Long noncoding RNA *SMUL* suppresses *SMURF2* production-mediated muscle atrophy via nonsense-mediated mRNA decay

Bolin Cai,[1,4,5] Zhenhui Li,[1,3,4,5] Manting Ma,[1,4,5] Jing Zhang,[1,4] Shaofen Kong,[1,4] Bahareldin Ali Abdalla,[1,4] Haiping Xu,[1,4] Endashaw Jebessa,[1,4] Xiquan Zhang,[1,4] Raman Akinyanju Lawal,[2] and Qinghua Nie[1,4]

[1]College of Animal Science, Lingnan Guangdong Laboratory of Modern Agriculture & State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou 510642, Guangdong, China; [2]The Jackson Laboratory, 600 Main Street, Bar Harbor, ME, USA; [3]Laboratory of Neurobiology and Behavior, The Rockefeller University, New York, NY 10065, USA; [4]Guangdong Provincial Key Lab of Agro-Animal Genomics and Molecular Breeding, and Key Laboratory of Chicken Genetics, Breeding and Reproduction, Ministry of Agriculture, Guangzhou 510642, Guangdong, China

As the world population grows, muscle atrophy leading to muscle wasting could become a bigger risk. Long noncoding RNAs (lncRNAs) are known to play important roles in muscle growth and muscle atrophy. Meanwhile, it has recently come to light that many putative small open reading frames (sORFs) are hidden in lncRNAs; however, their translational capabilities and functions remain unclear. In this study, we uncovered 104 myogenic-associated lncRNAs translated, in at least a small peptide, by integrated transcriptome and proteomic analyses. Furthermore, an upstream ORF (uORF) regulatory network was constructed, and a novel muscle atrophy-associated lncRNA named *SMUL* (Smad ubiquitin regulatory factor 2 [*SMURF2*] upstream lncRNA) was identified. *SMUL* was highly expressed in skeletal muscle, and its expression level was downregulated during myoblast differentiation. *SMUL* promoted myoblast proliferation and suppressed differentiation in vitro. In vivo, *SMUL* induced skeletal muscle atrophy and promoted a switch from slow-twitch to fast-twitch fibers. In the meantime, translation of the *SMUL* sORF disrupted the stability of *SMURF2* mRNA. Mechanistically, *SMUL* restrained *SMURF2* production via nonsense-mediated mRNA decay (NMD), participating in the regulation of the transforming growth factor β (TGF-β)/SMAD pathway and further regulating myogenesis and muscle atrophy. Taken together, these results suggest that *SMUL* could be a novel therapeutic target for muscle atrophy.

INTRODUCTION

As the world population expands, muscle atrophy has become a major health challenge. Specifically, muscle atrophy conditions can present themselves during both aging and prolonged periods of muscle inactivity.[1,2] Prevention of this condition could provide a significant clinical benefit.

Muscle growth is a highly ordered process in which myoblasts withdraw from the cell cycle, express muscle-specific genes, and prevent the expression of other cell- or tissue-specific genes, including myoblast proliferation and differentiation as well as myotube formation and maturity, which are controlled by a series of myogenic regulatory factors.[3–5] After birth, the number of muscle fibers in animals is basically fixed, and their skeletal muscle development is mainly regulated by the size and composition of muscle fibers. Protein synthesis and catabolism are the main factors affecting skeletal muscle mass.[6] When protein synthesis exceeds its degradation, muscle hypertrophy occurs and the cross-sectional area of muscle fiber increases. Furthermore, excessive protein degradation induces muscle atrophy, which leads to a decrease in muscle fiber cross-sectional area and muscle strength. Skeletal muscle is heterogeneous,[7] and changes in its mass can often cause the transformation of muscle fiber types. In response to environmental demands, skeletal muscle can remodel by activating signaling pathways to reprogram gene expression to sustain muscle performance.[7] Recently, multiple studies have found that long noncoding RNAs (lncRNAs) can determine epigenetic or transcriptional regulation in skeletal muscle development.[8–12]

Increasing high-throughput studies have revealed the presence of pervasive transcription across 70%–90% of the human genome. Tens of thousands of genomic loci produce transcripts, but more than 98% of these, known as noncoding RNAs (ncRNAs), are not translated to protein.[13] Noncoding transcripts can generally be divided into small noncoding RNAs and lncRNAs on the basis of their size. Compared with small noncoding RNAs, lncRNAs are a novel class of mRNA-like transcripts, can vary in length from 200 nt to 100 kb, and make up the largest proportion of mammalian

noncoding transcriptome.[14,15] The translation capacity of lncRNAs has long been overlooked with the traditional approaches to classify the potential protein-coding capacity of RNA transcripts, which have been based on canonical open reading frames (ORFs), codon conservation, and similarity to known protein domains.[16,17] The cutoff of 100 aa was an artificial standard for annotation of ORFs as protein-coding transcripts or not. Therefore, the distinction between lncRNAs and mRNAs is not always clear.[16,18] Increasing studies revealed that lncRNAs were often associated with ribosomes, providing key clues to unanticipated translation potential of lncRNAs.[19,20] Previous studies have demonstrated that lncRNAs can give rise to functional peptides. lncRNA-Six1, located upstream of protein-coding gene *Six1*, can produce a 7.26-kDa peptide, affecting myoblast proliferation and differentiation.[21] Myoregulin (MLN) is a conserved 46-aa peptide encoded by a skeletal muscle-specific lncRNA, and co-localized with sarcoplasmic/endoplasmic reticulum Ca^{2+} adenosine triphosphatase (SERCA). MLN decreases Ca^{2+} release in skeletal muscle and affects muscle performance.[22] To date, the function that translation of lncRNAs contributes to muscle atrophy is unclear.

Smad ubiquitin regulatory factor 2 (*SMURF2*) is a homology to E6 carboxyl terminus (HECT) domain-containing E3 ubiquitin ligase that specifies substrates for ubiquitination and degradation by the proteasome.[23,24] *SMURF2* regulates key biological processes during development and homeostasis via controlling the transforming growth factor β (*TGF-β*)/mothers against decapentaplegic (*SMADs*) signaling pathway.[25–29] However, the precise mechanisms regulating *SMURF2* abundance remain largely unknown.

Recently, combined with proteomics analysis, some hidden proteins encoded by lncRNA have been discovered.[30–32] To systematically identify the lncRNA with potential translation ability involved in myogenesis, we characterized the lncRNA and mRNA expression profiles as well as peptidomic profiles during chicken primary myoblast (CPM) proliferation and differentiation. Based on these results, a potential translated lncRNA, LNC_004915, located upstream of the *SMURF2* gene, was identified and named *SMUL* (*SMURF2* upstream lncRNA). *SMUL* is highly expressed in skeletal muscle, and its expression decreases with myoblast differentiation. Gain- and loss-of-function analysis revealed that *SMUL* promoted myoblast proliferation and inhibited myogenic differentiation. *In vivo*, *SMUL* induced skeletal muscle atrophy and activated the fast-twitch fiber phenotype. Mechanistically, *SMUL* induces nonsense-mediated mRNA decay (NMD) of *SMURF2* through its translation process and activates the TGF-β/SMAD pathway. Altogether, our studies uncover a functional lncRNA that modulates skeletal muscle development, and they provide a novel therapeutic target for treating muscle atrophy.

RESULTS

Identification and characterization of small ORF (sORF)-encoded lncRNAs by integrated transcriptome and proteomic analysis

CPMs were cultured in differentiation medium to induce myoblast differentiation *in vitro*. To systematically identify the myogenesis-associated lncRNAs, RNA sequencing (RNA-seq) was performed during myoblast proliferation (CPMs cultured in growth medium [GM]) and the third day of differentiation (CPMs cultured in differentiation medium [DM]) (Figure 1A). In total, 3,106 differentially expressed genes (DEGs) and 72 differentially expressed lncRNAs (DE-lncRNAs) were identified during myoblast proliferation and differentiation (Figures 1B and 1C; Tables S1 and S2). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis found that the DEGs were enriched in RNA transport, ribosome biogenesis in eukaryotes, biosynthesis of amino acids, and the cell cycle (Figures S1A and S1B).

To identify the protein-coding potential of lncRNA sORF, we performed an integrated analysis using RNA-seq and proteomics data during CPM proliferation and differentiation. Interestingly, we found 104 myogenic lncRNAs that translated at least one small peptide, indicating that translated sORFs are hidden in myogenic lncRNAs (Figure 1D; Table S3). The length distribution of those small peptides is relatively uniform, with most in the range of 21–40 aa (Figure 1E). Furthermore, more than 80% of those are located on exons (58.04%) and are intergenic (27.27%), implying that they may perform their biological functions by acting as regulatory elements (Figure 1F).

Cis-acting is one of the main ways that lncRNA participates in epigenetic regulation.[21,33,34] Upstream ORF (uORF), as an important *cis*-element, appears in nearly half of mammalian transcripts and extensively regulates downstream gene expression.[35,36] To study the regulatory function of translated lncRNA, we constructed a uORF (which is transcribed by translated lncRNA) regulatory network (Figure 1G). A total of 49 translated lncRNA-transcribed ORFs were found to locate on the same chain and within 1.5 kb upstream of the main protein-coding ORF (mORF), which is predicted to function as a uORF (Figure 1G). *SMURF2* is well known to modulate the TGF-β/SMAD pathway, which widely regulates animal development.[25–29] Therefore, the LNC_004915-*SMURF2* pair served as a candidate for further study.

