HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2022 March 21.

Published in final edited form as:

Cell Rep. 2022 February 15; 38(7): 110361. doi:10.1016/j.celrep.2022.110361.

Distinct *MUNC* lncRNA structural domains regulate transcription of different promyogenic factors

Roza K. Przanowska[1,2], **Chase A. Weidmann**[3,4], **Shekhar Saha**[1], **Magdalena A. Cichewicz**[1], **Kate N. Jensen**[1], **Piotr Przanowski**[1,2], **Patrick S. Irving**[3], **Kevin A. Janes**[1,2], **Michael J. Guertin**[1,5], **Kevin M. Weeks**[3], **Anindya Dutta**[1,6,7,*]

[1]Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA 22908, USA

[2]Department of Biomedical Engineering, University of Virginia School of Engineering, Charlottesville, VA 22908, USA

[3]Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599, USA

[4]Department of Biological Chemistry and Center for RNA Biomedicine, University of Michigan Medical School, Ann Arbor, MI 48103, USA

[5]R. D. Berlin Center for Cell Analysis and Modeling, University of Connecticut, Farmington, CT 06030, USA

[6]Department of Genetics, University of Alabama, Birmingham, AL 35233, USA

[7]Lead contact

SUMMARY

Many lncRNAs have been discovered using transcriptomic data; however, it is unclear what fraction of lncRNAs is functional and what structural properties affect their phenotype. *MUNC* lncRNA (also known as ^{DRR}eRNA) acts as an enhancer RNA for the *Myod1* gene in *cis* and stimulates the expression of other promyogenic genes in *trans* by recruiting the cohesin complex. Here, experimental probing of the RNA structure revealed that *MUNC* contains multiple structural domains not detected by prediction algorithms in the absence of experimental information. We show that these specific and structurally distinct domains are required for induction of promyogenic genes, for binding genomic sites and gene expression regulation, and for binding

*Correspondence: duttaa@uab.edu.

AUTHOR CONTRIBUTIONS

R.K.P. conceived of the project and, together with A.D., designed the experiments. C.A.W. and K.M.W. designed, analyzed, and interpreted all SHAPE experiments. P.S.I. performed all SHAPE-MaP sequencing. R.K.P. and P.P. performed all other bioinformatic analyses with help and guidance from M.J.G. M.A.C. prepared *MUNC* ChIRP-seq libraries. R.K.P. cloned all *MUNC* mutant vectors with help of K.N. and support from K.A.J. S.S. performed SMC3 RIP-qPCR experiment, which was designed and analyzed by R.K.P. R.K.P. performed RNA-seq, SHAPE-MaP, differentiation assays, RT-qPCR, western blots, MHC immunofluorescence, and ChIRP-qPCR. R.K.P., C.A.W., K.N.J., and A.D. wrote the manuscript with input from all authors. All authors reviewed and edited the manuscript and approved the final draft.

DECLARATION OF INTERESTS

K.M.W. is an advisor to and holds equity in Ribometrix. All other authors declare no competing interests.

SUPPLEMENTAL INFORMATION

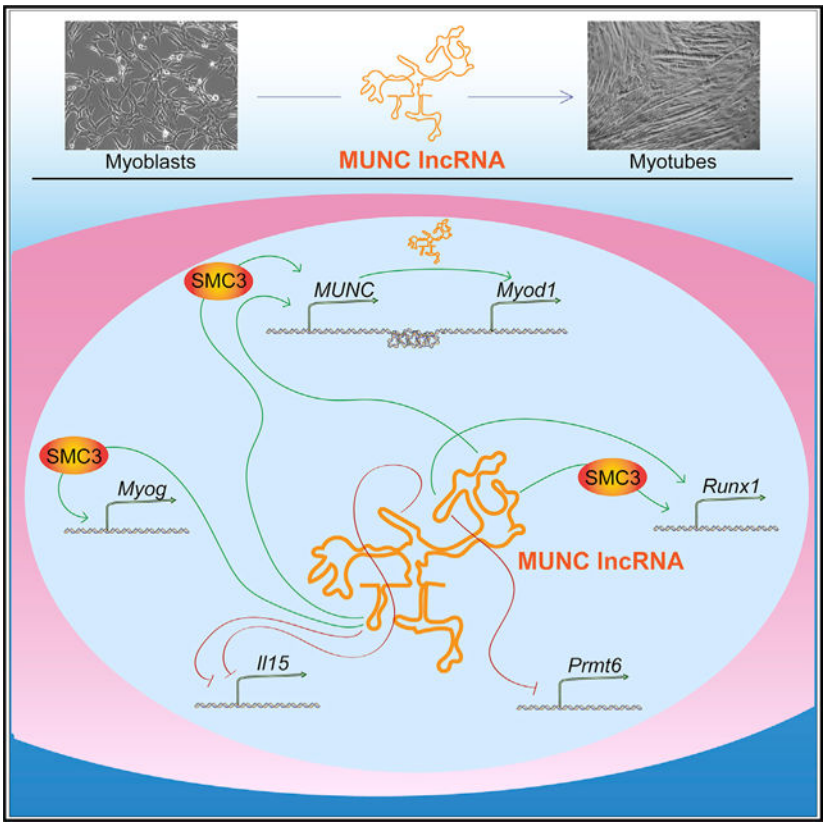
Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110361.

the cohesin complex. *Myod1* induction and cohesin interaction comprise only a subset of *MUNC* phenotype. Our study reveals unexpectedly complex, structure-driven functions for the *MUNC* lncRNA and emphasizes the importance of experimentally determined structures for understanding structure-function relationships in lncRNAs.

In brief

Przanowska et al. identify the secondary structure of *MUNC* lncRNA and determine structural domains responsible for gene expression regulation during skeletal muscle differentiation. This study establishes the power of integrating experimentally driven secondary structure modeling with structure-function analyses to identify functional domains and mechanisms of action of lncRNAs.

Graphical Abstract



INTRODUCTION

In 2004, the International Human Genome Sequencing Consortium found that roughly 70% of the nucleotides in the human genome are transcribed but not translated, thus comprising non-coding RNAs (ncRNAs) (International Human Genome Sequencing Consortium, 2004). Since then, many ncRNAs have been discovered and characterized, and a few functionally characterized, but even fewer have experimentally determined structural information. The majority of long non-coding RNAs (lncRNAs), defined as transcripts longer than 200

nucleotides, share signatures of coding RNAs: they are transcribed by RNA polymerase II, spliced, capped at the 5′ end with 7-methylguanosine, and polyadenylated at the 3′ end (Derrien et al., 2012). lncRNAs are not transient intermediaries to functional proteins, but functional macromolecules that drive cellular programs. It is thus important to understand how intrinsic higher-order lncRNA structures enable their function.

Most published studies have predicted the secondary structures of lncRNAs computationally by leveraging thermodynamic properties of base pairing, base stacking, and other atomic interactions. Although modeling can be accurate for short RNA sequences, accuracy drops significantly as the length of the transcript increases (Miao et al., 2015; Deigan et al., 2009). Data from probing experiments can be used to restrain folding algorithms to yield more accurate structural models, especially for long RNAs (Li et al., 2020a; Deigan et al., 2009; Mustoe et al., 2019). SHAPE-MaP (2′-hydroxyl acylation analyzed by primer extension coupled with mutational profiling) chemical probing strategies have proven especially useful, as nearly every nucleotide is probed in a single experiment, and RNAs of nearly any length can be studied (Merino et al., 2005; Smola et al., 2015a, 2015b; Busan et al., 2019). Chemical probing studies have revealed that individual RNAs have distinct “structural personalities” (Weeks, 2021), a feature that likely applies to lncRNAs and to individual domains within large lncRNAs. High-resolution methods, especially of ribosome (Watson et al., 2020) and viral (Jaafar and Kieft, 2019) systems have revealed enormous complexity in RNA structure and that large RNAs tend to form smaller domains.

The lack of structure-function studies in the lncRNA field is a major limitation for determining the molecular mechanisms by which lncRNAs exert their functions. One of the best-characterized lncRNAs is *Xist* lncRNA, which is a master regulator of X chromosome inactivation (Cerase et al., 2015; Brown et al., 1991). Deletion studies showed that a 5′ conserved repeat region (RepA) of *Xist* is indispensable for gene silencing (Wutz et al., 2002). NMR studies found that a 26-nucleotide fragment that includes the RepA sequence forms a stem-loop structure (Duszczyk et al., 2008). More recent studies have further demonstrated the influence of structure on the functions of lncRNAs: dynamic and flexible structures in *Xist* act as landing pads for proteins (Fang et al., 2015; Smola et al., 2016), *MEG3* pseudoknot structures (“kissing loops”) modulate the p53 response (Uroda et al., 2019), an unstructured region in the *SLNCR1* lncRNA nucleates a non-canonical transcription complex that promotes melanoma invasion (Schmidt et al., 2020), and *GAS5* lncRNA contains three structural domains that independently regulate cell survival under different conditions (Frank et al., 2020).

The *MUNC* lncRNA (also known as ^{DRR}eRNA) plays an important role in myogenesis, the process of skeletal muscle tissue formation (Mueller et al., 2015; Cichewicz et al., 2018; Tsai et al., 2018; Mousavi et al., 2013). *MUNC* is located 5-kb upstream of the *Myod1* transcription start site and has two functional isoforms (Mueller et al., 2015). Both are upregulated in skeletal muscles compared with other tissues and during differentiation of skeletal muscle myoblasts (Mueller et al., 2015). Based on the results of genomic transcriptome profiling (RNA-seq) in differentiating murine myoblasts, the abundance of *MUNC* lncRNA is comparable to that of *Myod1* mRNA. *MUNC* was initially thought to be a classic enhancer RNA (eRNA) that acts to maintain open chromatin and induce

expression of the *Myod1* gene in *cis*. However, *MUNC* depletion using small interfering RNA (siRNA) reduces *Myod1* transcription and myoblast differentiation, and since siRNA acts post-transcriptionally, this effect is inconsistent with a *cis*-acting effect on *Myod1* transcription. Stable overexpression of *MUNC* lncRNA from heterologous loci stimulates the expression of multiple promyogenic RNAs, including *Myod1*, meaning that the *MUNC* lncRNA operates in *trans* (Mueller et al., 2015). *MUNC* overexpression in C2C12 cells that lack the *Myod1* gene induces the expression of other myogenic genes, demonstrating that *MUNC* is capable of regulating genes in a MYOD1-independent manner (Cichewicz et al., 2018). Thus, *MUNC* regulates gene expression in *cis* and in *trans*. It has been suggested that the *trans* functions of *MUNC* are mediated through the recruitment of the cohesin complex to target promoters (Tsai et al., 2018). Here, we describe the first structure-function study of *MUNC* lncRNA. We characterized multiple RNA functional domains in *MUNC* and found that different domains mediate distinct features of *MUNC* promyogenic activity.

RESULTS

MUNC lncRNA promotes promyogenic pathway differentiation

The *MUNC* gene encodes two functional lncRNA isoforms that are approximately equal in abundance (Mueller et al., 2015), which we refer to as spliced and unspliced since they differ by the inclusion of an intron. We stably overexpressed each isoform separately in C2C12 murine myoblasts and cultured the resulting cell lines in proliferating (GM) or differentiating (DM3) conditions for 3 days (Figure 1A). Each isoform induced expression of *Myod1*, *Myog*, and *Myh3,* although expression of the spliced isoform resulted in higher levels of these transcripts (Figure 1B).

RNA-seq analysis reveals that although the two isoforms are both promyogenic, they induce different sets of genes. Hierarchical clustering of the RNA-seq data confirmed that all the cells in proliferating conditions were distinct from all the cells grown in differentiating conditions (Figure S1A). The principal components most responsible for distinguishing proliferating and differentiating expression profiles in control cells are exacerbated upon *MUNC* overexpression, but, intriguingly, the spliced and unspliced isoforms contributed to changes in different axes (Figure 1C). In proliferating conditions, overexpression of either spliced or unspliced *MUNC* pushed cells toward an expression profile more similar to that of differentiated control cells than to control cells grown under proliferating conditions, and cells that overexpressed the spliced isoform had transcriptomes more similar to that of differentiated control cells than did cells that expressed the unspliced isoform. In differentiation medium, when endogenous *MUNC* is expressed, overexpression of the spliced isoform induced more differentiation-related genes in the direction of the scores on the second axis of the principal component analysis (PCA2), whereas the unspliced isoform had a lesser effect in the direction of the scores on the first axis of the principal component analysis (PCA1) (Figure 1C). *MUNC* overexpression altered expression of many genes, although fewer genes responded to *MUNC* overexpression than when cells were switched to differentiating conditions (Figure S1B). Thus, *MUNC* overexpression promotes a pro-differentiation gene expression profile, and this phenotypic shift can even be observed in C2C12 cells in proliferating (non-differentiating) conditions.

To identify the transcription factors that regulate the genes altered by *MUNC* overexpression, we employed binding analysis for regulation of transcription (BART), which identifies transcription factors enriched in a set of promoters relative to the factors' genome-wide binding site distribution (Wang et al., 2018). We first identified the 20 transcription factors most highly activated during normal differentiation (Figure 1D). Nineteen of these regulators were also activated by overexpression of at least one *MUNC* isoform in proliferating cells, and 15 were activated by both (Figure 1D). The top 10 predicted transcription factors are known to contribute to myogenesis (Figure S1C and S1D). The cohesin complex is involved in regulation of genes upregulated by the *MUNC* spliced isoform (Figure S1C), supporting the hypothesis that *MUNC* recruits this complex (Tsai et al., 2018).

