Deficiency in the nuclear long noncoding RNA causes myogenic defects and heart remodeling in mice

Citation for published version:

Digital Object Identifier (DOI):
10.15252/embj.201899697

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
EMBO Journal

Publisher Rights Statement:
This is an open access article under the terms of the Creative Commons Attribution Non Commercial-NoDerivs 4.0 License, which permits use and distribution in any medium, provided the original work is properly cited, the use is noncommercial and no modifications or adaptations are made.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Deficiency in the nuclear long noncoding RNA Charme causes myogenic defects and heart remodeling in mice

Monica Ballarino¹, Andrea Cipriano¹, Rossella Tita¹, Tiziana Santini², Fabio Desideri¹, Mariangela Morlando¹, Alessio Colantoni¹, Claudia Carrieri³, Carmine Nicoletti⁴, Antonio Musarò²,⁴, Dònal O’Carroll³,⁴ & Irene Bozzi⁴

Abstract

Myogenesis is a highly regulated process that involves the conversion of progenitor cells into multinucleated myofibers. Besides proteins and miRNAs, long noncoding RNAs (lncRNAs) have been shown to participate in myogenic regulatory circuitries. Here, we characterize a murine chromatin-associated muscle-specific lncRNA, Charme, which contributes to the robustness of the myogenic program in vitro and in vivo. In myocytes, Charme depletion triggers the disassembly of a specific chromosomal domain and the downregulation of myogenic genes contained therein. Notably, several Charme-sensitive genes are associated with human cardiomyopathies and Charme depletion in mice results in a peculiar cardiac remodeling phenotype with changes in size, structure, and shape of the heart. Moreover, the existence of an orthologous transcript in human, regulating the same subset of genes, suggests an important and evolutionarily conserved function for Charme. Altogether, these data describe a new example of a chromatin-associated lncRNA regulating the robustness of skeletal and cardiac myogenesis.

Keywords CRISPR/Cas9; epigenetic control; heart development; lncRNAs; myogenesis

Subject Categories Chromatin, Epigenetics, Genomics & Functional Genomics; RNA Biology; Vascular Biology & Angiogenesis

DOI 10.15252/embj.201899697 | Received 23 April 2018 | Revised 17 July 2018 | Accepted 21 July 2018 | Published online 3 September 2018

The EMBO Journal (2018) 37: e99697

Introduction

Myogenesis is a highly regulated multistep process, in which the conversion of progenitor cells into multinucleated and functional myofibers can be easily reproduced in vitro. Besides the well-characterized protein factors and miRNAs (Buckingham & Rigby, 2014), also long noncoding RNAs (lncRNAs) have been shown to participate in myogenic regulatory circuitries (Rinn & Chang, 2012; Fatica & Bozzi, 2014; Ballarino et al, 2015). lncRNAs, through their ability to interact with nucleic acids as well as with proteins (Guttman & Rinn, 2012; Rinn & Chang, 2012; Batista & Chang, 2013; Cipriano & Ballarino, 2018), act as scaffolds for the formation of specific functional complexes where different proteins and RNA and even DNA molecules can be tethered together (Wutz et al, 2002; Clemson et al, 2009; Tsai et al, 2010; Ariel et al, 2014; Hacisuleyman et al, 2014; Ribeiro et al, 2017). Nuclear lncRNAs have been described to control the transcriptional myogenic program through diverse modes of action (Ballarino et al, 2016): as enhancer-associated transcripts (Mousavi et al, 2014; Mueller et al, 2015; Ounzain et al, 2015), as guides for the recruitment of PRCs and MLL epigenetic regulators or DNA methyltransferases (DNMT; Wang et al, 2015, 2016), and also as allosteric inhibitors (Han et al, 2014; Wang et al, 2016). Important functions have also been ascribed to cytoplasmic lncRNAs through their ability to control mRNA stability and translation (Wang et al, 2013), to modulate miRNA function acting as competing endogenous RNA (Cesana et al, 2011; Salmena et al, 2011; Han et al, 2015), or to encode for micropeptides (Anderson et al, 2015; Nelson et al, 2016). Interestingly, de-regulation of these molecules was often associated with skeletal (Ballarino et al, 2015) and cardiac (Han et al, 2014; Ounzain & Pedrazzini, 2015; Uchida & Dimmeler, 2015; Wang et al, 2016) diseases.