lncRNA *SMUL* is reduced in myoblast differentiation and has translation ability

To evaluate the full-length sequence of *SMUL* (LNC_004915), the 5′ and 3′ ends of this lncRNA were determined by a rapid amplification of cDNA ends (RACE) system (Figure 2A). The Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI) showed that *SMUL* was 1,061 nt long, located at chromosome 18 from position 6,970,951 to 6,967,249, relatively conserved in *Numida meleagris*, *Meleagris gallopavo*, and *Coturnix japonica*, and that its 3′ end was separated from the *SMURF2* coding sequence (CDS) by 1,241 bp (Figure 2B; Figures S2A and S2B). *SMUL* was highly expressed in polyadenylated RNA, demonstrating that *SMUL* is a polyadenylated lncRNA, which is similar to mRNA, and has a poly(A) tail (Figure S2C). The expression level of *SMUL* was significantly downregulated during myogenic differentiation, whereas its target gene (*SMURF2*) was gradually increased with the differentiation of myoblasts (Figures 2C and 2D). In the meantime, *SMUL* and *SMURF2* were highly expressed in breast and leg muscle
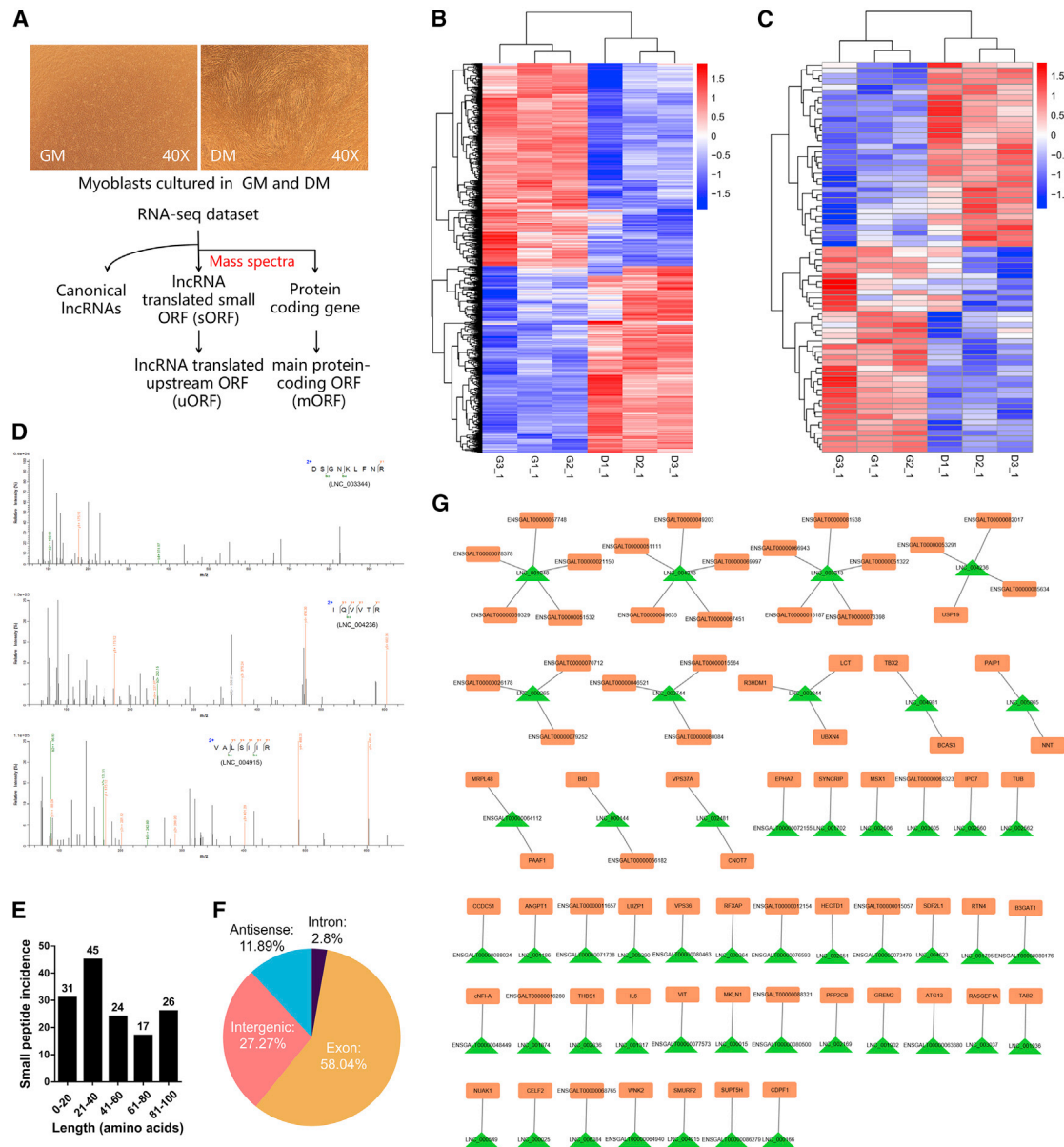
Figure 1. Translated lncRNAs are hidden in myogenesis

(A) Overview of experimental strategy for transcriptome (RNA-seq) and peptidome. The upper panel shows microscopic images of CPMs during proliferation (CPMs cultured in growth medium [GM]) and the third day of differentiation (CPMs cultured in differentiation medium [DM]). The lower panel is the discovery pipeline for translated lncRNAs. (B and C) Heatmaps of differentially expressed mRNAs (B) and lncRNAs (C) during myoblast proliferation and differentiation, with rows representing mRNAs or lncRNAs and columns representing different myoblast stages. (D) Mass spectrometry results of small peptide encoded by three translated lncRNAs (LNC_003344, LNC_004236, and LNC_004915) as examples. (E) Amino acid length distribution of lncRNA translated peptides detected by mass spectra. (F) Incidence of lncRNA translated peptides detected by mass spectra in each category within RefSeq mRNAs. (G) uORF regulatory network of translated lncRNAs and their potential target genes. lncRNAs containing peptides identified by mass spectrometry were used for the regulatory network analysis.

(Figures 2E and 2F), implying that they may play an important role in skeletal muscle development. Cell-fractionation assays demonstrated that *SMUL* is mainly present in the nuclei of CPMs (Figures 2G and 2H). A similar result was also confirmed by *in situ* RNA hybridization (Figure 2I). Interestingly, our peptidome data revealed

that *SMUL* encoded a small peptide, suggesting that *SMUL* has translation potential.

To verify this prediction, a series of constructs was generated in which a GFP mutant (GFPmut) ORF (in which the start codon ATGGTG
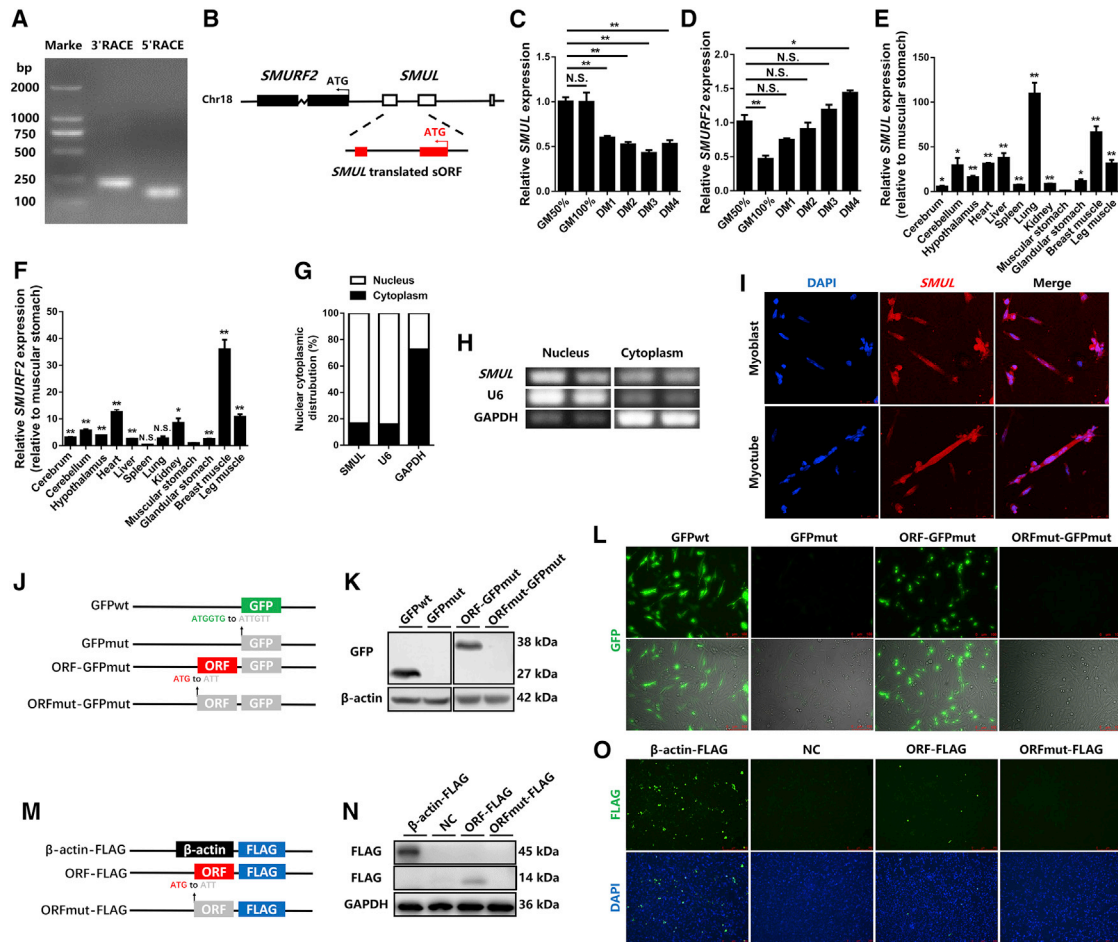
Figure 2. lncRNA *SMUL* is downregulated in myogenic differentiation and has translation capacity
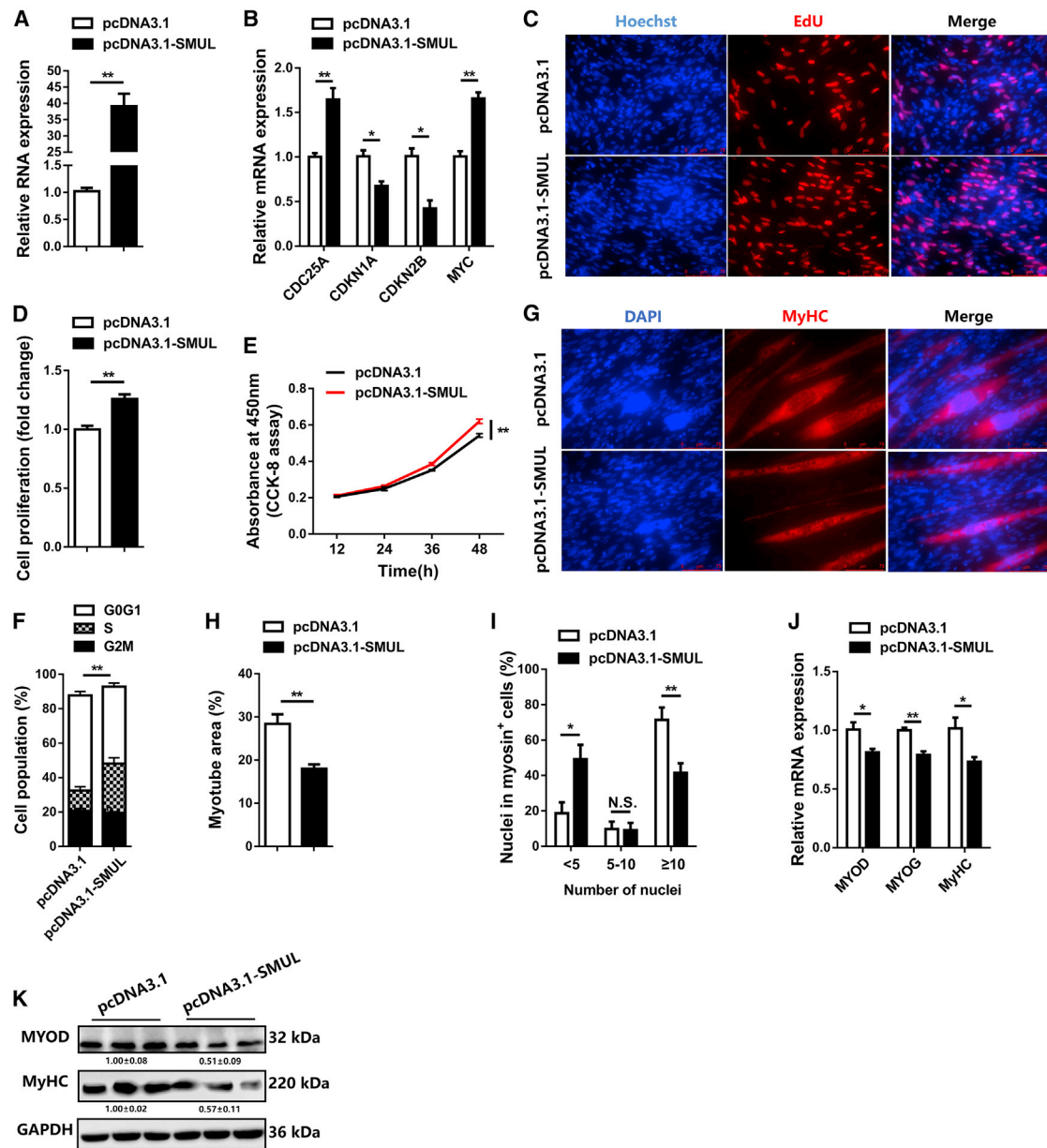
(A) Results of *SMUL* 5′ RACE and 3′ RACE. (B) Schematic image of the locus for *SMUL* (white), *SMUL* translated sORF (red), and *SMURF2* (black). Thin arrows represent the direction of transcription. (C and D) Relative *SMUL* (C) and *SMURF2* (D) expression during CPM differentiation (n = 4). (E and F) Tissue expression profiles of *SMUL* (E) and *SMURF2* (F). The horizontal axis and vertical axis indicate different tissues and their relative expression values, respectively (n = 4). (G and H) The distribution of *SMUL* in the cytoplasm and nuclei of CPMs was determined by quantitative real-time PCR (G) and semi-quantitative real-time PCR (H). Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and *U6* serve as cytoplasmic and nuclear localization controls, respectively. (I) RNA *in situ* hybridization of *SMUL* in chicken primary myoblasts. Special FISH probes against *SMUL* were modified by Cy3 (red). The nucleus was stained by 4′,6-diamidino-2-phenylindole (DAPI) (blue). (J) Diagram of the *GFP* fusion constructs used for transfection. The initiation codon ATGGTG of the *GFP* (GFPwt) gene is mutated to ATTGTT (GFPmut). The initiation codon ATG of the *SMUL* sORF is mutated to ATT. (K and L) Western blotting with anti-GFP (K) and GFP fluorescence (L) were performed following 48 h after transfection of the indicated plasmids. (M) Diagram of the *FLAG* fusion constructs used for transfection. The initiation codon ATG of the *SMUL* sORF is mutated to ATT. (N and O) Western blotting with anti-FLAG (N) and immunostained using anti-FLAG (O) were performed following 48 h of transfection of the indicated plasmids. Data are presented as mean ± SEM. Statistical significance of differences between means was assessed using an independent sample t test. *p < 0.05, **p < 0.01.