The *MUNC* isoforms produced distinct changes in gene expression when overexpressed in proliferating cells. Gene Set Enrichment Analysis indicated that the genes upregulated by the spliced isoform in proliferating cells are involved in myogenesis, similar to the enrichments observed between differentiating and proliferating control cells (Figure 1E). Interestingly, overexpression of the unspliced *MUNC* isoform in proliferating cells did not result in significant enrichment for genes associated with myogenesis (Figure 1E), suggesting that either the spliced is simply better than the unspliced isoform at inducing promyogenic genes or that the two isoforms regulate distinct sets of genes. In support of the latter hypothesis, 645 genes were upregulated by both isoforms of *MUNC* (e.g., *Myod1*, *Myog*, and *Myh3*; Figure 1B), whereas there were 2,730 genes upregulated by one isoform and not the other (Figures 1F and S1E). Gene ontology (GO) analysis revealed that both non-overlapping and overlapping upregulated gene sets are involved in muscle-related pathways (Figure 1G) and that the genes upregulated by the expression of the spliced isoform were more significantly enriched for muscle-related pathways than those upregulated by expression of the unspliced isoform. The same patterns were observed for genes downregulated in proliferating cells and for genes up- and downregulated by *MUNC* overexpression in differentiating conditions (Figure S1E). In summary, both isoforms of the *MUNC* lncRNA induce expression of promyogenic genes, likely through activation of a common set of transcription factors. The genes regulated by the two isoforms are distinct, and the spliced isoform has a more pronounced effect on promyogenic and differentiation gene expression profiles than does the unspliced isoform.

SHAPE-MaP of *MUNC* reveals distinct secondary structures

To determine the structural features of *MUNC* lncRNAs and to find the structural modules important for its promyogenic activity, we analyzed the structures of the ectopic spliced and unspliced isoforms using SHAPE-MaP chemical probing (Smola et al., 2015a, 2015b; Weeks, 2021). The RNAs were probed using the SHAPE reagent 5-nitroisatoic anhydride (5NIA) in C2C12 cells (in cell) and after gentle extraction from C2C12 cells (cell free). SHAPE measures local nucleotide flexibility; thus, unpaired nucleotides are preferentially acylated at their 2′-hydroxyl groups. SHAPE-modified nucleotides are identified as mutations and short deletions in cDNAs created during relaxed fidelity MaP reverse transcription. The resulting *MUNC* mutation profiles enabled us to model the secondary structures of the spliced and unspliced *MUNC* isoforms.

The SHAPE analysis of the spliced isoform identified multiple regions likely to form well-defined local domains (Figures 2A and 2B). The SHAPE-supported model differs substantially from previously reported *MUNC* lncRNA structures predicted using bioinformatic algorithms (Cichewicz et al., 2018). SHAPE reactivities from two independent experiments performed 3 years apart showed good agreement (Pearson's R = 0.94; Figure S2A) and yielded similar pairing probabilities (Figure S2B). In-cell and cell-free SHAPE data for the *MUNC* spliced isoform are highly correlated (Pearson's R = 0.92; Figure S2C), suggesting that the in-cell structure is similar to that of the cell-free RNA. We also identified nucleotides with significant in-cell protection from or enhancement of modification relative to the cell-free structure (Figures 2B and S2D), highlighting regions of potential protein interactions or other changes in cells.

Comparisons of SHAPE reactivity profiles (Figure S3A) and base-pairing probabilities (Figure S3B) for the spliced isoform (518 nt; exon 1 and exon 2) and the unspliced isoform (1,083 nt; exon 1, intron, and exon 2) indicate that the isoforms share six structurally homologous domains (referred to as common hairpin [CH] domains; Figures 2B and S3C). CH1 and CH6 domains are at the 5′ and 3′ ends of the transcripts, respectively, whereas CH2 is close to the 3′ end of exon 1. We also discovered regions of well-defined structure unique to the spliced isoform (Figure 2B): one hairpin (SH1), two in-cell protected regions (PR1 and PR2), and one well-defined loop (L1). All of the selected structures (with the exception of CH2) are contained within low-SHAPE low-Shannon regions (51 nt median window, <0.4 SHAPE, <0.06 entropy), with SH1 being the most strongly low-SHAPE low-Shannon region.

Distinct structural domains of the spliced isoform of *MUNC* regulate myogenesis

We next tested how disruption of structural domains in the spliced isoform of *MUNC* identified using SHAPE-based modeling affected the expression of promyogenic factors. A series of variants lacking specific structural domains or containing defined mutations were overexpressed in C2C12 cells (Figures 3A and S4A–4C). The wild-type isoform induced production of *Myod1*, *Myog*, and *Myh3* mRNAs in proliferating cells. Deletion or mutation of all tested sites decreased *Myod1* induction, and ΔCH1 and ΔCH4 deletions completely inhibited *Myod1* induction (Figure 3B). *Myog* induction required the CH1, CH2, CH5, and CH6 domains and wild-type PR2 site (Figure 3C). The motifs dispensable for *Myog* induction were dispensable for *Myh3* induction as well, with the exception of CH4 (Figure 3D). Notably, the ΔCH1 mutant did not induce any of the three promyogenic factors (Figures 3A–3D), suggesting that CH1 plays a critical role in activity of the spliced isoform of *MUNC*. Overexpression of the wild-type spliced isoform in differentiating cells also induced the three promyogenic factors, and ΔCH1 and ΔCH4 impaired in induction of all three promyogenic factors in differentiating cells (Figures S4D–S4F).

The mutations that disrupted CH4, CH3, and L1 impaired *Myod1* induction but promoted expression of *Myog*, suggesting that the lncRNA activates *Myog* independently of *Myod1* in proliferating cells (Figure 3E). Different *MUNC* domains also influenced *Myh3* and *Myod1* expression (Figure 3F). Mutations in domains of spliced *MUNC* affected *Myh3* and *Myog* expression similarly, except ΔCH4, which had a deleterious effect on *Myod1* and *Myh3*

expression but not *Myog* expression (Figure 3G). These correlations suggest that *MUNC* induces *Myog* independently of *Myod1*. whetreas *Myh3* induction is mostly dependent on *Myog* with some additional dependence on *Myod1*.

In some cases, functional domains can cooperate when supplied either in *cis* or in *trans* (Uroda et al., 2019), so we tested whether ΔCH1 and ΔCH4 variants complement each other without being physically connected. Co-overexpression of ΔCH1 and ΔCH4 in *trans* did not rescue defects in *Myod1* or *Myh3* induction in proliferating (Figures 3H–3K) or differentiating cells (Figures S4G–S4J). We conclude that the local proximity of CH1 and CH4 elements on the same RNA molecule is important for induction of expression of *Myod1* and of *Myh3*.

Overexpression of *MUNC* leads to phenotypical changes in C2C12 cells

Overexpression of the spliced isoform of *MUNC* induced production of MYOD1 protein in proliferating conditions and differentiation conditions (Figures 4A, 4B, S5A, and S5B). Phenotypical changes and an increase in the frequency of myosin heavy-chain (MHC)-positive cells were also observed upon overexpression of the spliced isoform of the lncRNA (Figures 4C and 4D). Induction of MYOD1 protein was not observed in proliferating cells when the *MUNC* construct lacked CH1, CH4, or CH5 domains (Figure 4B). In differentiation conditions, ΔCH1 and ΔCH4 did not induce MYOD1 protein expression, although other constructs did (Figures S5A and S5B). Although *Myog* and *Myh3* mRNA are induced by expression of the wild-type spliced isoform in proliferating conditions, MYOG and MHC protein expression required differentiation.

Given that protein levels are affected by *MUNC* spliced overexpression in a domain-specific manner, we tested C2C12 differentiation efficiency as measured by percentage of MHC-positive cells upon overexpression of the various *MUNC* constructs. Expression of the wild-type spliced *MUNC* isoform increased the percentage of MHC-positive cells at 3 days of differentiation (Figures 4C and 4D). The effect of *MUNC* lncRNA on differentiation was especially dependent on CH1 and CH4 domains (Figures 4C and 4D).

MUNC transcripts lacking CH1 and CH4 are structurally similar to wild type

When structure-function studies by making targeted deletions in RNAs is being performed, it is important to assess whether local sequence changes alter the RNA structure. We performed cell-free SHAPE-MaP on wild-type *MUNC*-spliced, ΔCH1, and ΔCH4 constructs to assess whether these mutants retained the overall structure of the native lncRNA. The SHAPE-informed secondary structure models and reactivity data of ΔCH1 and ΔCH4 were highly similar to each other and to the wild-type spliced isoform (Figures 5A–5C). The reactivity data for these two constructs were highly correlated with the data for the wild-type spliced isoform (Pearson's R values 0.93 and 0.94, respectively; Figures 5D and 5E). The structures of all retained structural domains and conserved helices were preserved in both ΔCH1 and ΔCH4 mutants, supporting that global lncRNA structure is maintained. These data further support the overall accuracy of our structural models, as specific, well-defined motifs can be deleted without affecting the global structure. We conclude that CH1 and CH4 domains are critical for the observed promyogenic phenotypes.

MUNC binds to specific genomic sites and regulates expression of adjoining genes in a domain-specific manner

We next investigated whether *MUNC* physically associates with the genes it regulates (Figure S6A). Using the chromatin isolation by RNA purification (ChIRP) assay, we identified *MUNC*-binding sites at 410 genomic loci. More than 90% of identified targets overlapped with previously published results (Tsai et al., 2018) (Figure S6A), and the locations of binding sites relative to transcription start sites of nearby genes were very similar between the two datasets (Figures S6B and S6C). To understand whether *MUNC* lncRNA binding to chromatin affects gene expression regulation, we compared our ChIRP-seq targets with RNA-seq profiles of wild-type cells in proliferating and differentiating conditions with those of cells that overexpress *MUNC* and with expression microarray data from cells where *MUNC* RNA had been depleted (Mueller et al., 2015). Hierarchical clustering demonstrated that the expression profiles of *MUNC*-targeted genes more closely resemble the expression in proliferating conditions than differentiating conditions when *MUNC* is depleted (Figure S6D). This confirmed that *MUNC* is required for expression changes during normal myogenesis. In support of the hypothesis that *MUNC* binding to a promoter region directly regulates expression of the downstream gene, overexpression of *MUNC* in proliferating cells results in expression profiles of *MUNC*-bound genes, similar to the profile in differentiating conditions (Figure S6E). These results support *MUNC*'s role as a promyogenic factor and that the *MUNC*-binding sites are functional and important for skeletal muscle differentiation.

To establish which of the *MUNC* targets are important for myogenesis, we separately evaluated genes activated by and genes repressed by *MUNC*. The *MUNC*-activated genes were defined as *MUNC*-bound genes that are induced during normal differentiation, repressed by depletion of *MUNC*, and induced by *MUNC* overexpression. The *MUNC*-repressed genes were defined as *MUNC*-bound genes that are repressed during normal differentiation, induced by *MUNC* depletion, and repressed by *MUNC* overexpression. There were 37 *MUNC*-activated genes and 22 *MUNC*-repressed genes that were regulated in a concordant manner during myogenesis (Figure S6A). ChIRP-qPCR for 10 of them showed that *MUNC* binding to the sites adjoining these genes was significantly increased after *MUNC* overexpression, and RT-qPCR confirmed significant up- or downregulation of these genes (Figure S6F).

We next evaluated the effects of *MUNC* mutants on binding to the promoter regions and on gene expression (FigureS6A, S7A, and S7B). In particular, the CH1 domain is critically important for *MUNC* binding and activation of *Myod1, Myog*, and *Dcn* and repression of *Fnbpl1* and *Il15*. The CH5 domain is important for *MUNC*-mediated regulation of *Myod1, Myog, Dcn, Met*, and *Runx1*. Although all domains were required for *Myod1* induction by *MUNC* (Figure 2B), specific domains were important for binding to the two *MUNC*-binding sites in the *Myod1* promoter region. CH4, PR1, and L1 are important for binding to one site, and CH2, SH1, and PR2 are important for binding to the other; CH1, CH5, and CH6 are required at both sites. In another example, PR2 is required for *MUNC* to bind to and downregulate the *Fam137b* gene, whereas CH1 is required for downregulation of *Fnbpl1*. Thus, different *MUNC* structural domains regulate different sets of target genes. In addition,

binding to and regulation of some genes require the function of multiple domains. This argues against a common mechanism of *MUNC* function on all target genes.

To disentangle the roles of *MUNC* sequence and structure, we designed four independent mutants (with compensatory mutations where necessary) each for the CH1 and CH4 domain (Figures 6B and 6C, accordingly) to change the sequences of these domains without disrupting the structure. These mutants were overexpressed in C2C12 cells (Figure S7C). For the CH1 domain there was no specific sequence important for *Myod1* or *Myh3* promyogenic factors, but we discovered one mutation (CH1mutD) in the terminal loop that affected the ability of *MUNC* to induce *Myog* (Figure 6B). This second separation-of-function mutant (the first being ΔCH4) further suggested that *MUNC* regulates *Myog* independently of *Myod1*. We also observed partial inhibition of *Dcn* and *Met* induction by *MUNC* in the same mutant (Figure S7D). The fact that we observed an effect in only one of four compensatory mutants and only on some genes but not others suggests that the structure plays a major role with some input from the sequence on specific target genes. For the CH4 domain we found no effect of sequence on the induction of the three promyogenic factors (Figure 6C) or regulation of *MUNC* target genes (Figure S7E).