1 Department of Biology and Biotechnology Charles Darwin, Sapienza University of Rome, Rome, Italy
2 Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, Rome, Italy
3 MRC Centre for Regenerative Medicine, University of Edinburgh, Edinburgh, UK
4 DAHFMO-Unit of Histology and Medical Embryology, Sapienza University of Rome, Rome, Italy
5 Institute Pasteur Fondazione Cenci-Bolognetti, Sapienza University of Rome, Rome, Italy
6 Institute of Molecular Biology and Pathology, CNR, Sapienza University of Rome, Rome, Italy
*Corresponding author. Tel: +39 06 4991 2202; Fax: +39 06 4991 2500; E-mail: irene.bozzi@uniroma1.it
†These authors contributed equally to this work
We recently identified a novel nuclear lncRNA, Charme (Chromatin architect of muscle expression), conserved in human and with restricted expression in skeletal and cardiac muscles (Ballarino et al., 2015). Here, we show that the depletion of Charme produces the downregulation of myogenic genes, many of which involved in human cardiomyopathies (Becker et al., 2011; Harvey & Leinwand, 2011). Through a combination of ChIRP and FISH approaches, we identified a Charme/chromatin interaction that is required for the correct expression of a subset of relevant myogenic genes. Notably, the functional knockdown of Charme in mice produced morphological alterations of both skeletal and cardiac muscles and resulted in a shorter life span. The structural and functional conservation of Charme in human further corroborates its relevance in the control of proper muscle differentiation and homeostasis.

Results

Charme is achromatin-associated long noncoding RNA

Previous transcriptome analysis from murine C2C12 myoblasts and myotubes discovered novel lncRNAs with muscle restricted expression (Ballarino et al., 2015). Among them, Inc-405, here named Charme (Chromatin architect of muscle expression; Figs 1A and EV1A), is a novel isoform of the 5430431A17Rik transcript of particular interest for (i) its specific expression in both skeletal and cardiac muscles, (ii) its predominant nuclear localization, and (iii) the presence of an orthologue transcript in human.

The expression of Charme was studied by 5′-RACE, Northern blot, and RT–PCR analyses (Figs 1B and EV1B): It is predominantly localized in the nucleus, mainly associated with the chromatin, with a minor fraction being present in the cytoplasm. In C2C12 cells, RNA FISH showed a well-defined punctate localization of Charme in the nuclei of myofibers defined as multinucleated cells expressing the myogenic marker myosin heavy chain (MHC). Charme was instead absent in proliferating myoblasts, identified as mononucleated cells lacking MHC staining (Fig 1C). Charme expression starts at day 1 upon shift to differentiation medium (DM), reaching high levels at day 3 (Fig 1D). In agreement with RNA-seq data, RT–PCR analyses revealed the existence of an isoform (pCharme) still retaining intron 1 and devoid of intron 2 (Figs 1A and D, and EV1C and D). Notably, actinomycin experiments indicated that this isoform was as stable as the fully spliced Charme species (mCharme; Fig EV1E). RNA FISH with intron probes detected an average of three nuclear spots which co-localize with the exonic signals (Appendix Fig S1A and B). This pattern, whose specificity was tested by both RNaseA and RNAi sensitivity (Appendix Fig S1C), well correlates with the aneuploidy of C2C12 cells having three copies of chromosome 7, where Charme locus resides. The number, the intensity, and the focal plane distribution of the signals obtained with exonic vs. intronic probes allowed us to establish that the major fraction of Charme is retained at the sites of its own transcription as a precursor species.