was mutated to ATTGTT) was fused to the C terminus of the ORF and the ORF mutant (ORFmut) in the full-length *SMUL* transcript (Figure 2J). Western blotting analysis using anti-GFP antibodies showed that *SMUL*-GFP fusion protein exhibited the predicted relative molecular masses in *SMUL* ORF-GFPmut-transfected cells, but not in *SMUL* ORFmut-GFPmut-transfected cells (Figure 2K). Substantial expression of the *SMUL*-GFP fusion protein was observed in *SMUL* ORF-GFPmut-transfected cells, while mutation of the *SMUL* ORF start codon (in which the start codon ATG was mutated to ATT) abolished the expression of the *SMUL*-GFP fusion protein (Figure 2L). Meanwhile, we further cloned a FLAG epitope tag in-

frame with the C terminus of the ORF and the ORFmut within the full-length *SMUL* transcript (Figure 2M). These results are consistent with those found using the *SMUL*-GFP fusion protein (Figures 2N and 2O), indicating that *SMUL*, which is annotated as an lncRNA, has translation capacity.

SMUL negatively regulates muscle development and induces muscle atrophy

In order to determine the function of lncRNA *SMUL* in myogenesis, the overexpression vector of *SMUL* at full length was constructed and transfected into CPMs (Figure 3A). Overexpression of *SMUL*

Figure 3. *SMUL* facilitates myoblast proliferation but suppresses myoblast differentiation *in vitro*

(A) Relative *SMUL* expression after introducing *SMUL* into CPMs (n = 4). (B) Relative mRNA levels of several cell cycle genes after overexpression of *SMUL* (n = 4). (C) Proliferation of transfected CPMs was assessed by 5-ethynyl-2′-deoxyuridine (EdU) incorporation (n = 3). (D) Proliferation rate of myoblasts with *SMUL* overexpression (n = 3). (E) CCK-8 assays were performed in CPMs with *SMUL* overexpression (n = 6). (F) Cell cycle analysis of myoblasts at 48 h after overexpression of *SMUL* (n = 4). (G–I) MyHC immunostaining (G) (n = 3), myotube area (%) (H) (n = 3), and myoblast fusion index (I) (n = 3) of CPMs transduced with *SMUL* overexpression. Cells were differentiated for 72 h after transfection. The nuclei were visualized with DAPI. (J and K) Relative mRNA (J) (n = 4) and protein (K) (n = 3) expression levels of myoblast differentiation marker genes from pcDNA3.1-SMUL-transfected CPMs. The numbers shown below the bands were folds of band intensities relative to control. Band intensities were quantified by ImageJ and normalized to *GAPDH*. Data are expressed as a fold change relative to the control. Results are shown as mean ± SEM. Statistical significances of differences between means were assessed using an independent sample t test. *p < 0.05, **p < 0.01.

increased the expression level of cell cycle-promoting genes, including *CDC25A* and *MYC*, while it decreased the expression level of cell cycle-inhibiting genes such as *CDKN1A* and *CDKN2B* (Figure 3B). Using 5-ethynyl-2′-deoxyuridine (EdU) and Cell Counting Kit-8

(CCK-8) assays, we found that *SMUL* promoted myoblast proliferation (Figures 3C–3E). Furthermore, *SMUL* overexpression reduced the number of cells that progressed to G_0/G_1, thus increasing the number of S phase cells (Figure 3F). Meanwhile, the opposite result
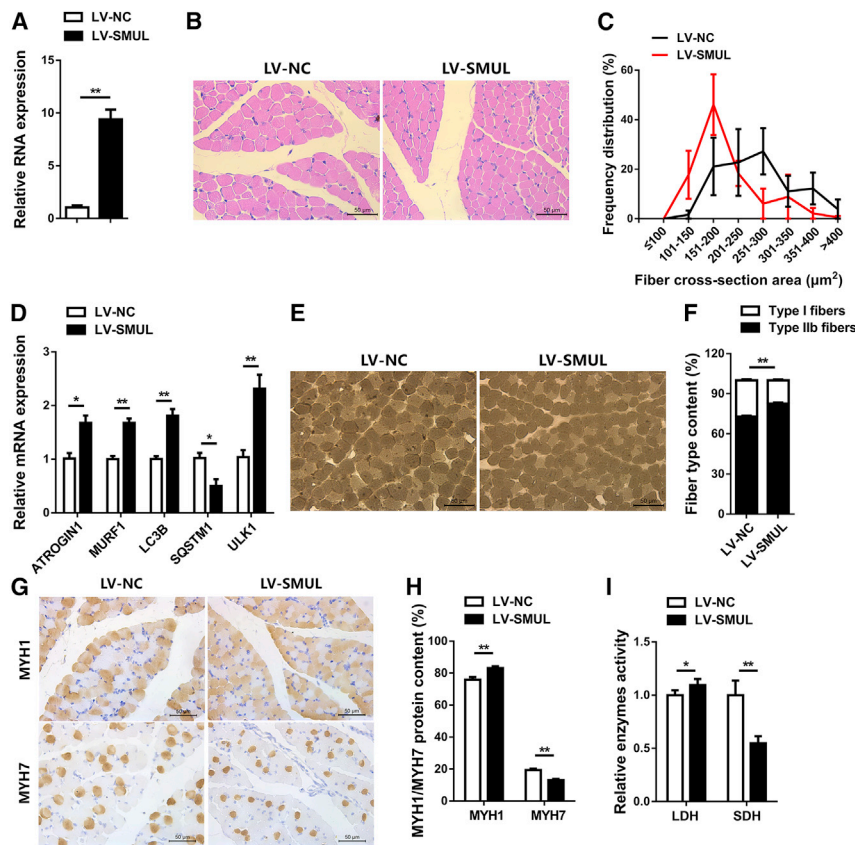
Figure 4. SMUL induces skeletal muscle atrophy and activates fast-twitch muscle phenotype in vivo

(A) Relative *SMUL* expression in gastrocnemius muscle after infection with *SMUL*-expressing lentivirus (LV-SMUL) or negative control (LV-NC) (n = 4). (B and C) H&E staining (B) (n = 4) and frequency distribution of fiber cross-section area (CSA) (C) (n = 3) of transverse sections of gastrocnemius muscle infected with the indicated lentivirus. (D) Relative mRNA expression of the atrophy and autophagy-related genes in gastrocnemius muscle after infection with the indicated lentivirus (n = 4). (E and F) ATPase staining (E) (n = 4) and fiber type content (F) (n = 4) of gastrocnemius muscle injected with the listed lentivirus. (G and H) Immunohistochemistry analysis of MYH1/MYH7 (G) (n = 4) and MYH1/MYH7 protein content (H) (n = 4) of gastrocnemius muscle injected with the listed lentivirus. (I) Relative enzymes activity of lactate dehydrogenase (LDH) and succinate dehydrogenase (SDH) in gastrocnemius muscle infected with the indicated lentivirus (n = 4). Data are presented as means ± SEM, and the statistical significance of differences between means was assessed using an independent sample t test. *p < 0.05, **p < 0.01.

was observed by *SMUL* lncRNA interference (Figures S3A–S3F). To further investigate the potential function of *SMUL* in myoblast differentiation, immunofluorescence staining was performed. Immunofluorescence staining showed that overexpression of *SMUL* suppressed myotube formation and myoblast fusion (Figures 3G–3I). Moreover, the expression levels of myoblast differentiation marker genes, including *MYOD*, *MYOG*, and *MyHC*, were significantly downregulated with *SMUL* overexpression (Figures 3J and 3K). In contrast, *SMUL* interference facilitated myoblast differentiation (Figures S3G–S3K).