To further explore sequence influence on the *MUNC* phenotype, we performed motif enrichment analysis using MEME (Bailey and Elkan, 1994) and found 12 significant motifs present in 76.17% of *MUNC* ChIRP-seq peaks. Next, we used FIMO (Grant et al., 2011) to identify significant motifs present in studied *MUNC* domains. We identified 4 significant motifs present in *MUNC* and 31.25% of *MUNC* ChIRP-seq peaks: 1 motif in CH1 domain (5.86% of peaks), one motif in CH5 domain (6.64% of peaks, including *Myod1*), and two motifs in CH6 domain (23.05% of peaks, including *Myod1* and *Fam173b*) (Figure S7E). However, mutations in the motif in CH1 (CH1mutA, CH1mutB, CH1mutC, and beginning of CH1mutD) had no effect on the promyogenic genes and *MUNC* target genes. Therefore, the sequence motif in CH1 is not responsible for *MUNC* function, and there was no similar motif in the other domain studied in detail, CH4. On the basis of the results from eight independent sequence mutants, we conclude that *MUNC* function attributed to CH1 and CH4 is primarily driven by structure, with the only dependence on sequence seen in the role of the terminal loop in CH1 for *Myog* induction.

MUNC is known to interact with the cohesin complex subunits including structural maintenance of chromosomes 3 (SMC3) to regulate gene expression (Tsai et al., 2018). RNA-immunoprecipitation with an anti-SMC3 antibody and RIP-qPCR of *MUNC* demonstrated that deletions of CH1, CH2, CH3, CH5, CH6, or SH1, but not CH4, PR1, or PR2, significantly decreased *MUNC* binding to SMC3 (Figure 6G).

We also noted that approximately equal numbers of genes adjoining *MUNC*-binding sites are up- or downregulated after *MUNC* overexpression and that the genes are distributed over all chromosomes and not clustered around the *MUNC* locus (Figure 7A). Together all these data show that *MUNC* can activate or repress genes that are proximate to sites where it is bound both in *cis* and in *trans* through distinct domains. We also proved that multiple mechanisms may be important for DNA/protein binding and gene regulation by *MUNC* (Figures 7B–7E). SMC3 binds in the promoter regions of *Myod1, Myog*, and *Runx1*

(Tsai et al., 2018); however, *MUNC*-SMC3 binding is not sufficient for all *MUNC*-mediated gene expression regulation. For example, mutation of the PR2 domain did not change *MUNC*-SMC3 binding but decreased induction of *Myod1* and *Myog* by *MUNC* (Figure 7F). Conversely, although *MUNC*-SMC3 binding was diminished upon deletion of SH1 or CH3, *MUNC* could still induce *Myog* (Figure 7G) and *Myod1* and *Myog* and *Runx1* (Figure 7H), respectively. Thus, in some cases *MUNC*-SMC3 binding may not be sufficient or required for regulating gene expression by *MUNC*.

DISCUSSION

Analyses of protein structure to define domains responsible for specific functions have proved essential for understanding complex functions. Analogous structure-directed functional studies for lncRNAs are rare. Here, we performed a structure-function analysis of the two isoforms of the *MUNC* lncRNA, which function during differentiation of skeletal muscle myoblasts (Mueller et al., 2015). Critically, the secondary structures determined based on modeling directed by SHAPE-MaP data differed substantially from those predicted using computational-only RNA folding algorithms (Cichewicz et al., 2018; Tsai et al., 2018), emphasizing the need for experimental data-directed structural studies and modeling of lncRNAs (Weeks, 2021).

We performed a structure-guided analysis of the importance of domains that are common to the two *MUNC* isoforms and that are unique to the spliced isoform. The mechanism of action of *MUNC* lncRNA is more complex than expected for an eRNA, as we showed that different RNA domains of *MUNC* are important for binding to and regulating different genes. Importantly, we showed that the deletion of CH1 and CH4 domains were surgical and did not alter the overall structure of the lncRNA. We do not know if the same is true for the other deletions, but the persistence of many functions with these deletions suggests that they, too, did not disrupt the overall structure of *MUNC*.

Both the spliced isoform and the unspliced isoform promote myogenesis and share multiple well-defined structural motifs but also contain isoform-specific structural domains. Although the two isoforms differ by the presence of a single intron, they clearly have non-identical effects. The two isoforms do activate a common subset of genes, as well as a larger, non-overlapping set of genes, and both sets of genes are enriched in muscle related pathways. The spliced transcript is also a much stronger promyogenic factor than the unspliced isoform. Multiple factors might influence the two isoforms' divergent behaviors. First, the spliced isoform is more stable structurally: it has a larger fraction of well-determined (low entropy) structures than the unspliced transcript. Second, different sites on the two isoforms may be bound by varying cellular factors: the relative changes in reactivity profiles between in-cell and cell-free conditions, a measure of interactions with cellular factors, are much higher for the spliced than for the unspliced isoform. Because most structural domains present in the spliced isoform also exist within the unspliced one, we hypothesize that the intron in the unspliced transcript binds to additional factors that allow it to target different genes. Last, the included intron in the unspliced isoform may contribute more to *Myod1* transcription in *cis* when functioning as an eRNA, while the more stably folded regions that assemble after *MUNC* splicing could regulate other targets in *trans*.

To evaluate the mechanisms by which the spliced isoform of *MUNC* regulates gene expression, we focused on *Myod1*, *Myog*, and *Myh3*, which encode promyogenic transcription factors. We previously reported that *MUNC* overexpression leads to increases in levels of these mRNAs (Mueller et al., 2015; Cichewicz et al., 2018). By careful comparison of effects of different *MUNC* mutants, we showed that *MUNC* directly regulates expression of all three mRNAs, although its effect on *Myh3* appears to be *Myog*- and *Myod1*-dependent in proliferating murine myoblasts. This observation confirms *MUNC* function is not limited to the secondary effects of MYOD1 expression resulting from *MUNC* enhancer RNA activity.

To interrogate whether different regions of *MUNC* act independently or whether they must be in *cis*, we focused on CH1 and CH4. These two domains are the most important of the domains tested for *Myod1* induction, and deletion of either did not affect the structure of the other as shown by SHAPE-MaP analyses. The mutants with deletions of CH1 and of CH4 did not complement each other in *trans* to rescue *Myod1* induction. Therefore, CH1 and CH4 motifs must exist on the same molecule to stimulate *Myod1* expression in *trans*.

Our structural and functional studies revealed that *MUNC* has multiple effector domains and does not regulate expression of all its target genes by a common mechanism. Because *MUNC* could be recruiting multiple transcription complexes, we focused on genes that are directly regulated by *MUNC* binding near their promoter. We used ChIRP to determine the sites in the genome where *MUNC* binds and looked for cases where *MUNC* also regulates adjoining gene expression to find potentially direct targets of *MUNC*. Distinct combinations of *MUNC* domains are required for binding to and induction of *Dcn* (CH1 and CH5), *Met* (CH2 and CH5), and *Runx1* (CH5 and SH1), genes that are known to stimulate muscle differentiation (Li et al., 2020b; Umansky et al., 2015; Wang et al., 2005; Kishioka et al., 2008). Similarly, different domains were critical for downregulation of *Il15* (CH1, CH3, and PR2), *Prmt6* (SH1), and *Grip1* (CH4), genes that are known to have a negative impact on skeletal muscles (Quinn et al., 1995; Choi et al., 2019; Wu et al., 2005). Thus, the multi-modal mechanisms of regulation of gene expression by *MUNC* involve different domains of *MUNC* and likely different protein partners. *MUNC* recruits varying cooperative machinery depending on the genomic context, and this behavior is likely exemplary of other lncRNAs.

In conclusion, this study established the power of integrating experimentally driven secondary structure modeling with structure-function analyses to identify functional domains and mechanisms of action of lncRNAs. Indeed, all well-determined domains chosen for validation based on our models proved to be functional. There is now a growing body of evidence that suggests that well-determined (low SHAPE, low entropy) motifs tend to correlate with functional motifs in large RNAs (Weeks, 2021; Boerneke et al., 2019). Our study clarifies the role of RNA structure in *MUNC* function. *MUNC* was initially described as a *cis*-acting eRNA for *Myod1* and was soon thereafter shown to have *trans*-acting functions beyond those expected for a classical eRNA (Mueller et al., 2015; Cichewicz et al., 2018; Tsai et al., 2018). We show that this underlying functional complexity of *MUNC* is integrated with its RNA structural complexity. Our results confirm *MUNC* regulation of *Myod1* and complex regulation of other genomic targets in *trans*. In

addition, our study shows that binding of SMC3, and by association cohesin, is not sufficient for *MUNC* function and may not be essential for regulating all of the target genes such as *Myod1*, *Myog*, and *Runx1*.

The compact and highly organized structure of *MUNC* lncRNA differentiates it from longer lncRNAs like *Xist* and *HOTAIR,* that contain long, unstructured, repetitive sequence regions that appear to function as landing pads for protein multimerization (Wang et al., 2017; Smola et al., 2016). *MUNC* instead appears to have scaffolding functions more similar to those of *ANRIL* (Zhang et al., 2020). *MUNC* may act as an eRNA in *cis* to recruit a regulatory protein complex (like cohesin) as it does for *Myod1* or in *trans* as demonstrated for *Myog* and *Runx1*. *MUNC* binding can also inhibit some target gene transcription as for *Il15*, *Prmt6*, and *Grip1* by binding to their promoter region and diminishing activity of a transcriptional regulatory factor; in these cases, *MUNC* may act as a decoy or may directly occlude a binding site. Last, since *MUNC* is encoded 5 kb upstream of the transcription start site of *Myod1*, the transcription of *MUNC* itself may positively regulate the transcription of *Myod1* and other loci in close three-dimensional proximity by maintaining active chromatin structure. *MUNC* may also play a role in stabilizing the genome organization and control the spreading of post-translational modifications to nearby chromatin. Regardless, experimentally derived structure models are essential for rapid characterization of functionally important motifs within RNAs of interest.

Limitations of the study

In our study, we have defined *MUNC* secondary structure based on SHAPE-MaP and the functionality of different structural domains. A limitation is the current lack of information on the tertiary structure of *MUNC* that would allow further functional evaluation. We have performed SHAPE-MaP on only two mutants, each lacking a domain important for promyogenic phenotype. This work clearly emphasizes the value of structure-informed functional studie and, to understand lack of binding to specific genomic sites followed by alteration of gene regulation, further mutants should be studied. Likewise, it would be useful to perform ChIRP-seq on all mutants to see changes in genome-wide binding of *MUNC* upon introduction of specific mutations. Additionally, we focused on *MUNC* interactions with only one protein, SMC3, a subunit of cohesin complex. Other protein partners of *MUNC* are known (Tsai et al., 2018). Now, in light of the available secondary structure for *MUNC* and information about nucleotides protected from reactivity in cells, it would be possible to precisely determine protein-binding motifs within *MUNC*. The ribonucleoprotein networks analyzed by mutational profiling (RNP-MaP) method could be used to map cooperative interactions among multiple proteins bound to *MUNC* at nucleotide resolution (Weidmann et al., 2021). Finally, this study did not explore intra- and intermolecular RNA-RNA interactions of *MUNC* that would inform additional insights into the mechanism of action of this intriguing lncRNA.

STAR⋆METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Anindya Dutta (duttaa@uab.edu).

Materials availability—Plasmids and cell lines generated in this study will be made available upon request. We may require a payment and/or a completed materials transfer agreement in case there is potential for commercial application.

Data and code availability

- RNA-seq, ChIRP-seq and SHAPE-MaP data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

For proliferating conditions, C2C12 cells (female murine myoblasts) were cultured in DMEM-high glucose medium (GE Healthcare Life Sciences co.) with 20% fetal bovine serum (Gibco). For differentiation, the serum was 2% horse serum (GE Healthcare Life Sciences co.).

METHOD DETAILS

RNA-seq—RNA samples were isolated from proliferating (GM) or differentiating (DM3) control cells or cells that overexpress a *MUNC* construct by TRIzol extraction using Direct-zol RNA MiniPrep Plus Kit including DNase treatment. RNA-seq was performed by Hudson Alpha on poly(A)-enriched RNA using the Illumina HiSeq 2500 instrument. RNA-seq data was aligned to the mouse assembly GRCm38/mm10 using STAR v2.5 (Dobin et al., 2013) and quantified by HTSeq (Anders et al., 2015). DESeq2 R package (Love et al., 2014) was then applied to identify differentially expressed genes with a adjusted p of <0.05. Bioinformatic prediction for functional factors (including transcription factors and chromatin regulators) that bind at *cis*-regulatory regions was performed using BART 2.0 (Wang et al., 2018). Gene set enrichment analysis was performed as previously described (Subramanian et al., 2005). Gene Ontology was performed using GeneTrail2 (Stöckel et al., 2016). All RNA-seq library data files are available under GEO accession number GSE174203 as a part of the SuperSeries GSE174218.

