *6*, pii: a018614.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. *81*, 145–166.

Ritchie, M.E., Phipson, B., Wu, D., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Simionescu-Bankston, A., and Kumar, A. (2016). Noncoding RNAs in the regulation of skeletal muscle biology in health and disease. J. Mol. Med. *94*, 853–866.

Uchida, S., and Bolli, R. (2017). Short and long noncoding RNAs regulate the epigenetic status of cells. Antioxid. Redox Signal. doi:10.1089/ars.2017.7262.

Uchida, S., and Dimmeler, S. (2015). Long noncoding RNAs in cardiovascular diseases. Circ. Res. *116*, 737–750.

Uchida, S., Gellert, P., and Braun, T. (2012). Deeply dissecting stemness: making sense to non-coding RNAs in stem cells. Stem Cell Res. *8*, 78–86.

Uezumi, A., Ojima, K., Fukada, S., et al. (2006). Functional heterogeneity of side population cells in skeletal muscle. Biochem. Biophys. Res. Commun. *341*, 864–873.

Ustanina, S., Carvajal, J., Rigby, P., et al. (2007). The myogenic factor Myf5 supports efficient skeletal muscle regeneration by enabling transient myoblast amplification. Stem Cells *25*, 2006–2016.

Wang, L., Zhao, Y., Bao, X., et al. (2015). LncRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. Cell Res. *25*, 335–350.

Werber, M., Wittler, L., Timmermann, B., et al. (2014). The tissue-specific transcriptomic landscape of the mid-gestational mouse embryo. Development *141*, 2325–2330.

Xia, J., Gill, E.E., and Hancock, R.E. (2015). NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nat. Protoc. *10*, 823–844.

Yang, Z., MacQuarrie, K.L., Analau, E., et al. (2009). MyoD and E-protein heterodimers switch rhabdomyosarcoma cells from an arrested myoblast phase to a differentiated state. Genes Dev. *23*, 694–707.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite cells and the muscle stem cell niche. Physiol. Rev. *93*, 23–67.

Yu, X., Zhang, Y., Li, T., et al. (2017). Long non-coding RNA Linc-RAM enhances myogenic differentiation by interacting with MyoD. Nat. Commun. *8*, 14016.

Zhou, L., Sun, K., Zhao, Y., et al. (2015). Linc-YY1 promotes myogenic differentiation and muscle regeneration through an interaction with the transcription factor YY1. Nat. Commun. *6*, 10026.

Zhou, J., Zhang, S., Wang, H., et al. (2017). LncFunNet: an integrated computational framework for identification of functional long noncoding RNAs in mouse skeletal muscle cells. Nucleic Acids Res. *45*, e108.

Zhu, M., Liu, J., Xiao, J., et al. (2017). Lnc-mg is a long non-coding RNA that promotes myogenesis. Nat. Commun. *8*, 14718.