To test whether *SMUL* regulates skeletal muscle development *in vivo*, 1-day-old chicks were injected with lentiviral-mediated *SMUL* overexpression (LV-SMUL) or lentiviral-mediated *SMUL* knockdown (LV-shSMUL) (Figure 4A; Figure S3K). Overexpression of *SMUL* not only elevated the proportion of small myofibers (<200 μm²), but it also upregulated the expression level of muscle atrophy marker genes (*ATROGIN1* and *MURF1*) and activated autophagy (Figures 4B-4D). However, on the contrary, knockdown of *SMUL* promoted muscle fiber hypertrophy (Figures S3M–S3O). Muscle fiber type is closely related to the occurrence of muscle atrophy. Compared with type I (slow-twitch; oxidative) muscle fibers, type II (fast-twitch; glycolytic) muscle fibers are more susceptible to atrophy.[37,38] Numerous studies have found that muscle atrophy is often accompa-

nied by the skeletal muscle transition from slow-twitch to fast-twitch fibers.[39–42] Here, overexpression of *SMUL* not only increased the percentage of type II (fast-twitch; glycolytic) muscle fibers, it also reduced the appearance of type I (slow-twitch; oxidative) muscle fibers (Figures 4E and 4F). Immunohistochemical results also showed that *SMUL* promoted MYH1/fast protein level and suppressed the expression level of MYH7/slow protein (Figures 4G and 4H). Moreover, the activity of lactate dehydrogenase (LDH) was enhanced, while the activity of succinate dehydrogenase (SDH) was suppressed with *SMUL* overexpression, suggesting that *SMUL* promoted the glycolytic capacity of skeletal muscle and induced the fast-twitch muscle phenotype (Figure 4I). In contrast, knockdown of *SMUL* induced fast-twitch muscle fiber transformation to slow-twitch muscle fibers and suppressed muscle atrophy (Figures S3P–S3T).

SMUL sORF disrupted the stability of SMURF2 mRNA via NMD

To investigate the molecular mechanism of *SMUL*, we detected the expression level of its target gene (*SMURF2*). Both *in vitro* and *in vivo*, overexpression of *SMUL* decreased the mRNA and protein levels of *SMURF2*, while *SMURF2* mRNA and protein were upregulated after *SMUL* knockdown (Figures 5A–5D). Expression vectors of wild-type (WT), ATG mutant (in which the start codon ATG was mutated to ATT), TGA mutant (in which the termination codon TGA was mutated to TGG), and both ATG and TGA mutant (in which both the start codon ATG was mutated to ATT and the termination codon TGA was mutated to TGG) of *SMUL* sORF were constructed to explore whether *SMUL* works through its sORF (Figures 5E and 5F). Importantly, only wild-type
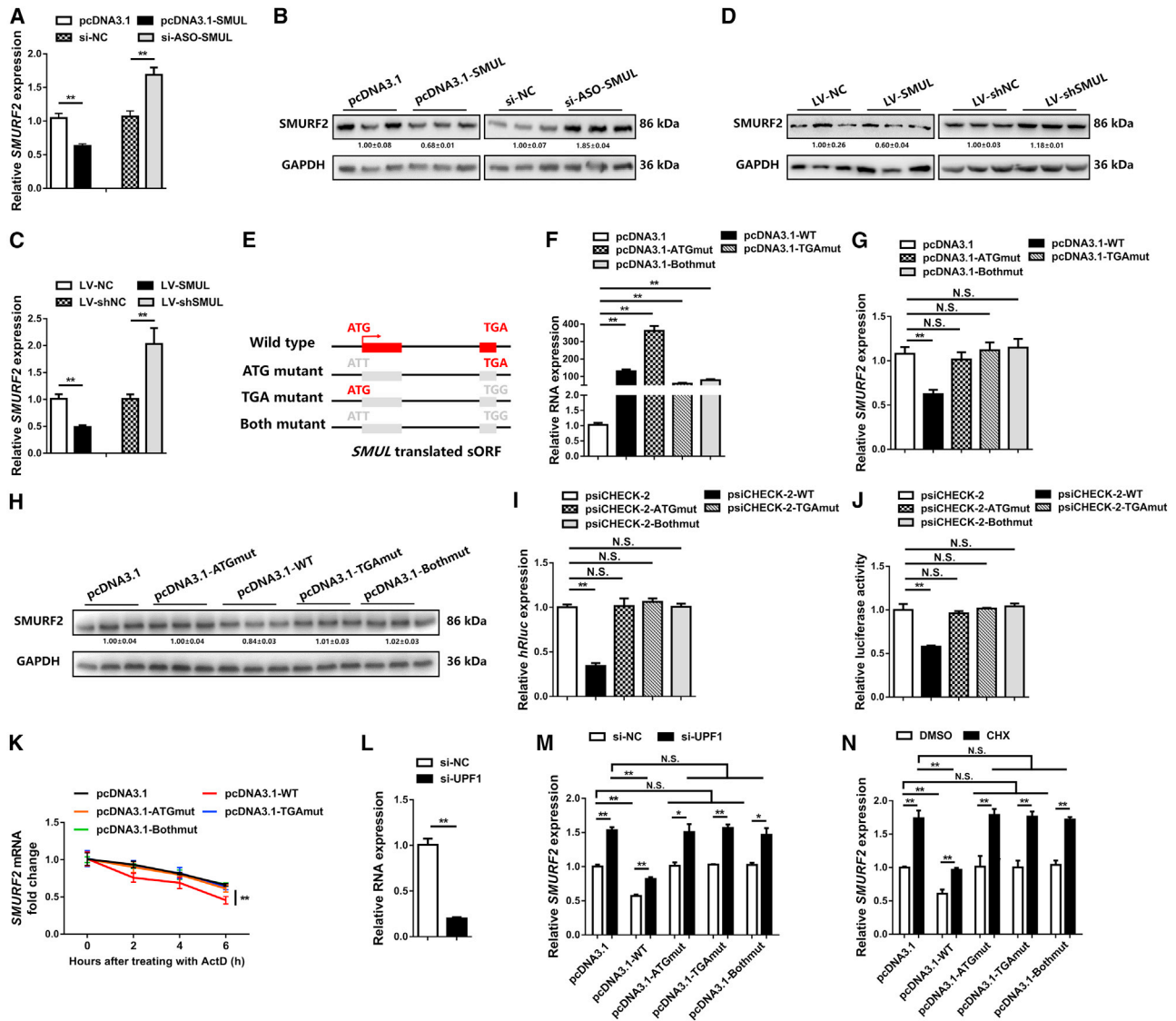
Figure 5. *SMUL* sORF disrupts *SMURF2* mRNA stability through nonsense-mediated mRNA decay (NMD) mechanism

(A–D) Relative mRNA (A and C) (n = 4) and protein (B and D) (n = 3) expression levels of *SMURF2* with *SMUL* overexpression or knockdown *in vitro* (A and B) and *in vivo* (C and D). The numbers shown below the bands were folds of band intensities relative to control. Band intensities were quantified by ImageJ and normalized to *GAPDH*. Data are expressed as a fold change relative to the control. (E) Diagram of the *SMUL* sORF wild-type and mutant constructs used for transfection. For the ATG mutant, the initiation codon ATG of the *SMUL* sORF was mutated to ATT. For the TGA mutant, the termination codon TGA of the *SMUL* sORF was mutated to TGG. For both the ATG and TGA mutant, the initiation codon ATG and termination codon TGA of the *SMUL* sORF was mutated to ATT and TGG. (F–H) Relative expression levels of *SMUL* sORF (F) (n = 4) and *SMURF2* (G and H) (n = 4 and n = 3) induced by the listed nucleic acids in CPMs. The numbers shown below the bands were folds of band intensities relative to control. Band intensities were quantified by ImageJ and normalized to *GAPDH*. Data are expressed as a fold change relative to the control. (I and J) Relative *hRlu* expression and luciferase activity after wild-type and mutant *SMUL* sORFs are inserted into the 5′ UTR of the *hRluc* gene (n = 4 and n = 5). (K) *SMURF2* stability assay after transfection with the indicated plasmids (n = 4). (L) Relative expression levels of *UPF1* with *UPF1* interference (n = 4). (M and N) Relative *SMURF2* expression following NMD pathway inhibition via siRNA-mediated depletion of the regulator *UPF1* (M) (n = 4) or cycloheximide (CHX)-mediated inhibition of protein synthesis (N) (n = 4) induced by the listed nucleic acids in CPMs. Results are expressed as the mean ± SEM, and the statistical significance of differences between means was assessed using an independent sample t test. *p < 0.05, **p < 0.01.

SMUL sORF inhibited the expression level of *SMURF2*, while the other mutant-types had no effect (Figures 5G and 5H). To further study the role of *SMUL* sORF in post-transcriptional regulation, the wild-type and mutant types of *SMUL* sORF were inserted into the 5′ UTR of *hluc* of the psiCHECK-2 dual-luciferase reporter vector. A reporter gene assay showed that the *SMUL* sORF suppressed the translation of the downstream gene (Figures 5I and 5J).
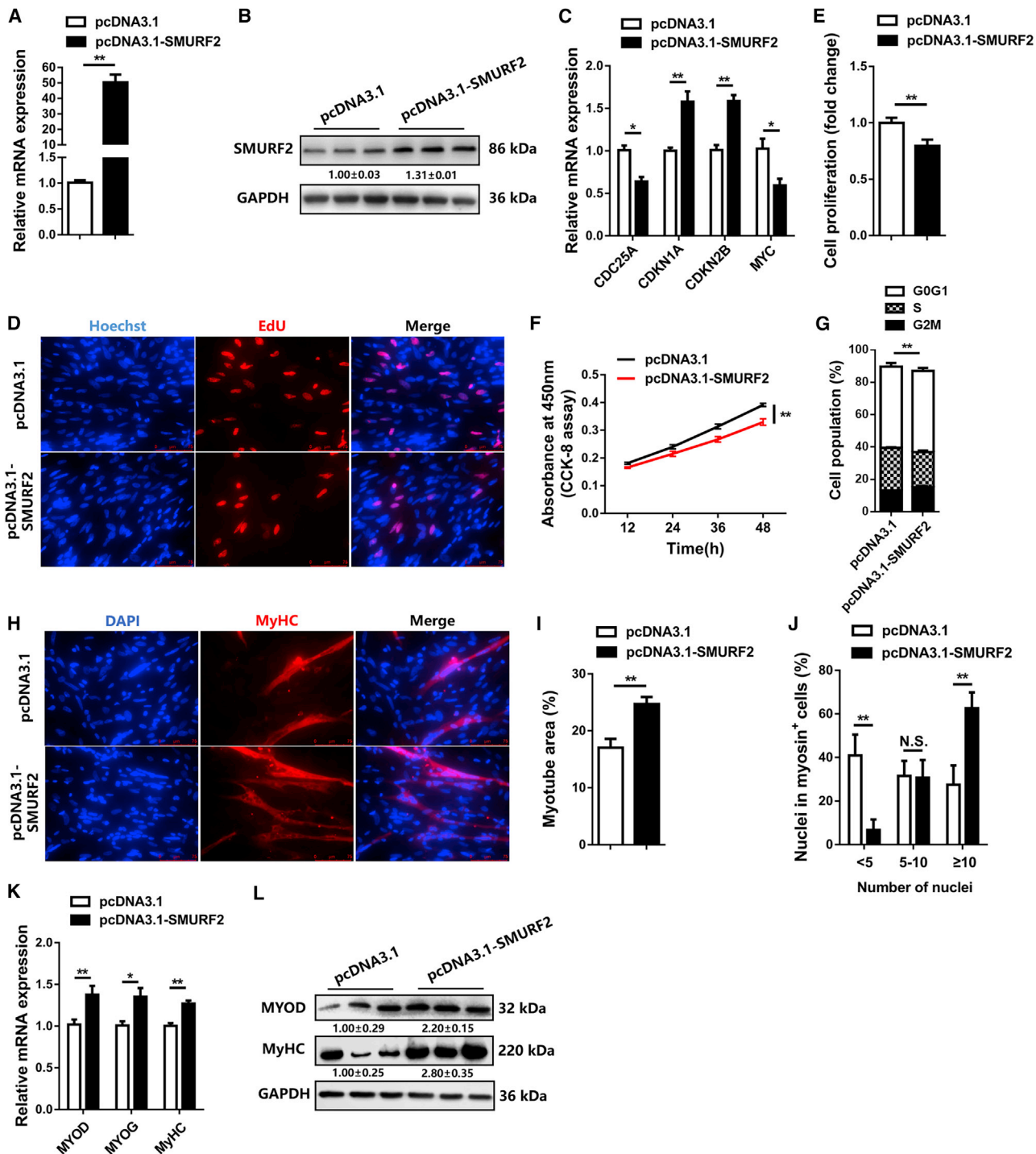
Figure 6. *SMURF2* inhibits myoblast proliferation and promotes myoblast differentiation