SHAPE-MaP

Cell-free SHAPE: Control or *MUNC* construct overexpressing C2C12 cells were grown to 70% confluency in two 15-cm dishes. Both plates were washed once in PBS before scraping and lysis in 2.5 mL of proteinase K buffer (40 mM Tris, pH 8, 200 mM NaCl, 1.5% sodium dodecyl sulfate, and 0.5 mg/mL proteinase K). Proteins were digested for 45

min at 23°C with intermittent mixing. Nucleic acids were extracted twice with 1 volume of phenol:chloroform:isoamyl alcohol (25:24:1) that was pre-equilibrated with 1.1× RNA folding buffer (110 mM HEPES, pH 8, 110 mM NaCl, 5.5 mM $MgCl_2$). Excess phenol was removed through two subsequent extractions with 1 volume chloroform. The final aqueous layer was buffer exchanged into 1.1 × RNA folding buffer using PD-10 desalting columns (GE Healthcare Life Sciences). The resulting RNA solution was incubated at 37°C for 20 min before being split into two equal volumes. The SHAPE reagent, 250 mM 5NIA (AstaTech) in DMSO was added to one half, and DMSO was added to the other. Samples were incubated at 37°C for 10 min. RNA was precipitated with 1/10 volume of 2 M NH_4OAc and 1 volume of isopropanol. After one wash with 75% ethanol, the resulting pellet was dried and resuspended in 88 μL of water and 10 μL of 10× TURBO DNase buffer and 4 units of TURBO DNase (Thermo Fisher) were added. The mixture was incubated at 37°C for 1 h. RNA was purified (GeneJET RNA Cleanup and Concentration Micro Kit, Fisher) and eluted into 20 μL of nuclease-free water.

In-cell SHAPE: Control C2C12 cells or cells expressing *MUNC* constructs were grown to 70% confluency in four wells of a 6-well plate. After washing with PBS, 900 μL of standard growth medium was added. Next, 100 μL of 250 mM 5NIA was added to two wells and 100 μL of DMSO (control) were added to the other two wells. Plates were incubated at 37°C for 10 min. Media was aspirated, cells were washed once with PBS, and total RNA was extracted using TRIzol (Thermo Fisher). RNA pellets were dried and resuspended in 88 μL nuclease-free water, treated with TURBO DNase, and purified with GeneJET RNA Cleanup and Concentration Micro Kit as described for the cell-free experiment.

MaP reverse transcription: 1 μg of each RNA sample was subjected to MaP reverse transcription, which requires Superscript II and addition of betaine and Mn^{2+} to the RT buffer (Siegfried et al., 2014; Smola et al., 2015b), using a *MUNC*-specific reverse primer (Table S1). The cDNA generated was buffer exchanged over Illustra microspin G-50 columns (GE Healthcare). For second-strand cDNA synthesis, output DNA (corresponding to 167 ng of total RNA) was used as a template for 25 μL PCR reactions (Q5 Hot-start polymerase, NEB) with primers made to amplify spliced (1–518 bp) and unspliced (1–792 and 584–1083 bp) *MUNC* isoforms. Reactions included 1x Q5 reaction buffer, 250 nM each primer, 100 μM dNTPs, 0.02 units/μL Q5 Hot-start polymerase. PCR was conducted as follows: 98°C for 30 s, then 25 cycles of 98°C for 10 s, 60°C (*MUNC* spliced) or 69°C (*MUNC* unspliced) for 30 s, and 72°C for 35 s, followed by 72°C for 2 min. Step 1 PCR products were run on a 2% gel and purified using Zymoclean Gel DNA Recovery Kit and eluted in 10 μL of nuclease-free water. Purified PCR products were measured using a Qubit dsDNA HS Assay Kit, and 1 ng was used for tagmentation using Nextera XT DNA Library Preparation Kit (Illumina). After neutralization with NT buffer, multiplex indices were added using the Nextera XT DNA Library Preparation Kit (Illumina). PCR was performed as follows: 72°C for 3 min, 95°C for 30 s, then 12 cycles of 95°C for 10 s, 55°C for 30 s, 72°C 30 s, and, finally, 72°C for 5 min. Step 2 PCR products were purified using a 0.8x ratio of Agencourt AMPure XP beads (Beckman Coulter) and eluted in 20 μL of nuclease-free water.

Sequencing of MaP libraries: Size distributions and purities of fragmented *MUNC* amplicon libraries were verified (2100 Bioanalyzer, Agilent). Libraries (about 120 amol of each) were sequenced on a MiSeq instrument (Illumina) with 2 × 250 or 2 × 300 paired-end sequencing. Libraries derived from total cytoplasmic RNA were sequenced with 2 × 300 paired-end sequencing on a MiSeq instrument, combining reads from multiple runs until desired RNA sequencing depth was achieved. All SHAPE-MaP libraries data files are available under GEO accession number GSE174217 as a part of the SuperSeries GSE174218.

Mutation counting and SHAPE profile generation with ShapeMapper 2 software: FASTQ files from sequencing runs were directly input into the ShapeMapper 2 software (Busan and Weeks, 2018) for read alignment and mutation counting. To ensure that mutation rates were not affected by reduced fidelity at reverse transcription initiation sites, target FASTA files input to ShapeMapper 2 had primer-overlapping sequences and the first 5 nucleotides transcribed were set to lowercase, which eliminates these positions from analysis. ShapeMapper 2 was run with the –min-depth 4000 flag and all other values set to defaults. In each experiment, the 5NIA-treated samples were designated as the “modified” samples and DMSO-treated samples as “unmodified” samples.

***Modeling* MUNC *structure using SuperFold*:** The SuperFold analysis software (Smola et al., 2015b) was used with experimental SHAPE data to inform RNA structure modeling by RNAStructure (Reuter and Mathews, 2010). Default parameters were used to generate base-pairing probabilities for all nucleotides (with a max pairing distance of 600 nt) and minimum free energy structure models.

***Identification of in-cell changes in* MUNC *spliced SHAPE reactivity*:** SHAPE reactivities of in-cell and cell-free treated RNAs were normalized to each other using a median difference minimization strategy to improve normalization between in-cell and cell-free samples and enable sensitivity to single-nucleotide level differences. First, the log relative reactivities for each dataset were calculated as follows:

$$\mathrm{LOGREL}_{NT} = \log_{10}\left(\frac{5NIA\ \text{mutation rate}_{\mathrm{NT}}}{DMSO\ \text{mutation rate}_{\mathrm{NT}}}\right)$$

$$\text{Initial scaling factor} = \mathrm{MEDIAN}(\mathrm{LOGREL}_{\mathrm{cell-free}}) - \mathrm{MEDIAN}(\mathrm{LOGREL}_{\mathrm{in-cell}})$$

The $\mathrm{LOGREL}_{\mathrm{in\text{-}cell}}$ values were adjusted up by the initial scaling factor, and differences were calculated for each nucleotide:

$$\mathrm{Diff}_{NT} = \mathrm{LOGREL}_{\mathrm{NT-cell-free}} - \mathrm{LOGREL}_{\mathrm{NT-in-cell}}$$

The final scaling factor (added to in-cell LOGREL values) was calculated as the value that minimizes the median for all nucleotides of $|\mathrm{Diff}_{\mathrm{NT}}|$. New $\mathrm{Diff}_{\mathrm{NT}}$ values were computed with the final scaling factor, and Z-scores were computed for each nucleotide:

$$\text{Zscore} = \frac{\text{Diff}_{\text{NT}} - \overline{\text{Diff}_{\text{NTs}}}}{\sigma_{\text{NTs}}}$$

Only nucleotides with Z-scores > 1.645 standard deviations from the mean (90% confidence intervals) in both replicates were considered significant shifts in SHAPE reactivity.

Graphical display of SHAPE data: Secondary structure projection images were generated using the (VARNA) visualization applet for RNA (Darty et al., 2009).

***Structural analyses of* MUNC *mutants*:** Cells overexpressing the wild-type spliced *MUNC* isoform and ΔCH1 and ΔCH4 mutants were used for cell-free SHAPE as described above. MaP reverse transcription was performed using a *MUNC*-specific reverse primer. For second-strand cDNA synthesis we used the same reverse primer for all three *MUNC* constructs but a different forward primer was used for ΔCH1 that overlapped the deletion junction. All the further steps were performed as described above.

Cell lines generation—The pLPCX plasmid carrying the sequence of spliced *MUNC* (Mueller et al., 2015) was used as a template to obtain mutants via PCR followed by In-Fusion cloning. The constructs were linearized and transfected into the C2C12 cells using Lipofectamine 3000 (Life Technologies). After 24 h, pools of stably transfected cells were selected with 2 μg/ml puromycin. The procedure was repeated at least twice for each construct to ensure that the observed phenotype was not an effect of random selection of a less or more differentiation-potent population. For complementation experiments ΔCH4 was cloned into pLHCX vector. At 24 h after transfection, pools of stably transfected cells were selected with 300 μg/ml hygromycin. Oligonucleotide sequences are listed in Table S1.

RNA isolation and RT-PCR—RNA was isolated by TRIzol extraction using Direct-zol RNA MiniPrep Plus Kit including DNase treatment. cDNA synthesis for mRNA expression levels measurement was performed using GoScript RT cDNA synthesis kit (VWR) with random hexamer priming. After cDNA synthesis, qPCR was performed with StepOnePlus™ Real-Time PCR System using PowerUp™ SYBR® Green Master Mix (Thermo Fisher). All primers used in this study are listed in Table S1.

Western blotting—Cells were lysed in IPH buffer (50 mM Tris-Cl, 0.5% NP-40%, 50 mM EDTA). Samples were run on a 10% polyacrylamide SDS-PAGE gel and transferred to nitrocellulose membranes. Membranes were blocked for 30 min in 5% milk containing PBST and incubated overnight with primary antibody in 1% milk. After washing, secondary antibody incubation was carried out for 1 h at 1:4000 dilution before washing and incubation with Millipore Immobilon HRP substrate. Antibodies used were as follows: MYOD1 sc-32758 (Santa Cruz Biotechnology), MHC 22287–1-AP (Proteintech), MYOGENIN sc-12732 (Santa Cruz Biotechnology), and HSP90 sc-13119 (Santa Cruz Biotechnology).

Immunofluorescence assay—Cells were plated on glass coverslips and collected after 3 days of differentiation. The coverslips were fixed with 4% paraformaldehyde in PBS for 15 min, permeabilized in 0.5% Triton X-100 in PBS, and blocked in 5% goat serum. The

coverslips were incubated with primary antibody MHC 22287–1-AP (Proteintech) overnight at 4°C and then with Alexa Fluor 555-conjugated secondary antibody (Thermo Fisher Scientific) for 1 h. Cells were stained with Hoechst 33342 (1 μg/mL; Invitrogen) for 2 min at room temperature, washed, and then mounted with ProLong Gold (Invitrogen). The primary and secondary antibodies were diluted 1:400 and 1:1000, respectively. Microscopy was performed using the Zeiss Axio Observer Live Cell microscope and ImageJ Software for analysis (Schneider et al., 2012).

ChIRP-seq and ChIRP-qPCR—Antisense probes complementary to the genomic *MUNC* sequence (Table S1) were labeled with biotin-16dUTP (Roche, 11093070910) using a terminal transferase reaction (NEB, M0315). Probes were purified with the QIAquick Nucleotide Removal Kit (QIAGEN, Cat: 28304). ChIRP was performed as described (Chu et al., 2012). ChIRP libraries were prepared using DNA SMART ChIP-Seq kit (Takara Bio., 634865) with 1 ng of DNA as starting material. The quality and quantity of final libraries were assessed using an Agilent Technologies 2100 Bioanalyzer. Libraries were sequenced on an Illumina MiSeq platform (Genome Analysis and Technology Core, University of Virginia School of Medicine). All ChIRP-seq libraries data files are available under GEO accession number GSE174195 as a part of the SuperSeries GSE174218.

ChIRP-seq data from our study and from Tsai et al. (Tsai et al., 2018) were independently analyzed. First, data was aligned to the mouse assembly GRCm38/mm10 using bowtie2 version 2.3.4.1 (Langmead and Salzberg, 2012). Peak calling was done using MACS2 version 2.1.1.20160309 (Zhang et al., 2008) with a q value cutoff of 0.05. Peaks were assigned to gene-centric genomic regions with GREAT (Mclean et al., 2010). To ensure detection of only true positives, we designed an additional set of biotinylated probes based on the structure of the spliced *MUNC* (Table S1). Next, we performed ChIRP followed by qPCR with StepOnePlus™ Real-Time PCR System using PowerUp™ SYBR® Green Master Mix (Thermo Fisher). Oligonucleotide sequences are listed in Table S1.

RIP-qPCR—The RNA immunoprecipitation assay (RIP) was carried out as previously described (Klattenhoff et al., 2013) with slight modification. Briefly, 1×10^7 proliferating cells from each cell line were lysed with lysis buffer, incubated on ice for 20 min, and centrifuged at 2,500×g for 10 min. The nuclear pellet was lysed with RIP lysis buffer (25 mM Tris, pH 7.4, 150 mM KCl, 1 mM DTT, 0.5% NP-40, 1 mM PMSF, 10 mM NaF, 0.25% sodium deoxycholate, phosphatase inhibitor (Roche), protease inhibitor (Thermo Fisher), and 100 U/ml RNase inhibitor) and incubated on ice for 30 min. The cell lysates were sonicated for a total of 30 s (10 s on, 10 s off) with 10% amplitude. The lysates were centrifuged at 10,000×g for 15 min, and 1 mg of cell lysates were incubated with 4 μg of SMC3 antibody (Abcam) at 4°C overnight. The immuno-complexes were captured with 30 μl Protein G Plus Agarose beads (Thermo Fisher) and incubated at 4°C for 2 h. The SMC3-bound RNA-protein complexes were washed three times with RIP lysis buffer. Next, 1 ml of TRIzol was added directly to the pellet, and RNA was precipitated with ethanol and glycogen followed by cDNA synthesis and qPCR analysis. Fold enrichment was calculated by taking the ratio of *MUNC* enrichment in SMC3 immunoprecipitated over a negative control long noncoding RNA *H19*.