(A and B) Relative mRNA (A) (n = 4) and protein (B) (n = 3) expression levels of *SMURF2* with *SMURF2* overexpression. The numbers shown below the bands were folds of band intensities relative to control. Band intensities were quantified by ImageJ and normalized to *GAPDH*. Data are expressed as a fold change relative to the control. (C) The relative mRNA expression of several cell cycle genes induced by *SMURF2* overexpression (n = 4). (D) EdU proliferation assays for CPMs with the overexpression of *SMURF2* (n = 3); (E) the numbers of proliferative cells were also counted (n = 3). (F) Cell growth was measured after *SMURF2* overexpression (n = 6). (G) Cell cycle analysis of CPMs with *SMURF2* overexpression (n = 4). (H–J) MyHC staining (H) (n = 3), myotube area (%) (I) (n = 3), and myoblast fusion index (J) (n = 3) of CPMs at 72 h after overexpression of *SMURF2*. (K and L) Relative mRNA (K) (n = 4) and protein (L) (n = 3) expression of the differentiation marker genes after transfection with pcDNA3.1-SMURF2. The numbers

(legend continued on next page)

SMUL sORF is located upstream of the *SMURF2* mORF, and its presence restrains *SMURF2* production. We hypothesized that *SMUL* may suppress *SMURF2* expression due to the introduction of a premature termination codon (PTC) at the 5′ UTR of *SMURF2*, which induces NMD. RNA half-life was measured to study whether *SMUL* disrupts *SMURF2* mRNA stability. Both *SMUL* full-length and *SMUL* sORF accelerated the degradation of *SMURF2* mRNA, while the stability of *SMURF2* mRNA was maintained with *SMUL* lncRNA interference (Figure 5K; Figures S4A and S4B). To validate that NMD occurs, *UPF1* (an essential component of the NMD machinery) was knocked down by small interfering RNA (siRNA) (Figure 5L). As predicted, NMD inhibition significantly increased the expression level of the *SMURF2* transcript and rescued the inhibitory effect from *SMUL* full-length and *SMUL* sORFs (Figure 5M; Figures S4C and S4D). To rule out possible artifacts of *UPF1* depletion unrelated to NMD, we also indirectly inhibited NMD by blocking translation with cycloheximide (CHX), which generated similar results (Figure 5N; Figures S4E and S4F).

The biological functions of wild-type and mutant *SMUL* sORFs in myogenesis were also evaluated. Similar results to *SMUL* full-length sORF were found with wild-type *SMUL* sORF overexpression, which contains the same sORF (Figure S5). In contrast, the mutant constructs, both of which do not express *SMUL* sORF, did not promote myoblast proliferation and restrain myoblast differentiation (Figure S5). Taken together, these data revealed that *SMUL* sORF suppressed *SMURF2* production-regulated myogenesis via NMD.

SMURF2 is involved in myogenesis by negatively regulating the TGF-β/SMAD pathway

To further identify the role of *SMURF2* in myoblast proliferation and differentiation, *SMURF2* overexpression and knockdown in CPMs were performed (Figures 6A and 6B; Figures S6A and S6B). *SMURF2* overexpression significantly suppressed myoblast proliferation (Figures 6C–6F). This also resulted in a large number of G_0/G_1 cells and few S phase cells (Figure 6G). Furthermore, overexpression of *SMURF2* significantly promoted the formation of myotubes and induced myoblast fusion, as well as upregulated the expression levels of myoblast differentiation marker genes (Figures 6H–6L). However, the interference of *SMURF2* induced myoblast proliferation and suppressed myoblast differentiation (Figures S6C–S6L).

To investigate the function of *SMURF2* in skeletal muscle *in vivo*, lentiviral-mediated *SMURF2* overexpression (LV-SMURF2) and lentiviral-mediated *SMURF2* knockdown (LV-shSMURF2) were introduced into gastrocnemius muscle of 1-day-old chicks (Figures 7A and 7B; Figures S6M and S6N). Overexpression of *SMURF2* promoted muscle fiber hypertrophy, as well as reduced the expression of atrophy-related genes and suppressed autophagy (Figures 7C–7E). In

addition, *SMURF2* overexpression induced the slow-twitch muscle phenotype and improved the oxidative capacity of gastrocnemius muscle (Figures 7F–7J). Inversely, gastrocnemius muscle fibers were induced to atrophy, and the fast-twitch muscle program was activated with *SMURF2* knockdown (Figures S6O–S6V).

It is well established that *SMURF2* is an upstream regulatory gene of the TGF-β/SMAD pathway, and this pathway is widely involved in the regulation of cell growth, differentiation, and development. We further assessed the TGF-β/SMAD pathway and found that *SMURF2* overexpression downregulated the expression of *TGFB1* and inhibited the phosphorylation of SMAD2 and SMAD3 (Figures 7K and 7L). Conversely, this pathway was activated with the knockdown of *SMURF2* (Figures 7K and 7L). More importantly, the contrary results were found with *SMUL* overexpression and knockdown (Figures 7M and 7N), clearly demonstrating that *SMUL* participated in the TGF-β/SMAD pathway by modulating *SMURF2*.

DISCUSSION

With the development of genome research, it has been found that the transcription of the animal genome is not only ubiquitous, but also incredibly complex. In the past, lncRNA was considered part of the "noise" of the transcribed genome, without translation potential. However, increasing evidence has shown that some lncRNAs function biologically through translation.[31,32,43,44] In the present study, we first compared transcriptome and peptidome data during myoblast proliferation and differentiation in chickens and discovered up to 104 lncRNAs that translate at least one small peptide.

So far, only a limited number of lncRNAs have been well characterized, with a diverse array of mechanisms.[45] Previous studies on the regulation mechanism of lncRNAs have been mainly focused on how lncRNAs adsorb microRNAs (miRNAs) to regulate gene expression by acting as competing endogenous RNAs (ceRNAs).[10–12,46,47] However, how translated lncRNAs function in concert with their target genes is still not clear. To uncover this novel regulation of lncRNAs and identify the key molecular players in chicken muscle development, in this study we constructed a uORF (which is transcribed by lncRNA) regulatory network.

NMD serves as a cellular surveillance mechanism and can deplete transcripts that show signs of premature translational termination.[48] Genome-wide research in multiple organisms[49–52] demonstrated that uORF-bearing transcripts are particularly susceptible to targeted degradation by NMD, attributed to the 5′-near termination events at uORF termination codons. In this study, we found that the termination codon of the lncRNA *SMUL*-transcribed sORF is located at the 5′ UTR of *SMURF2*. *SMUL*'s translation disrupted the stability of *SMURF2* and suppressed *SMURF2* production. More importantly, *SMUL*-induced NMD was confirmed directly and indirectly by

---

shown below the bands are fold changes of band intensities relative to the control. Band intensities were quantified by ImageJ and normalized to *GAPDH*. Data are expressed as a fold change relative to the control. Data are presented as means ± SEM. Statistical significance of differences between the means was assessed using an independent sample t test. *p < 0.05, **p < 0.01.
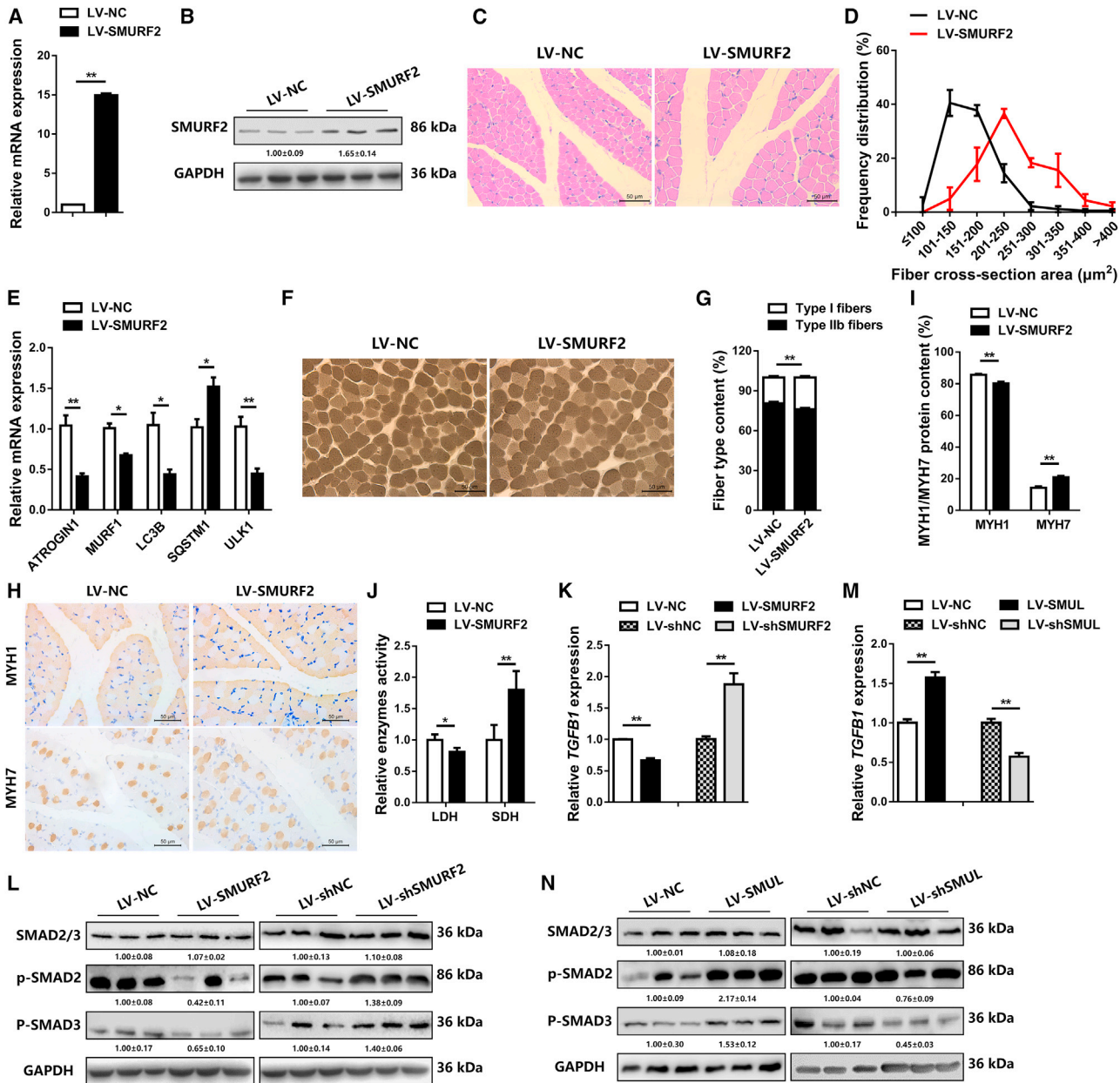
Figure 7. *SMURF2* rescues skeletal muscle atrophy and drives a switch from fast-twitch to slow-twitch fibers via the TGF-β/SMAD pathway

(A and B) Relative *SMURF2* mRNA (A) (n = 3) and protein (B) (n = 4) expression in gastrocnemius muscle with infection of *SMURF2*-expressing lentivirus (LV-SMURF2) or LV-NC. (C and D) H&E staining (C) (n = 4) and frequency distribution of CSA (D) (n = 4) of gastrocnemius muscle injected with the listed lentivirus. (E) Relative mRNA expression of the atrophy and autophagy-related genes in gastrocnemius muscle after infection with the indicated lentivirus (n = 4). (F and G) ATPase staining (F) (n = 4) and fiber type content (G) (n = 4) of gastrocnemius muscle infected with the indicated lentivirus. (H and I) Immunohistochemistry analysis of MYH1/MYH7 (H) (n = 4) and MYH1/MYH7 protein content (I) (n = 4) of gastrocnemius muscle injected with the indicated lentivirus. (J) Relative enzymes activity of LDH and SDH in gastrocnemius muscle infected with the listed lentivirus (n = 4). (K and M) Relative *TGFB1* mRNA expression in gastrocnemius muscle infected with the indicated lentivirus (n = 4). (L and N) Protein expression levels of the TGF-β/SMAD pathway with infection of the indicated lentivirus (n = 3). Results are expressed as the mean ± SEM. Statistical significance of differences between means was assessed using independent sample t test. *p < 0.05, **p < 0.01.

NMD inhibition models, which demonstrated that lncRNA *SMUL*-transcribed sORF can work as a uORF, modulating *SMURF2* mRNA via NMD by introducing a PTC.

Although several lncRNAs have been shown to have roles in myoblast proliferation and differentiation *in vitro*,[11,12,53,54] little is known about their function in skeletal muscle atrophy and fiber remodeling *in vivo*.

In this study, we found that *SMUL* is highly expressed in skeletal muscle, and its expression decreases with myoblast differentiation. Gain- and loss-of-function analysis revealed that *SMUL* promotes myoblast proliferation and suppresses myoblast differentiation *in vitro*.

lncRNAs are important in regulating gene expression. Thus, lncRNAs have received wide attention as pivotal players in a variety of physiological and pathological processes, including skeletal muscle atrophy.[12] The ubiquitin-proteasome system (UPS) and autophagy-lysosomal system are well known to regulate skeletal muscle atrophy.[4,55–58] In this study, *in vivo* experiments demonstrated that *SMUL* promotes skeletal muscle atrophy by activating proteasomal degradation and autophagy, as well as drives the transformation of slow-twitch muscle fibers to fast-twitch muscle fibers. From a pathway perspective, *SMUL* modulates myogenesis and muscle atrophy by restraining *SMURF2* to activate the TGF-β/SMAD pathway.

Our findings present a novel model showing that lncRNA can play a regulatory role through its translation process rather than its translation products to modulate myogenesis and muscle atrophy. These findings, which are expected to shape the direction of future developmental research, open a small gate to further study on the lncRNA mechanism of action.

MATERIALS AND METHODS

Ethics statement
All animal studies were sanctioned by the Institutional Animal Care and Use Committee at the South China Agricultural University. All experiments were performed according to the regulations and guidelines established by this committee and by international standards for animal welfare.

Cell culture, transfection, and treatment
CPMs were isolated from leg muscles of embryonic day 11 (E11) as previously described.[59] The obtained cells were cultured in growth medium consisting of Roswell Park Memorial Institute (RPMI) 1640 medium (Gibco, Gaithersburg, MD, USA) with 20% fetal bovine serum (FBS; Gibco). To induce myogenic differentiation, growth medium was removed and replaced with differentiation medium (RPMI 1640 medium containing 2% horse serum) after myoblasts achieved 90% cell confluence. Cells were incubated at 37°C with an atmosphere of 5% CO_2.

All transient transfections used Lipofectamine 3000 reagent (Invitrogen, USA) following the manufacturer's protocol.

To indirectly block NMD, primary myoblasts were treated for 6 h with 100 μg/mL CHX (Sigma-Aldrich, St. Louis, MO, USA) or the vehicle dimethyl sulfoxide (DMSO) as a control. To detect the half-life of *SMURF2* mRNA, myoblasts were treated with 2 μg/mL actinomycin D (ActD, Sigma-Aldrich, St. Louis, MO, USA), which could inhibit transcription, and harvested at 0, 2, 4, and 6 h after treatment. *GAPDH* mRNA was applied as internal control, which was stable within 32 h.

RNA-seq
CPMs during proliferation and differentiation were used for RNA-seq. Total RNA was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. RNA quantity and quality were evaluated on an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany), and RNA integrity was further examined using agarose gel electrophoresis. Ribosomal RNA (rRNA) was removed by an Epicenter Ribo-Sero rRNA removal kit (Epicenter, USA), and rRNA free residue was cleaned up by ethanol precipitation. Subsequently, sequencing libraries were generated using the rRNA-depleted RNA by NEBNext Ultra directional RNA library prep kit for Illumina (NEB, USA) following the manufacturer's recommendations, and sequenced on an Illumina HiSeq 2500 platform by Beijing Novogene Bioinformation Technology. The raw data of RNA-seq were deposited in the Sequence Read Archive (SRA) database under accession no. PRJNA674499.

DEGs were subjected to enrichment analysis of GO functions and KEGG pathways, respectively.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis
Proliferating and differentiated CPMs were lysed respectively and then mixed with SDS sample buffer. Proteins were fractionated by SDS-PAGE on 12% Bis-Tris acrylamide NuPAGE gels using 2-(*N*-morpholino)ethanesulfonic acid (MES) SDS running buffer (Invitrogen). Lower protein bands (lower than 15 kDa) were excised and subjected to in-gel digestion. The resulting peptides were analyzed by a Q Exactive mass spectrometer coupled to a nano-LC (Advance LC, Michrom) via a nano-electrospray source with a column oven set at 37°C (AMR), as previously reported,[30] by Gene Denovo Biotechnology (Guangzhou, China). Briefly, samples were dissolved and then separated by an in-house-made 20-cm column (inner diameter, 100 μm, 3-μμ L-column, CERI, Japan) with a linear gradient at a flow rate of 250 nL/min. The Q Exactive was operated in a data-dependent mode with survey scans acquired at a resolution of 70,000 at m/z 400. Up to the top 10 most abundant ions with charge 2^+ or 3^+ from the survey scan were selected and fragmented by higher energy collision dissociation. The acquired MS/MS spectra were analyzed with the SEQUEST HT algorithm using a database derived from three-frame (forward only) translation (longer than 5 aa) of lncRNA sequences. The mass spectrometry proteomics data were deposited in the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX partner repository[60] with the dataset identifier PXD021030.

Prediction and construction of interaction network
lncRNAs encoding small peptide identified by mass spectrometry were selected for *cis*-target gene prediction. The sORF of translated lncRNA, which is on the same chain and located within 1.5 kb upstream to an mORF, was predicted to have the regulatory role as uORF. A translated lncRNA sORF-mORF interaction network was constructed using the Cytoscape 3.7.1 program.

RNA extraction, cDNA synthesis, and quantitative real-time PCR

Total RNA was extracted using TRIzol reagent (TaKaRa, Otsu, Japan) as recommended by the supplier. Nuclear and cytoplasmic RNA fractionation was performed by using the Paris kit (Ambion, Life Technologies, USA), following the manufacturer's protocol. cDNA synthesis for mRNA was carried out using the PrimeScript RT reagent kit with genomic DNA (gDNA) eraser (perfect real time) (TaKaRa, Otsu, Japan). Quantitative real-time PCR with an iTaq Universal SYBR Green supermix kit (Bio-Rad, USA) and analyses with the $2^{-\Delta\Delta Ct}$ method were performed as described previously.[61] Chicken β-actin was used as an internal control. Primer pairs for PCR and quantitative real-time PCR are shown in Table S4.

5′ and 3′ RACE

A SMARTer RACE cDNA amplification kit (Clontech, Osaka, Japan) was used to obtain the full-length sequence of *SMUL* as recommended by the supplier. Nested-PCR reactions were performed. The products of the RACE PCR were cloned into the pJET 1.2/blunt cloning vector (CloneJET PCR cloning kit; Fermentas, Glen Burnie, MD, USA) and sequenced by Sangon Biotech (Shanghai, China). All primers used in RACE are summarized in Table S4.

RNA *in situ* hybridization

RNA *in situ* hybridization was performed by using a fluorescence *in situ* hybridization (FISH) kit (C10910, RiboBio, Guangzhou, China), following the manufacturer's protocol. Special FISH probes against *SMUL*, which were modified by Cy3, were also designed and synthesized by Guangzhou RiboBio (Guangzhou, China).

Plasmid construction and RNA oligonucleotides

For EGFP fusion protein construction, the ORF and ORFmut (in which the start codon ATG was mutated to ATT) in the full-length *SMUL* transcript were amplified and cloned into the pEGFP-N1 mutation vector (pGFPmut) by using the *NheI* and *XhoI* restriction sites, in which the GFP start codon (ATGGTG) was mutated to ATTGTT (Clontech, Osaka, Japan).