Other bioinformatic analyses—Motif analysis was performed using MEME (Bailey and Elkan, 1994) and FIMO (Grant et al., 2011). A map of chromosomal locations was created with the idiogramFISH package for R (Roa F, 2021). A graph that summarizes all interactions presented in this study was created using the rTRM package for R (Diez et al., 2014).

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical information of each experiment, including the statistical methods, the p value and sample numbers (n) are shown in figure legends. All experiments were performed at least two times, with similar results. The statistical method used for comparison between experimental groups was a two-tailed Student's t test. Statistical significance was expressed as a p value.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank the Genome Analysis and Technology Core at University of Virginia (RRID: SCR_018883) for ChIRP-seq sequencing. This work was supported by grants from the NIH (R01 AR067712 to A.D., R35 GM122532 to K.M.W., and R35 GM128635 to M.J.G.), a David and Lucile Packard Foundation (#2009-34710 to K.A.J.), an American Cancer Society Postdoctoral Fellowship (ACS 130845-RSG-17-114-01-RMC to C.A.W.), a Predoctoral Fellowship from the American Heart Association (18PRE33990261 to R.K.P.), Wagner fellowships from the University of Virginia (to R.K.P. and M.A.C.), and the F99/K00 NCI Predoctoral to Postdoctoral Fellow Transition Award (F99/K00CA253732 to R.K.P.).

REFERENCES

Anders S, Pyl PT, and Huber W (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169. [PubMed: 25260700]

Bailey TL, and Elkan C (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol 2, 28–36. [PubMed: 7584402]

Boerneke MA, Ehrhardt JE, and Weeks KM (2019). Physical and functional analysis of viral RNA genomes by SHAPE. Annu. Rev. Virol. 6, 93–117. [PubMed: 31337286]

Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, and Willard HF (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349, 38–44. [PubMed: 1985261]

Busan S, and Weeks KM (2018). Accurate detection of chemical modifications in RNA by mutational profiling (MaP) with ShapeMapper 2. RNA 24, 143–148. [PubMed: 29114018]

Busan S, Weidmann CA, Sengupta A, and Weeks KM (2019). Guidelines for SHAPE reagent choice and detection strategy for RNA structure probing studies. Biochemistry 58, 2655–2664. [PubMed: 31117385]

Cerase A, Pintacuda G, Tattermusch A, and Avner P (2015). Xist localization and function: new insights from multiple levels. Genome Biol. 16, 166. [PubMed: 26282267]

Choi S, Jeong HJ, Kim H, Choi D, Cho SC, Seong JK, Koo SH, and Kang JS (2019). Skeletal muscle-specific Prmt1 deletion causes muscle atrophy via deregulation of the PRMT6-FOXO3 axis. Autophagy 15, 1069–1081. [PubMed: 30653406]

Chu C, Quinn J, and Chang HY (2012). Chromatin isolation by RNA purification (ChIRP). J. Vis. Exp 3912. 10.3791/3912.

Cichewicz MA, Kiran M, Przanowska RK, Sobierajska E, Shibata Y, and Dutta A (2018). MUNC, an enhancer RNA upstream from the MYOD gene, induces a subgroup of myogenic transcripts in trans independently of MyoD. Mol. Cell Biol 38, e00617–e00655.

Darty K, Denise A, and Ponty Y (2009). VARNA: interactive drawing and editing of the RNA secondary structure. Bioinformatics 25, 1974–1975. [PubMed: 19398448]

Deigan KE, Li TW, Mathews DH, and Weeks KM (2009). Accurate SHAPE-directed RNA structure determination. Proc. Natl. Acad. Sci. U S A 106, 97–102. [PubMed: 19109441]

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, 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. [PubMed: 22955988]

Diez D, Hutchins AP, and Miranda-Saavedra D (2014). Systematic identification of transcriptional regulatory modules from protein-protein interaction networks. Nucleic Acids Res. 42, e6. [PubMed: 24137002]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Duszczyk MM, Zanier K, and Sattler M (2008). A NMR strategy to unambiguously distinguish nucleic acid hairpin and duplex conformations applied to a Xist RNA A-repeat. Nucleic Acids Res. 36, 7068–7077. [PubMed: 18987004]

Fang R, Moss WN, Rutenberg-Schoenberg M, and Simon MD (2015). Probing xist RNA structure in cells using targeted structure-seq. PLoS Genet 11, e1005668. [PubMed: 26646615]

Frank F, Kavousi N, Bountali A, Dammer EB, Mourtada-Maarabouni M, and Ortlund EA (2020). The lncRNA growth arrest specific 5 regulates cell survival via distinct structural modules with independent functions. Cell Rep 32, 107933. [PubMed: 32697996]

Grant CE, Bailey TL, and Noble WS (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017–1018. [PubMed: 21330290]

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. Nature 431, 931–945. [PubMed: 15496913]

Jaafar ZA, and Kieft JS (2019). Viral RNA structure-based strategies to manipulate translation. Nat. Rev. Microbiol. 17, 110–123. [PubMed: 30514982]

Kishioka Y, Thomas M, Wakamatsu J, Hattori A, Sharma M, Kambadur R, and Nishimura T (2008). Decorin enhances the proliferation and differentiation of myogenic cells through suppressing myostatin activity. J. Cell Physiol 215, 856–867. [PubMed: 18163379]

Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H, Butty VL, Torrey L, Haas S, et al. (2013). Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell 152, 570–583. [PubMed: 23352431]

Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. [PubMed: 22388286]

Li B, Cao Y, Westhof E, and Miao Z (2020a). Advances in RNA 3D structure modeling using experimental data. Front Genet. 11, 574485. [PubMed: 33193680]

Li Y, Zhou H, Chen Y, Zhong D, Su P, Yuan H, Yang X, Liao Z, Qiu X, Wang X, et al. (2020b). MET promotes the proliferation and differentiation of myoblasts. Exp. Cell Res 388, 111838. [PubMed: 31930964]

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. [PubMed: 25516281]

Mclean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, and Bejerano G (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28, 495–501. [PubMed: 20436461]

Merino EJ, Wilkinson KA, Coughlan JL, and Weeks KM (2005). RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). J. Am. Chem. Soc. 127, 4223–4231. [PubMed: 15783204]

Miao Z, Adamiak RW, Blanchet MF, Boniecki M, Bujnicki JM, Chen SJ, Cheng C, Chojnowski G, Chou FC, Cordero P, et al. (2015). RNA-Puzzles Round II: assessment of RNA structure

prediction programs applied to three large RNA structures. RNA 21, 1066–1084. [PubMed: 25883046]

Mousavi K, Zare H, Dell'orso S, Grontved L, Gutierrez-Cruz G, Derfoul A, Hager GL, and Sartorelli V (2013). eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. Mol. Cell 51, 606–617. [PubMed: 23993744]

Mueller AC, Cichewicz MA, Dey BK, Layer R, Reon BJ, Gagan JR, and Dutta A (2015). MUNC, a long noncoding RNA that facilitates the function of MyoD in skeletal myogenesis. Mol. Cell Biol 35, 498–513. [PubMed: 25403490]

Mustoe AM, Lama NN, Irving PS, Olson SW, and Weeks KM (2019). RNA base-pairing complexity in living cells visualized by correlated chemical probing. Proc. Natl. Acad. Sci. U S A 116, 24574–24582. [PubMed: 31744869]

Quinn LS, Haugk KL, and Grabstein KH (1995). Interleukin-15: a novel anabolic cytokine for skeletal muscle. Endocrinology 136, 3669–3672. [PubMed: 7628408]

Reuter JS, and Mathews DH (2010). RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics 11, 129. [PubMed: 20230624]

Roa F, P.T.M. (2021). idiogramFISH: Shiny App. Idiograms with Marks and Karyotype Indices (UFG, Goiânia: Universidade Federal de Goiás).

Schmidt K, Weidmann CA, Hilimire TA, Yee E, Hatfield BM, Schneekloth JS, Weeks KM, and Novina CD (2020). Targeting the oncogenic long non-coding RNA SLNCR1 by blocking its sequence-specific binding to the androgen receptor. Cell Rep 30, 541–554.e5. [PubMed: 31940495]

Schneider CA, Rasband WS, and Eliceiri KW (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675. [PubMed: 22930834]

Siegfried NA, Busan S, Rice GM, Nelson JA, and Weeks KM (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). Nat. Methods 11, 959–965. [PubMed: 25028896]

Smola MJ, Calabrese JM, and Weeks KM (2015a). Detection of RNA-protein interactions in living cells with SHAPE. Biochemistry 54, 6867–6875. [PubMed: 26544910]

Smola MJ, Rice GM, Busan S, Siegfried NA, and Weeks KM (2015b). Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. Nat. Protoc. 10, 1643–1669. [PubMed: 26426499]

Smola MJ, Christy TW, Inoue K, Nicholson CO, Friedersdorf M, Keene JD, Lee DM, Calabrese JM, and Weeks KM (2016). SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. Proc. Natl. Acad. Sci. U S A 113, 10322–10327. [PubMed: 27578869]

Stöckel D, Kehl T, Trampert P, Schneider L, Backes C, Ludwig N, Gerasch A, Kaufmann M, Gessler M, Graf N, et al. (2016). Multi-omics enrichment analysis using the GeneTrail2 web service. Bioinformatics 32, 1502–1508. [PubMed: 26787660]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A 102, 15545–15550. [PubMed: 16199517]

Tsai PF, Dell'orso S, Rodriguez J, Vivanco KO, Ko KD, Jiang K, Juan AH, Sarshad AA, Vian L, Tran M, et al. (2018). A muscle-specific enhancer RNA mediates cohesin recruitment and regulates transcription in trans. Mol. Cell 71, 129–141.e8. [PubMed: 29979962]

Umansky KB, Gruenbaum-Cohen Y, Tsoory M, Feldmesser E, Goldenberg D, Brenner O, and Groner Y (2015). Runx1 transcription factor is required for myoblasts proliferation during muscle regeneration. PLoS Genet 11, e1005457. [PubMed: 26275053]

Uroda T, Anastasakou E, Rossi A, Teulon JM, Pellequer JL, Annibale P, Pessey O, Inga A, Chillón I, and Marcia M (2019). Conserved pseudoknots in lncRNA MEG3 are essential for stimulation of the p53 pathway. Mol. Cell 75, 982–995.e9. [PubMed: 31444106]

Wang X, Blagden C, Fan J, Nowak SJ, Taniuchi I, Littman DR, and Burden SJ (2005). Runx1 prevents wasting, myofibrillar disorganization, and autophagy of skeletal muscle. Genes Dev. 19, 1715–1722. [PubMed: 16024660]

Wang H, Zheng H, Wang C, Lu X, Zhao X, and Li X (2017). Insight into HOTAIR structural features and functions as landing pads for transcription regulation proteins. Biochem. Biophys. Res. Commun. 485, 679–685. [PubMed: 28235488]

Wang Z, Civelek M, Miller CL, Sheffield NC, Guertin MJ, and Zang C (2018). BART: a transcription factor prediction tool with query gene sets or epigenomic profiles. Bioinformatics 34, 2867–2869. [PubMed: 29608647]

Watson ZL, Ward FR, Méheust R, Ad O, Schepartz A, Banfield JF, and Cate JH (2020). Structure of the bacterial ribosome at 2 Å resolution. Elife 9, e60482. [PubMed: 32924932]

Weeks KM (2021). SHAPE directed discovery of new functions in large RNAs. Acc. Chem. Res. 54, 2502–2517. [PubMed: 33960770]

Weidmann CA, Mustoe AM, Jariwala PB, Calabrese JM, and Weeks KM (2021). Analysis of RNA-protein networks with RNP-MaP defines functional hubs on RNA. Nat. Biotechnol. 39, 347–356. [PubMed: 33077962]

Wu HY, Hamamori Y, Xu J, Chang SC, Saluna T, Chang MF, O'malley BW, and Kedes L (2005). Nuclear hormone receptor coregulator GRIP1 suppresses, whereas SRC1A and p/CIP coactivate, by domain-specific binding of MyoD. J. Biol. Chem. 280, 3129–3137. [PubMed: 15563453]

Wutz A, Rasmussen TP, and Jaenisch R (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat. Genet. 30, 167–174. [PubMed: 11780141]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. 9, R137. [PubMed: 18798982]

Zhang C, Ge S, Gong W, Xu J, Guo Z, Liu Z, Gao X, and Wei X (2020). LncRNA ANRIL acts as a modular scaffold of WDR5 and HDAC3 complexes and promotes alteration of the vascular smooth muscle cell phenotype. Cell Death Dis 11, 435. [PubMed: 32513988]

Highlights

- *MUNC* lncRNA has two promyogenic isoforms that control different sets of genes
- *MUNC* lncRNA folds into a defined secondary structure with separate domains
- Specific structural domains mediate different functions of *MUNC* during myogenesis
- A single activity cannot explain all the actions of *MUNC* on gene expression