For FLAG fusion protein construction, the ORF and ORFmut in the full-length *SMUL* transcript were amplified by PCR and then subcloned into *HindIII* and *XhoI* restriction sites of the pcDNA3.1-3xFLAG-C vector.

For overexpression vector construction, the full-length sequence, ORF wild-type, ORF ATG mutant (in which the start codon ATG was mutated to ATT), ORF TGA mutant (in which the termination codon TGA was mutated to TGG), and ORF both ATG and TGA mutant (in which both the start codon ATG was mutated to ATT and the termination codon TGA was mutated to TGG) of *SMUL* and *SMURF2* coding sequence (NCBI: XM_425380.6) were amplified and cloned into the expression plasmid pcDNA-3.1 (Promega, Madison, WI, USA) by using *HindIII* and *XhoI* restriction sites.

For psiCHECK-2 dual-luciferase reporter vectors constructed, the ORF wild-type, ORF ATG mutant, ORF TGA mutant, and ORF both ATG and TGA mutant of *SMUL* were amplified and subcloned into *NheI* restriction sites in the psiCHECK-2 dual-luciferase reporter vector (Promega, Madison, WI, USA).

For overexpression lentiviral vectors constructed, the full lengths of *SMUL* and *SMURF2* coding sequences were amplified and then cloned into the pLVX-mCMV-ZsGreen-IRES-Puro vector (Addgene, Cambridge, MA, USA) between the *SpeI* and *NotI* sites. Short hairpin RNAs (shRNAs) against *SMUL* or *SMURF2* were designed by Shanghai Hanbio Biotechnology and then subcloned into the pLVX-shRNA2-Puro vector (Addgene, Cambridge, MA, USA) by using the *BamHI* and *EcoRI* restriction sites.

SMUL is a lncRNA molecule present in the cytoplasm and nucleus. The siRNAs and antisense oligonucleotides (ASOs) that were used for the specific knockdown of *SMUL* in the cytoplasm and nucleus, respectively, were designed and synthesized by Guangzhou RiboBio (Guangzhou, China). The siRNAs against *SMURF2* and *UPF1* were also designed and synthesized.

The primers and oligonucleotide sequences used in this study are listed in Tables S4 and S5.

Immunoblotting and immunofluorescence

Western blot analysis was performed as previously described.[59] The primary antibodies used were anti-GFP (50430-2-AP, 1:1,000, Proteintech), anti-FLAG (AF519, 1:1,000, Beyotime), anti-MYOD (ABP53067, 1:500, Abbkine), anti-MyHC (B103, 0.5 μg/mL, DHSB), anti-SMURF2 (ab94483, 1 μg/mL, Abcam), anti-SMAD2/3 (70R-51804, 1:500, Fitzgerald), anti-phosphorylated (phospho-) SMAD2 (bs-3419R, 1:500, Bioss), anti-phospho-SMAD3 (bs-3425R, 1:500, Bioss), anti-β-actin (bsm-33036M, 1:1,000, Bioss), and anti-GAPDH (60004-1-Ig, 1:5,000, Proteintech). ProteinFind goat anti-mouse immunoglobulin G (IgG) (H+L), horseradish peroxidase (HRP) conjugate (HS201-01, 1:1,000, TransGen Biotech, Beijing, China) and ProteinFind goat anti-rabbit IgG (H+L), HRP conjugate (HS101-01, 1:500, TransGen Biotech) were used as a secondary antibody.

The immunofluorescence was performed using anti-FLAG (AF519, 1:1,000, Beyotime) or anti-MyHC (B103, 2.5 μg/mL, DHSB). Images were obtained with a fluorescence microscope (DMi8; Leica, Germany). The area of cells labeled with anti-MyHC was measured and calculated as previously described.[59]

Flow cytometry, EdU, and CCK-8 assays

For the flow cytometry analysis of the cell cycle, myoblasts were seeded in 12-well plates. After a 48-h transfection, the cultured cells in growth media were collected and fixed overnight in 70% ethanol at −20°C. With a cell cycle analysis kit (Thermo Fisher Scientific, USA), the cells were analyzed by a BD Accuri C6 flow cytometer (BD Biosciences, San Jose, CA, USA), and the data were processed using FlowJo software (7.6, Tree Star, Ashland, OR, USA).

For the EdU assay, primary myoblasts seeded in 24-well plates were cultured to 50% density and then transfected. Forty-eight hours after transfection, the cells were fixed and stained with a C10310 EdU Apollo *in vitro* imaging kit (RiboBio, China; 50 μM) as previously described.[21] A fluorescence microscope (DMi8; Leica, German) was used to capture three randomly selected fields to visualize the number of EdU-stained cells.

For the CCK-8 assay, primary myoblasts were seeded in a 96-well plate and cultured in growth medium. After being transfected, the proliferation of the cell culture was monitored at 12, 24, 36, and 48 h using the TransDetect CCK (TransGen Biotech, Beijing, China), according to the manufacturer's protocol. The data for absorbance at 450 nm were read by an iMark microplate absorbance reader (Bio-Rad). All of the data were acquired by averaging the results from six independent repeats.

Lentivirus production and transduction
To generate lentivirus, the recombinant lentiviral expression vectors were co-transfected with packaging plasmid psPAX2 (Addgene, USA) and envelope plasmid pMD2.G (Addgene, USA) into 293T cells using Lipofectamine 3000 reagent. Infectious particles were harvested at 48 and 72 h after transfection, filtered through 0.45-μm polyvinylidene fluoride (PVDF) membranes (Millipore, CA, USA), and concentrated by ultracentrifugation. The viral titer was evaluated by a gradient-dilution method.

Thirty-two 1-day-old chicks were randomly divided into four groups (n = 8): (1) LV-SMUL and LV-NC, (2) LV-shSMUL and LV-shNC, (3) LV-SMURF2 and LV-NC, and (4) LV-shSMURF2 and LV-shNC. Chicks received two intramuscular doses (at days 1 and 7) of lentivirus (10^6 titers) in two different sites of the gastrocnemius muscle. Thirteen days after the initial injection, chickens were euthanized. Subsequently, gastrocnemius muscles were collected after rapid dissection, then immediately frozen in liquid nitrogen and stored at −80°C.

Hematoxylin and eosin (H&E) staining, myosin-ATPase staining, and immunohistochemistry
For H&E staining, gastrocnemius muscle tissues were immersed in 4% paraformaldehyde and then embedded in paraffin and cut into 4-μm-thick transverse sections. Subsequently, the sections were stained with H&E.

Myosin-ATPase staining was carried out using an ATPase staining solution kit (G2380, Solarbio, Beijing, China), following the manufacturer's instructions. Myosin-ATPase staining was performed at pH 10.4. Under these alkaline conditions, MyHC I isoforms were inactivated while MyHC IIb isoforms were still functional, resulting in addition of black dye to MyHC IIb-positive muscle fibers.

An SP-POD kit (SP0041, Solarbio, Beijing, China) was used for immunohistochemistry as recommended by the supplier. The primary antibodies included anti-MYH1 (F59, 1:100, DHSB) and anti-MYH7 (S58, 1:300, DHSB) and were used for labeling the signals.

Enzyme activities assay
The glycolytic capacity of skeletal muscle was evaluated by the activity of LDH, while the oxidative capacity of skeletal muscle was evaluated by the activity of SDH. Enzyme activities were measured by commercial assay kits (BC0685 and BC0955) that were purchased from Beijing Solarbio Science & Technology.

Dual-luciferase reporter assay
The psiCHECK-2 recombinant vectors were transfected into primary myoblasts using Lipofectamine 3000 reagent (Invitrogen, Carlsbad, CA, USA) in 96-well plates. At 48 h after transfection, the firefly and Renilla luciferase activities were measured using a fluorescence/multi-detection microplate reader (BioTek, Winooski, VT, USA) and a Dual-Glo luciferase assay system kit (Promega, USA). The firefly luciferase activities were normalized to Renilla luminescence in each well.

Statistical analysis
Each experiment was repeated at least three times, and all data are expressed as means ± SEM. Where applicable, the statistical significance of the data was tested using one-sample or paired t tests. The types of tests and the p values, when applicable, are indicated in the figure legends.

SUPPLEMENTAL INFORMATION
Supplemental Information can be found online at https://doi.org/10.1016/j.omtn.2020.12.003.

AUTHOR CONTRIBUTIONS
Q.N. and X.Z. conceived and designed the study. B.C. and Z.L. performed the experiments, interpreted the data, and wrote the paper. M.M., J.Z., and S.K. performed the experiments. B.A.A., H.X., E.J., and R.A.L. interpreted the data. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS
The authors declare no competing interests.

REFERENCES
1. Sun, L., Si, M., Liu, X., Choi, J.M., Wang, Y., Thomas, S.S., Peng, H., and Hu, Z. (2018). Long-noncoding RNA Atrolnc-1 promotes muscle wasting in mice with chronic kidney disease. J. Cachexia Sarcopenia Muscle 9, 962–974.

2. Tieland, M., Trouwborst, I., and Clark, B.C. (2018). Skeletal muscle performance and ageing. J. Cachexia Sarcopenia Muscle 9, 3–19.

3. Buckingham, M. (2006). Myogenic progenitor cells and skeletal myogenesis in vertebrates. Curr. Opin. Genet. Dev. 16, 525–532.

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

5. Buckingham, M., and Rigby, P.W. (2014). Gene regulatory networks and transcriptional mechanisms that control myogenesis. Dev. Cell 28, 225–238.

6. Sartorelli, V., and Fulco, M. (2004). Molecular and cellular determinants of skeletal muscle atrophy and hypertrophy. Sci. STKE 2004, re11.

7. Bassel-Duby, R., and Olson, E.N. (2006). Signaling pathways in skeletal muscle remodeling. Annu. Rev. Biochem. 75, 19–37.

8. Hubé, F., Velasco, G., Rollin, J., Furling, D., and Francastel, C. (2011). Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation. Nucleic Acids Res. 39, 513–525.

9. Lu, L., Sun, K., Chen, X., Zhao, Y., Wang, L., Zhou, L., Sun, H., and Wang, H. (2013). Genome-wide survey by ChIP-seq reveals YY1 regulation of lincRNAs in skeletal myogenesis. EMBO J. 32, 2575–2588.

10. Zhu, M., Liu, J., Xiao, J., Yang, L., Cai, M., Shen, H., Chen, X., Ma, Y., Hu, S., Wang, Z., et al. (2017). lnc-mg is a long non-coding RNA that promotes myogenesis. Nat. Commun. 8, 14718.

11. Ma, M., Cai, B., Jiang, L., Abdalla, B.A., Li, Z., Nie, Q., and Zhang, X. (2018). lncRNA-Six1 is a target of miR-1611 that functions as a ceRNA to regulate Six1 protein expression and fiber type switching in chicken myogenesis. Cells 7, 243.

12. Li, Z., Cai, B., Abdalla, B.A., Zhu, X., Zheng, M., Han, P., Nie, Q., and Zhang, X. (2019). lncIRS1 controls muscle atrophy via sponging miR-15 family to activate IGF1-PI3K/AKT pathway. J. Cachexia Sarcopenia Muscle 10, 391–410.

13. Song, X., Cao, G., Jing, L., Lin, S., Wang, X., Zhang, J., Wang, M., Liu, W., and Lv, C. (2014). Analysing the relationship between lncRNA and protein-coding gene and the role of lncRNA as ceRNA in pulmonary fibrosis. J. Cell. Mol. Med. 18, 991–1003.

14. Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 25, 1915–1927.

15. Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature 489, 101–108.

16. Dinger, M.E., Pang, K.C., Mercer, T.R., and Mattick, J.S. (2008). Differentiating protein-coding and noncoding RNA: challenges and ambiguities. PLoS Comput. Biol. 4, e1000176.

17. Ulitsky, I., and Bartel, D.P. (2013). lincRNAs: genomics, evolution, and mechanisms. Cell 154, 26–46.

18. Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., et al. (1996). Life with 6000 genes. Science 274, 546–, 563–567.

19. van Heesch, S., van Iterson, M., Jacobi, J., Boymans, S., Essers, P.B., de Bruijn, E., Hao, W., MacInnes, A.W., Cuppen, E., and Simonis, M. (2014). Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. Genome Biol. 15, R6.

20. Wilson, B.A., and Masel, J. (2011). Putatively noncoding transcripts show extensive association with ribosomes. Genome Biol. Evol. 3, 1245–1252.

21. Cai, B., Li, Z., Ma, M., Wang, Z., Han, P., Abdalla, B.A., Nie, Q., and Zhang, X. (2017). lncRNA-Six1 encodes a micropeptide to activate Six1 in cis and is involved in cell proliferation and muscle growth. Front. Physiol. 8, 230.

22. Anderson, D.M., Anderson, K.M., Chang, C.L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R., and Olson, E.N. (2015). A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell 160, 595–606.

23. Nath, D., and Shadan, S. (2009). The ubiquitin system. Nature 458, 421.

24. Rotin, D., and Kumar, S. (2009). Physiological functions of the HECT family of ubiquitin ligases. Nat. Rev. Mol. Cell Biol. 10, 398–409.

25. Bonni, S., Wang, H.R., Causing, C.G., Kavsak, P., Stroschein, S.L., Luo, K., and Wrana, J.L. (2001). TGF-β induces assembly of a Smad2-Smurf2 ubiquitin ligase complex that targets SnoN for degradation. Nat. Cell Biol. 3, 587–595.

26. Inoue, Y., and Imamura, T. (2008). Regulation of TGF-β family signaling by E3 ubiquitin ligases. Cancer Sci. 99, 2107–2112.

27. Tang, L.Y., Yamashita, M., Coussens, N.P., Tang, Y., Wang, X., Li, C., Deng, C.X., Cheng, S.Y., and Zhang, Y.E. (2011). Ablation of Smurf2 reveals an inhibition in TGF-β signalling through multiple mono-ubiquitination of Smad3. EMBO J. 30, 4777–4789.

28. Cai, Y., Zhou, C.H., Fu, D., and Shen, X.Z. (2012). Overexpression of Smad ubiquitin regulatory factor 2 suppresses transforming growth factor-β mediated liver fibrosis. J. Dig. Dis. 13, 327–334.

29. Cao, S., Xiao, L., Rao, J.N., Zou, T., Liu, L., Zhang, D., Turner, D.J., Gorospe, M., and Wang, J.Y. (2014). Inhibition of Smurf2 translation by miR-322/503 modulates TGF-β/Smad2 signaling and intestinal epithelial homeostasis. Mol. Biol. Cell 25, 1234–1243.

30. Matsumoto, A., Pasut, A., Matsumoto, M., Yamashita, R., Fung, J., Monteleone, E., Saghatelian, A., Nakayama, K.I., Clohessy, J.G., and Pandolfi, P.P. (2017). mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature 541, 228–232.

31. Bazin, J., Baerenfaller, K., Gosai, S.J., Gregory, B.D., Crespi, M., and Bailey-Serres, J. (2017). Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. Proc. Natl. Acad. Sci. USA 114, E10018–E10027.

32. Lu, S., Zhang, J., Lian, X., Sun, L., Meng, K., Chen, Y., Sun, Z., Yin, X., Li, Y., Zhao, J., et al. (2019). A hidden human proteome encoded by "non-coding" genes. Nucleic Acids Res. 47, 8111–8125.

33. Yap, K.L., Li, S., Muñoz-Cabello, A.M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M.J., and Zhou, M.M. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol. Cell 38, 662–674.

34. Dimitrova, N., Zamudio, J.R., Jong, R.M., Soukup, D., Resnick, R., Sarma, K., Ward, A.J., Raj, A., Lee, J.T., Sharp, P.A., and Jacks, T. (2014). lincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. Mol. Cell 54, 777–790.

35. Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc. Natl. Acad. Sci. USA 106, 7507–7512.

36. Barbosa, C., Peixeiro, I., and Romão, L. (2013). Gene expression regulation by upstream open reading frames and human disease. PLoS Genet. 9, e1003529.

37. Sîrca, A., and Susec-Michieli, M. (1980). Selective type II fibre muscular atrophy in patients with osteoarthritis of the hip. J. Neurol. Sci. 44, 149–159.

38. Fiori, M.G., Andreola, S., Ladelli, G., and Scirea, M.R. (1983). Selective atrophy of the type IIb muscle fibers in rheumatoid arthritis and progressive systemic sclerosis (scleroderma). A biopsy histochemical study. Eur. J. Rheumatol. Inflamm. 6, 168–181.

39. Diffee, G.M., Caiozzo, V.J., Herrick, R.E., and Baldwin, K.M. (1991). Contractile and biochemical properties of rat soleus and plantaris after hindlimb suspension. Am. J. Physiol 260, C528–C534.

40. Mozdziak, P.E., Greaser, M.L., and Schultz, E. (1999). Myogenin, MyoD, and myosin heavy chain isoform expression following hindlimb suspension. Aviat. Space Environ. Med. 70, 511–516.

41. Fitts, R.H., Riley, D.R., and Widrick, J.J. (2000). Physiology of a microgravity environment invited review: microgravity and skeletal muscle. J Appl Physiol (1985) 89, 823–839.

42. Potthoff, M.J., Arnold, M.A., McAnally, J., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2007). Regulation of skeletal muscle sarcomere integrity and postnatal muscle function by Mef2c. Mol. Cell. Biol. 27, 8143–8151.

43. Huang, J.Z., Chen, M., Chen, D., Gao, X.C., Zhu, S., Huang, H., Hu, M., Zhu, H., and Yan, G.R. (2017). A peptide encoded by a putative lncRNA HOXA-AS3 suppresses colon cancer growth. Mol. Cell 68, 171–184.e6.

44. Nelson, B.R., Makarewich, C.A., Anderson, D.M., Winders, B.R., Troupes, C.D., Wu, F., Reese, A.L., McAnally, J.R., Chen, X., Kavalali, E.T., et al. (2016). A peptide

encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science *351*, 271–275.

45. Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol. Cell *43*, 904–914.

46. Zhang, Z.K., Li, J., Guan, D., Liang, C., Zhuo, Z., Liu, J., Lu, A., Zhang, G., and Zhang, B.T. (2018). A newly identified lncRNA MAR1 acts as a miR-487b sponge to promote skeletal muscle differentiation and regeneration. J. Cachexia Sarcopenia Muscle *9*, 613–626.

47. Zhang, Z.K., Li, J., Guan, D., Liang, C., Zhuo, Z., Liu, J., Lu, A., Zhang, G., and Zhang, B.T. (2018). Long noncoding RNA lncMUMA reverses established skeletal muscle atrophy following mechanical unloading. Mol. Ther. *26*, 2669–2680.

48. He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast. Mol. Cell *12*, 1439–1452.

49. Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev. *21*, 1833–1856.

50. Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. *36*, 1073–1078.

51. Iacono, M., Mignone, F., and Pesole, G. (2005). uAUG and uORFs in human and rodent 5′untranslated mRNAs. Gene *349*, 97–105.

52. Ramani, A.K., Nelson, A.C., Kapranov, P., Bell, I., Gingeras, T.R., and Fraser, A.G. (2009). High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective *Caenorhabditis elegans*. Genome Biol. *10*, R101.

53. Jin, J.J., Lv, W., Xia, P., Xu, Z.Y., Zheng, A.D., Wang, X.J., Wang, S.S., Zeng, R., Luo, H.M., Li, G.L., and Zuo, B. (2018). Long noncoding RNA *SYISL* regulates myogenesis by interacting with polycomb repressive complex 2. Proc. Natl. Acad. Sci. USA *115*, E9802–E9811.

54. Zhou, L., Sun, K., Zhao, Y., Zhang, S., Wang, X., Li, Y., Lu, L., Chen, X., Chen, F., Bao, X., et al. (2015). Linc-YY1 promotes myogenic differentiation and muscle regeneration through an interaction with the transcription factor YY1. Nat. Commun. *6*, 10026.

55. Bodine, S.C., Latres, E., Baumhueter, S., Lai, V.K., Nunez, L., Clarke, B.A., Poueymirou, W.T., Panaro, F.J., Na, E., Dharmarajan, K., et al. (2001). Identification of ubiquitin ligases required for skeletal muscle atrophy. Science *294*, 1704–1708.

56. Gomes, M.D., Lecker, S.H., Jagoe, R.T., Navon, A., and Goldberg, A.L. (2001). Atrogin-1, a muscle-specific F-box protein highly expressed during muscle atrophy. Proc. Natl. Acad. Sci. USA *98*, 14440–14445.

57. Sandri, M., Sandri, C., Gilbert, A., Skurk, C., Calabria, E., Picard, A., Walsh, K., Schiaffino, S., Lecker, S.H., and Goldberg, A.L. (2004). Foxo transcription factors induce the atrophy-related ubiquitin ligase atrogin-1 and cause skeletal muscle atrophy. Cell *117*, 399–412.

58. Sandri, M. (2013). Protein breakdown in muscle wasting: role of autophagy-lysosome and ubiquitin-proteasome. Int. J. Biochem. Cell Biol. *45*, 2121–2129.

59. Cai, B., Ma, M., Chen, B., Li, Z., Abdalla, B.A., Nie, Q., and Zhang, X. (2018). miR-16-5p targets *SESN1* to regulate the *p53* signaling pathway, affecting myoblast proliferation and apoptosis, and is involved in myoblast differentiation. Cell Death Dis. *9*, 367.

60. Ma, J., Chen, T., Wu, S., Yang, C., Bai, M., Shu, K., Li, K., Zhang, G., Jin, Z., He, F., et al. (2019). iProX: an integrated proteome resource. Nucleic Acids Res. *47* (D1), D1211–D1217.

61. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ method. Methods *25*, 402–408.