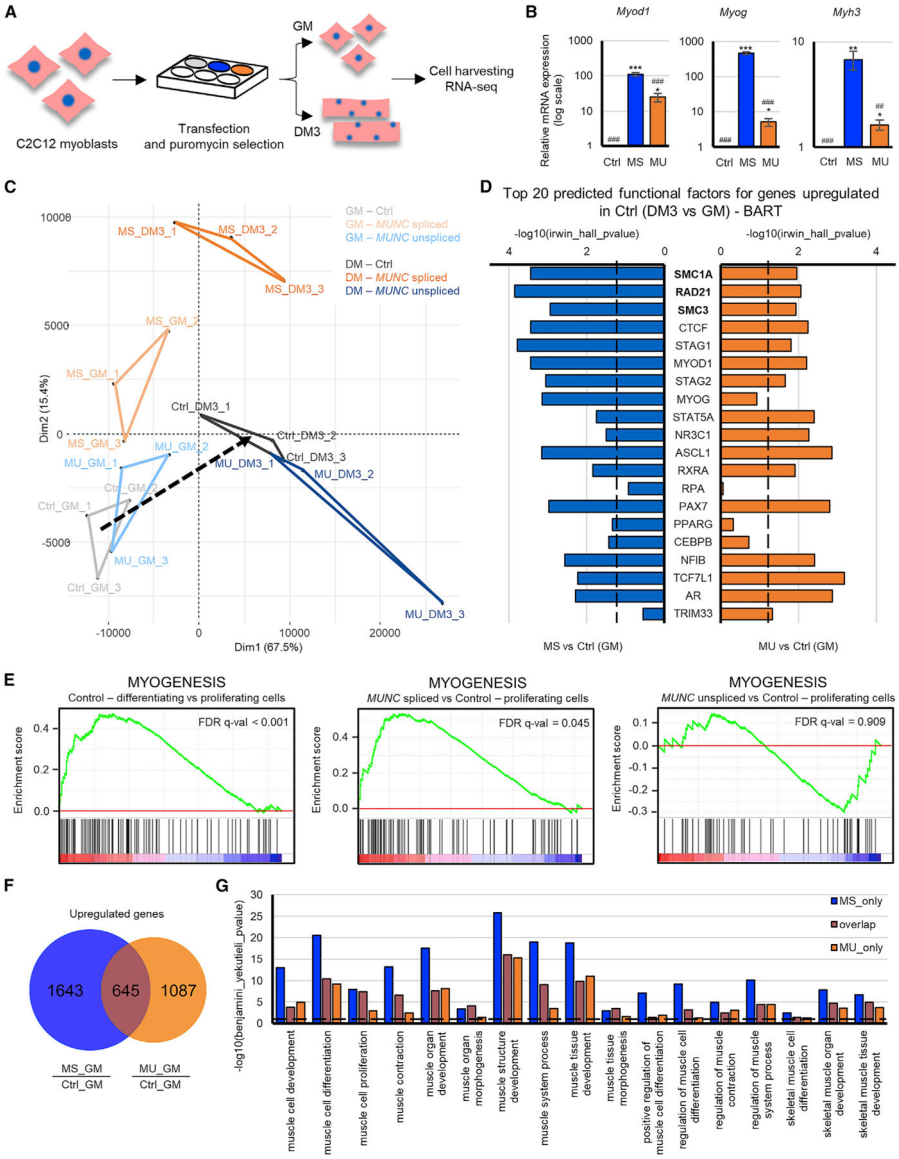


Figure 1. *MUNC* isoforms regulate different sets of genes involved in promyogenic pathways

(A) Scheme of experimental design.

(B) RT-qPCR analysis of Ctrl, MS- (*MUNC* spliced), and MU- (*MUNC* unspliced) overexpressing cells grown under proliferating conditions. Levels of mRNAs normalized to *Gapdh* and shown relative to Ctrl. $*/^{\#}p < 0.05$, $**/^{\#\#}p < 0.01$, $***/^{\#\#\#}p < 0.001$, two-tailed Student's t test versus Ctrl/MS. Data are representative of three biological replicates (mean ± SEM).

(C) PCA of RNA-seq data. Black dashed line indicates differentiation direction.

(D) BART analysis for upregulated genes shows enrichment of transcription factors that are predicted to upregulate genes during normal myogenesis. Black dashed line represents Irvin-Hall p value of 0.05. Subunits of cohesin complex are bolded.

(E) GSEA plots for myogenic genes.

(F) Venn diagram of the overlap between differentially upregulated genes.

(G) Muscle-related GO terms enriched in upregulated genes. Black dashed line represents Benjamini-Yekutieli p value of 0.05.
See also Figure S1.

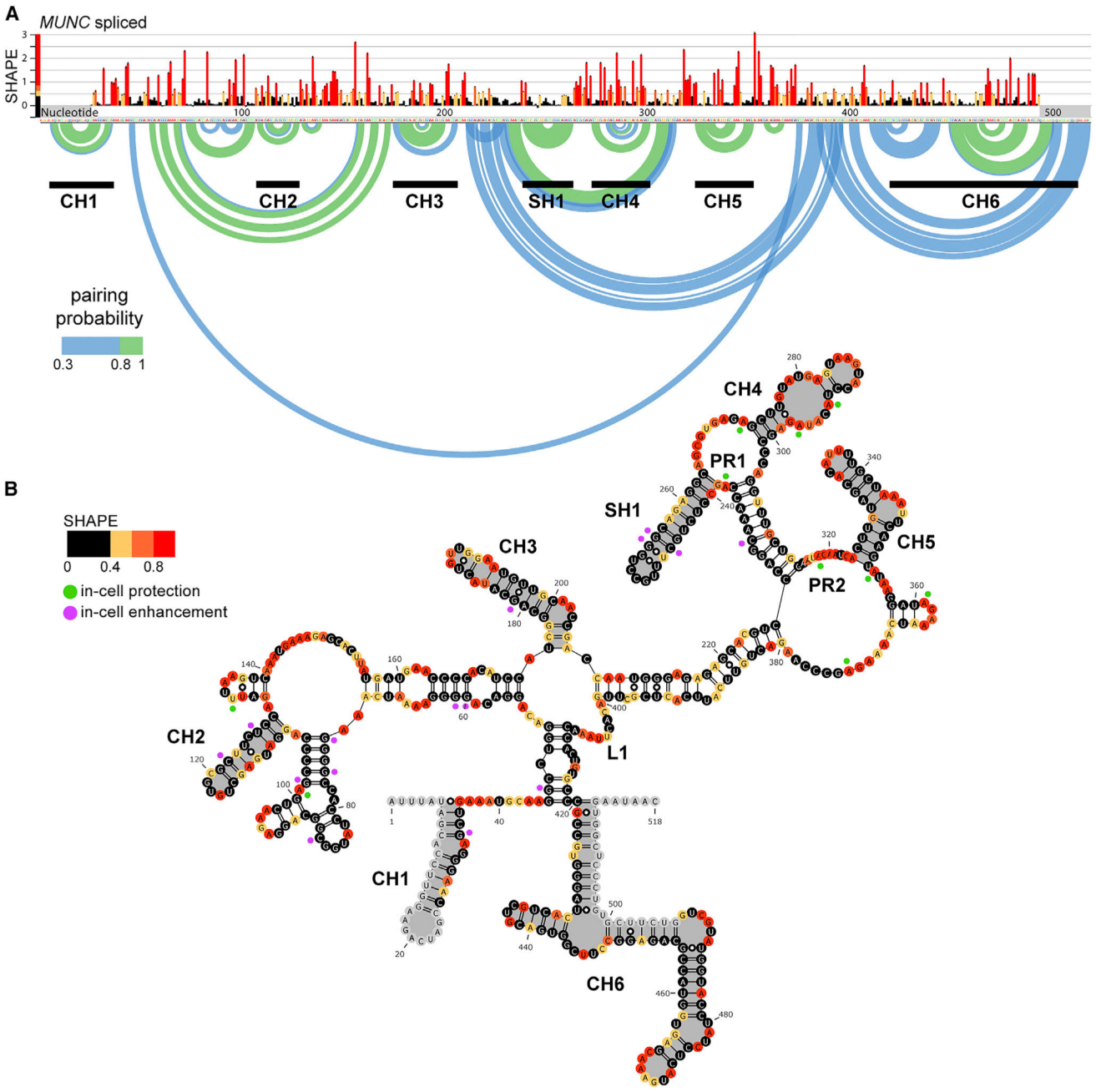


Figure 2. SHAPE-MaP of *MUNC* spliced isoform reveals a structured architecture

(A) SHAPE reactivity profile (top) and pairing probabilities (bottom) for the cell-free MS isoform (n = 2). Mean reactivities (±SE) are colored by relative value as shown in (B).

(B) Minimum free energy secondary structure model of MS. Structural domains are highlighted in gray.

See also Figures S2 and S3.

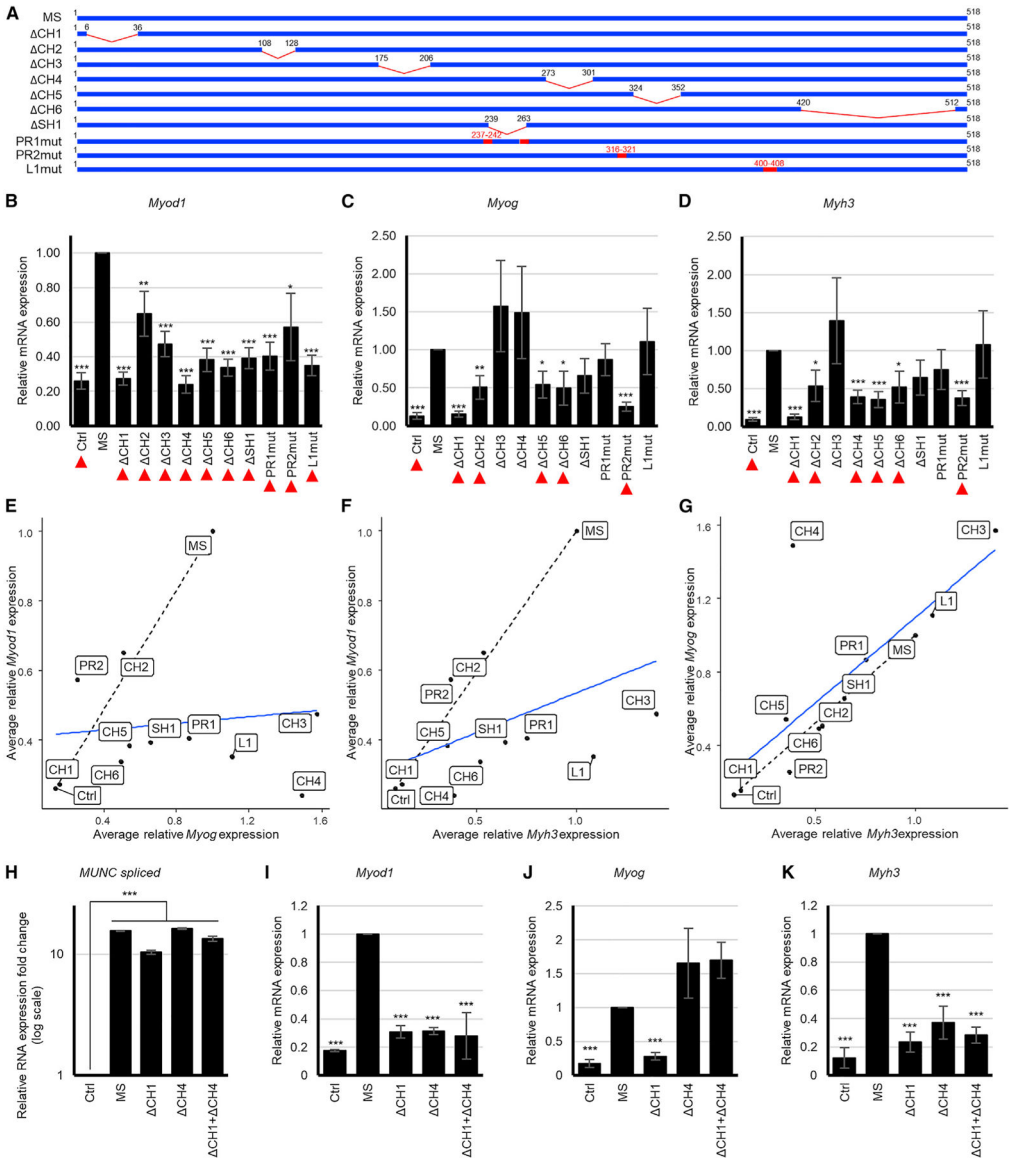


Figure 3. Distinct structural domains of the spliced isoform of *MUNC* are required for induction of promyogenic factors.

(A) Constructs used in the study.

(B–D) RT-qPCR analyses of (B) *Myod1*, (C) *Myog*, and (D) *Myh3* mRNAs normalized to *Gapdh* and shown relative to MS overexpression in proliferating control cells. Data are representative of at least two independent transfectants with at least three biological replicates for each of them (mean ± SEM).

(E–G) Plots of average induction of (E) *Myod1* versus *Myog*, (F). *Myod1* versus *Myh3*, and (G) *Myog* versusw *Myh3* in Ctrl cells and cells expressing MS constructs. Black line is the theoretical correlation line if the expression of one gene is dependent on another; blue line shows experimental correlation.

(H) RT-qPCR analysis of *MUNC* levels normalized to *Gapdh* and shown relative to Ctrl. Three biological replicates (mean ± SEM).

(I–K) RT-qPCR analyses of (I) *Myod1*, (J) *Myog*, and (K) *Myh3* mRNAs normalized to *Gapdh* and shown relative to MS. Data are representative of three biological replicates (mean ± SEM). *p < 0.05, **p < 0.01, ***p < 0.001, two-tailed Student's t test.
See also Figure S4.

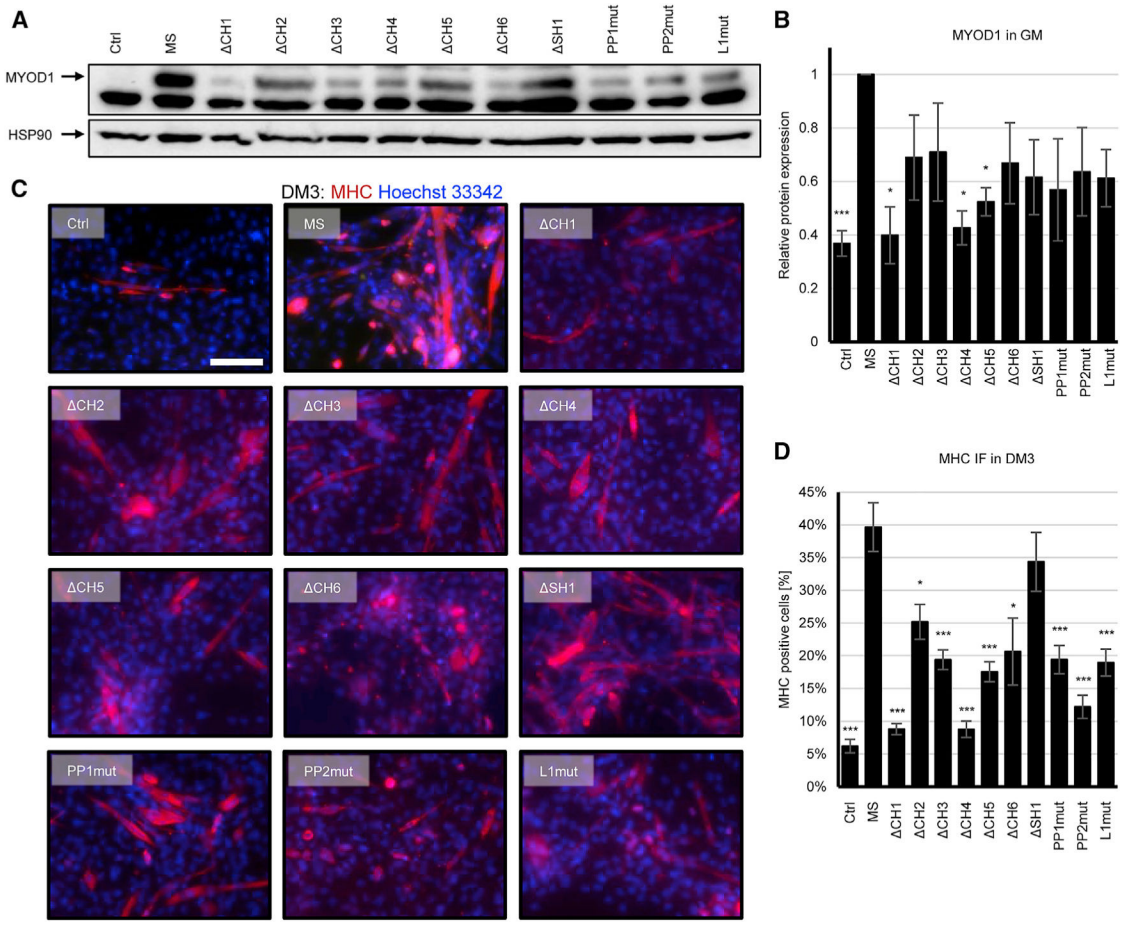


Figure 4. *MUNC* structural domains are necessary for myotube formation

(A) Representative western blot for MYOD1 in proliferating cells with or without indicated MS constructs. HSP90 served as a loading control.

(B) Quantification of MYOD1 protein in proliferating cells normalized to HSP90 and shown relative to MS. Data represent two independent transfectants with two biological replicates for each (one of four is shown in A; mean ± SEM).

(C) Representative immunofluorescence images. Scale bar, 200 μm.

(D) Quantification of the percentage of MHC-positive cells. Data represent two independent transfectants with at least 1,000 nuclei counted (mean ± SEM). *p < 0.05, **p < 0.01, ***p < 0.001, two-tailed Student's t test.

See also Figure S5.

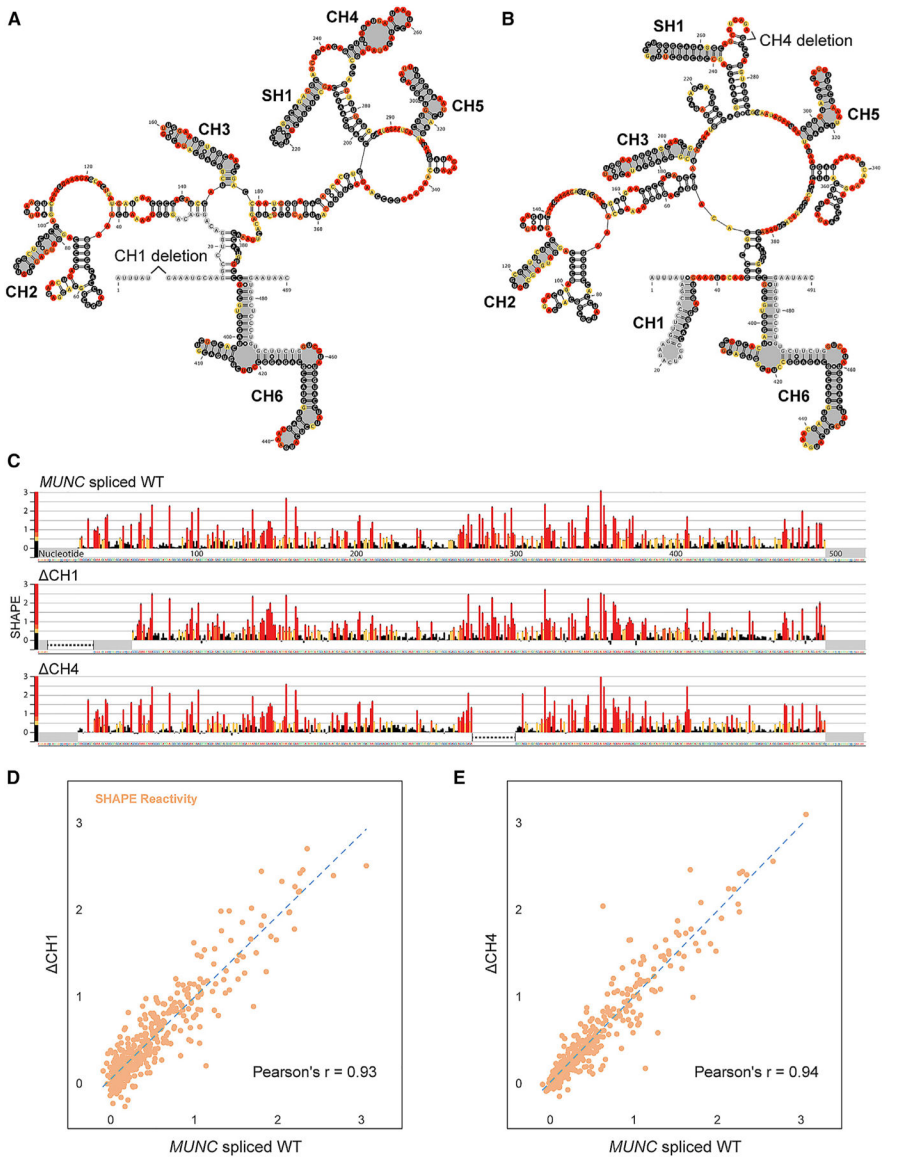


Figure 5. Deletion of CH1 or CH4 domains does not disrupt *MUNC* lncRNA structure

(A and B) Minimum free energy secondary structure model of (A) ΔCH1 and (B) ΔCH4, color coded for SHAPE reactivities. Structural domains are highlighted in gray.

(C) SHAPE reactivity profiles from cell-free SHAPE-MaP. Mean reactivities (±SE) are colored by relative value. Gray boxes indicate primers used for PCR. Black dashed lines represent deletions.

(D and E) Correlation plot of SHAPE reactivities for (D) wild-type (x-axis) and (E) mutant (y-axis) spliced isoforms.

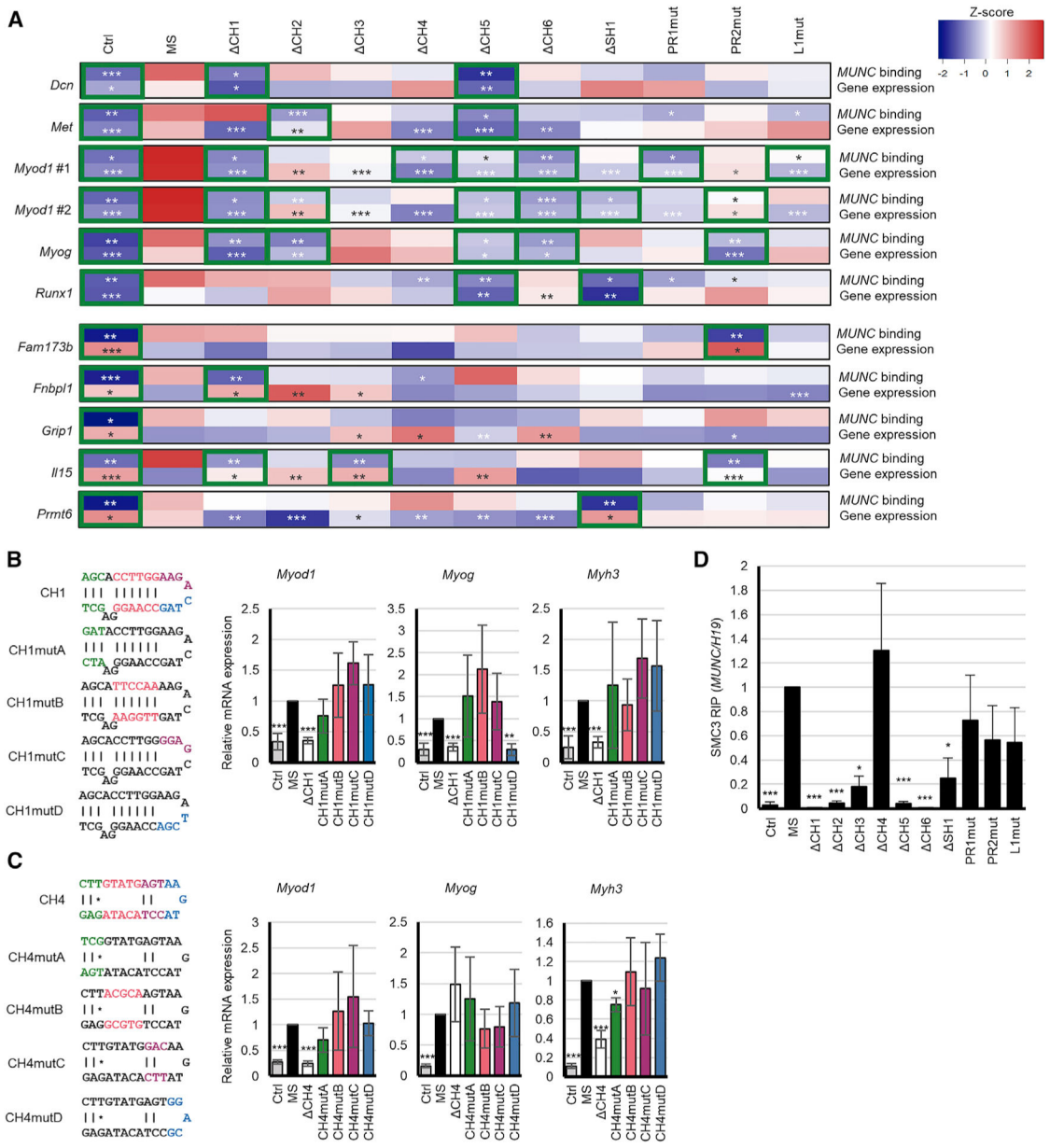


Figure 6. Distinct *MUNC* spliced domains are required for binding to specific genomic sites, regulating adjoining gene expression, and interacting with protein partners

(A) Heatmap of ChIRP-qPCR ("*MUNC* binding") and RT-qPCR ("gene expression") Z scores upon expression of MS constructs in proliferating cells. Mean values from three biological replicates are presented as *Z* score. Green outlined boxes identify domains required both for *MUNC* binding and gene expression regulation.

(B and C) Left: scheme of point mutations in CH1 (B) and CH4 (C) domains (with compensatory mutations to maintain base-pairing when necessary). Right: RT-qPCR of *Myod1*, *Myog*, and *Myh3* mRNAs normalized to *Gapdh* upon expression of *MUNC* constructs in proliferating cells. Shown relative to MS. Three independent transfectants (mean ± SEM).

(D) SMC3 association with *MUNC* constructs in proliferating cells. Levels of *MUNC* lncRNA in RIP-qPCR were normalized to *H19* lncRNA and shown relative to MS. Three biological replicates (mean ± SEM). *p < 0.05, **p < 0.01, ***p < 0.001, two-tailed Student's t test.

See also Figures S6 and S7.

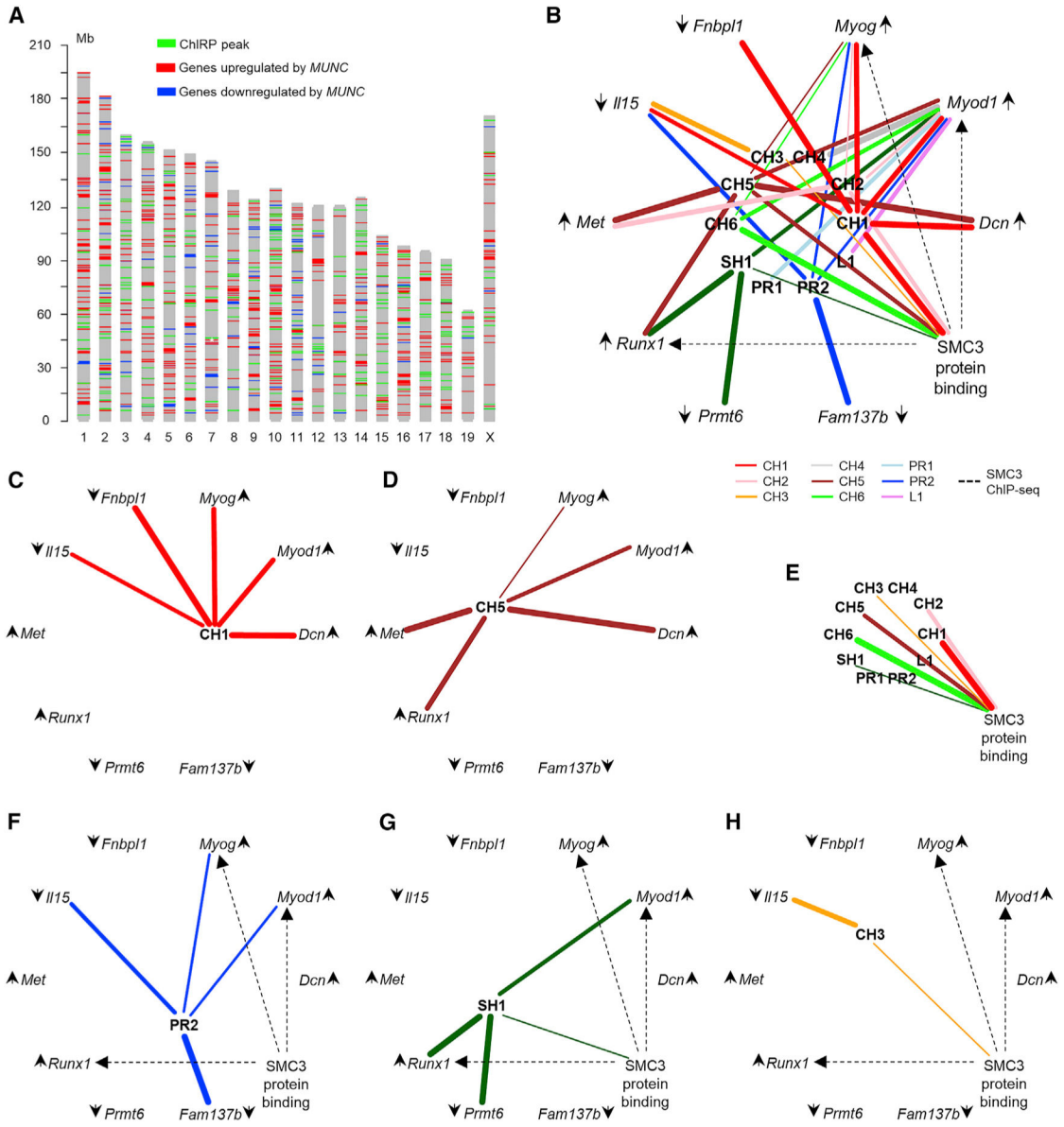


Figure 7. *MUNC* regulates genes in *cis* and in *trans* through multiple mechanisms

(A) Chromosomal locations of *MUNC* ChIRP peaks (green), genes upregulated by MS (red), and genes downregulated by MS (blue). White star marks *Myod1* locus.

(B) Graph summarizing all interactions identified in this study. Lines represent significant alterations relative to MS in DNA binding and gene regulation or SMC3 binding when a particular domain is mutated. Line width represents effect size; dashed lines indicate SMC3 binding sites in promoters of *Myod1*, *Myog*, and *Runx1*.

(C–G) Simplified graphs as in B, summarizing (C) interactions of CH1, (D) interactions of CH5, (E) domains required for optimal *MUNC*-SMC3 binding, (F) interactions of PR2, and (G) interactions of CH3.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal MyoD (5.8A)	Santa Cruz	Cat# sc-32758; RRID: AB_627978
Rabbit polyclonal MYH3	Proteintech	Cat# 22287-1-AP; RRID: AB_2879060
Mouse monoclonal myogenin (F5D)	Santa Cruz	Cat# sc-12732; RRID: AB_627980
Mouse monoclonal HSP90α/β (F-8)	Santa Cruz	Cat# sc-13119; RRID: AB_675659
Anti-rabbit IgG, HRP-linked Antibody	Cell Signaling	Cat# 7074; RRID: AB_2099233
Anti-mouse IgG, HRP-linked Antibody	Cell Signaling	Cat# 7076; RRID: AB_330924
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor Plus 555	Thermo Fisher Scientific	Cat# A32732; RRID: AB_2633281
Anti-SMC3 antibody	Abcam	Cat# ab9263; RRID: AB_307122
Bacterial and virus strains		
Stellar™ Competent Cells	Takara	Cat# 636763
Chemicals, peptides, and recombinant proteins		
Proteinase K	New England Biolabs	Cat# P8107S
Phenol/Chloroform/Isoamyl Alcohol (25:24:1 Mixture, pH 6.7/8.0, Liq.), Fisher BioReagents	Thermo Fisher Scientific	Cat# BP1752I-400
5-nitroisatoic anhydride (5NIA)	AstaTech	Cat# 69445
Dimethyl sulfate	Millipore Sigma	Cat# D186309
TURBO DNase	Thermo Fisher Scientific	Cat# AM2238
TRIzol™ Reagent	Thermo Fisher Scientific	Cat# 15596018
SuperScript II Reverse Transcriptase	Thermo Fisher Scientific	Cat# 18064014
Manganese (II) chloride tetrahydrate	Millipore Sigma	Cat# M3634
Betaine	Millipore Sigma	B0300-1VL
Q5 Hot-start high Fidelty DNA Polymerase	New England Biolabs	Cat# M0493
Puromycin	Sigma	Cat# P9620
Hygromycin	Life technologies	Cat# 10687-010
Lipofectamine 3000	Life Technologies	Cat# L3000015
DMEM-high glucose medium	GE Healthcare Life Sciences co.	Cat# SH30022.FS
FBS (20%)	Gibco	Cat# 10437-028
HS (2%)	GE Healthcare Life Sciences co.	Cat# SH30074.03
MilliporeSigma™ Immobilon™ Western Chemiluminescent HRP Substrate (ECL)	Millipore Sigma	Cat# WBKLS0500
Hoechst 33342	Life Technologies	Cat# H3570
ProLong™ Gold Antifade Mountant	Life Technologies	Cat# P10144
Biotin-16-dUTP	Roche	Cat# 11093070910
Terminal Transferase	New England BioLabs	Cat# M0315S

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HaltTM Protease Inhibitor	Thermo Fisher Scientific	Cat# 78438
SUPERase●In™ RNase Inhibitor (20 U/μL)	Ambion	Cat# AM2694
RNase H, from E. coli, 10 U/μL	Ambion	Cat# AM2293
Rnase A	New England Biolabs	Cat# T3018L
Dynabeads™ MyOne™ Streptavidin C1	Thermo Fisher Scientific	Cat# 65002
Phosphatase Inhibitor	Millipore Sigma	Cat# 5892970001
Pierce™ Protein G Plus Agarose	Thermo Fisher Scientific	Cat# 22851
Critical commercial assays		
NEBNext® Poly(A) mRNA Magnetic Isolation Module	New England BioLabs	Cat# E7490S
NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina	New England BioLabs	Cat# E7420S
Agencourt AMPure XP kit	Beckman Coulter	Cat# A63881
Direct-zol RNA MiniPrep Kit	Zymo Research	Cat# R2052
PD-10 Desalting Columns	GE Healthcare Life Sciences	Cat#17-0851-01
GeneJET RNA Cleanup and Concentration Micro Kit	Fisher	Cat# K0841
Illustra Microspin G-50 columns	GE Healthcare Life Sciences	Cat# 27533001
Zymoclean Gel DNA Recovery Kit	Genessee	Cat# 11-301
Qubit™ dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat# Q32851
Nextera XT DNA Library Preparation Kit	Illumina	Cat# FC-131-1024
Nextera XT Index Kit (24 indexes, 96 samples)	Illumina	Cat# FC-131-1001
In-Fusion HD Cloning kit	Takara	Cat# 638910
GoScript RT cDNA synthesis kit	VWR	Cat# PAA5004
PowerUp™ SYBR® Green Master Mix	Thermo Fisher Scientific	Cat# A25778
QIAquick Nucleotide Removal Kit	Qiagen	Cat# 28304
DNA SMART ChIP-Seq kit	Takara Bio	Cat# 634865
Deposited data		
C2C12 RNA-seq	This paper	GSE174203
MUNC SHAPE-MaP	This paper	GSE174217
MUNC ChIRP-seq	This paper	GSE174195
Experimental models: Cell lines		
C2C12 PAX7 negative murine myoblasts	N/A	N/A
C2C12 PAX7 negative *MUNC* spliced overexpressing murine myoblasts	(Mueller et al., 2015)	N/A
C2C12 PAX7 negative *MUNC* genomic overexpressing murine myoblasts	(Mueller et al., 2015)	N/A
C2C12 PAX7 negative *MUNC* spliced ΔCH1 overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced ΔCH2 overexpressing murine myoblasts	This paper	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
C2C12 PAX7 negative *MUNC* spliced ΔCH3 overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced ΔCH4 overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced ΔCH5 overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced ΔCH6 overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced DSH1 overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced PP1mut overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced L1mut overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced ΔCH1+ ΔCH4 overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH1mutA overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH1mutB overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH1mutC overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH1mutD overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH4mutA overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH4mutB overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH4mutC overexpressing murine myoblasts	This paper	N/A
C2C12 PAX7 negative *MUNC* spliced CH4mutD overexpressing murine myoblasts	This paper	N/A
Oligonucleotides		
MUNC SHAPE-MaP RT and PCR primers, see Table S1	This paper	N/A
MUNC cloning primers, see Table S1	This paper	N/A
ChIRP oligonucleotides, see Table S1	This paper	N/A
ChIRP-qPCR primers, see Table S1	This paper	N/A
RT-qPCR primers, see Table S1	This paper	N/A
Recombinant DNA		
pLPCX *MUNC* spliced	(Mueller et al., 2015)	N/A
pLHCX *MUNC* spliced	(Cichewicz et al., 2018)	N/A
Software and Algorithms		
STAR v2.5	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
HTSeq 0.6.1p1 (Python 2.7.5)	(Anders et al., 2015)	https://htseq.readthedocs.io/en/release_0.11.1/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R Studio	https://www.rstudio.com/products/rstudio/	N/A
DESeq2	(Love et al., 2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
BART 2.0	(Wang et al., 2018)	http://bartweb.org/
Gene Ontology (GO) by GeneTrail2r	(Stöckel et al., 2016)	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
ShapeMapper 2	(Busan and Weeks, 2018)	https://github.com/Weeks-UNC/shapemapper2
SuperFold	(Smola et al., 2015b)	https://github.com/Weeks-UNC/Superfold
RNAStructure	(Reuter and Mathews, 2010)	https://rna.urmc.rochester.edu/RNAstructure.html
VARNA visualization applet for RNA	(Darty et al., 2009)	http://varna.lri.fr/
ImageJ Software 1.50i (Java 1.6.0_24)	(Schneider et al., 2012)	https://imagej.nih.gov/ij/
bowtie2	(Langmead and Salzberg, 2012)	http://bowtiebio.sourceforge.net/bowtie2/index.shtml
MACS2	(Zhang et al., 2008)	https://pypi.org/project/MACS2/
GREAT	(Mclean et al., 2010)	http://great.stanford.edu/public/html/
MEME	(Bailey and Elkan, 1994)	https://memesuite.org/meme/
FIMO	(Grant et al., 2011)	https://memesuite.org/meme/doc/fimo.html
Other		
StepOnePlus™ Real-Time PCR System	Applied Biosystems	Cat# 4376600
Agilent Technologies 2100 Bioanalyzer	Agilent	N/A
Illumina HiSeq 2500 instrument	Illumina	N/A
Illumina MiSeq	Illumina	N/A
QuBIT	Thermo Fisher Scientific	N/A
Zeiss Axio Observer Live Cell microscope	Zeiss	N/A