Charme depletion affects myogenesis

To analyze the role of Charme in myogenesis, we treated C2C12 myotubes with two different LNA GAPmers (GAP-2 and GAP-2/3; Fig 1A), which provided 60–70% average reduction in both intron 1-harboring (pCharme) and mature (mCharme) Charme transcripts (Fig 1E). Upon knockdown, 50% decrease of the mRNAs for two diagnostic markers of differentiation, myosin creatine kinase (MCK) and the myosin heavy chain (MHC), was found (Fig 1E and Table EV1) indicating quite a clear impairment of myogenesis. Morphological parameters were also specifically affected: a consistent reduction in the caliber of differentiated MHC-positive myotubes (Fig 1F) was observed, together with a decrease of more than twofold of the fusion index (F.I.) and of the number of nuclei per fiber (Nu/MT). In parallel, the number of mononucleated (MonoMHC⁺) cells increased by a factor of 3 (Fig 1G). These results suggest that Charme inhibition mainly affects the ability of myogenic cells to fuse and to proceed to later stages of differentiation.

RNA-seq analysis on cells treated with GAP-2 or GAP-2/3, in parallel with a scramble control (Appendix Fig S2A and B) led to the identification of a set of 826 genes differentially expressed upon Charme depletion with two different sets of LNA GAPmers (Fig 2A and Appendix Fig S2C). Of these, 302 were upregulated and 524 downregulated in comparison with the scramble condition (Fig 2A and Table EV1). Real-time quantitative PCR on a selected number of targets confirmed the sequencing data (Appendix Fig S2D). The Gene Ontology (GO) term enrichment study performed with the FIDEA web tool (D’Andrea et al., 2013) suggested that Charme may act by activating genes involved in muscle function and contraction.
Figure 1.

**mmu_Charme**

11.3 kb 11.3 kb

**mCharme**

0.12 kb 0.12 kb

**pCharme**

GAP-2 GAP-2/3

**Carme MHC DAPI**

**DM**

Cyt Nu - Np Chr -

mCharme

pCharme

GAPDH

pre-GAPDH

RT-PCR

GM 1 3 5 - days

1-3

2-3

1-int1a

2-int2

RT-PCR

GAPDH

Charme amplicons

**mCharme**

**pCharme**

**MCK**

**MHC**

Expression fold

F.I. (%) Nu/MT MonoMHC+ (%)

GAP-scr

GAP-2

GAP-2/3

© 2018 The Authors

The EMBO Journal 37: e99697 | 2018 3 of 16

Monica Ballarino et al Charme lncRNA functions in myogenesis

The EMBO Journal

Published online: September 3, 2018
Charme lncRNA functions in myogenesis
Monica Ballarino et al

Published online: September 3, 2018
The EMBO Journal 37: e99697 | 2018 © 2018 The Authors

Figure 2.
Figure 2. Charme regulates myogenesis by interacting with the nctc region.

A. RNA-seq analysis of transcriptome changes upon Charme depletion. Heatmap was drawn using the heatmap3 R package and represents hierarchical clustering of the common genes differentially expressed upon Charme depletion with both GAP-2 and GAP-2/I GAPmers, compared to GAP-scr control. The expression levels correspond to mean-centered log₂-transformed FPKM values. See also Table EV1.

B. Gene Ontology (GO) term enrichment analyses performed by FDEA (D’Andrea et al., 2013) on genes downregulated (top) or upregulated (bottom) upon Charme depletion. Bars indicate the top 10 categories of cellular components in decreasing order of significance. Threshold (P-value < 0.01) is indicated by the dashed black line.

C. KEGG pathway enrichment analysis on genes downregulated by Charme depletion.

D. Quantification of recovered RNA (left) and DNA (right) upon Charme ChIRP with the pool of biotinylated probes reported in Table EV4.

E. RNA/DNA FISH for Charme RNA and nctc (left) or Inc-31 (right) loci in 2-day differentiated myotubes. Inserts show a magnification of the spots converted into binary images (I), 3D rendering (II), or rotated on Z-axis (III). Yellow arrows indicate signals overlapping. Right histogram reports the mean ± SD percentage of nuclei with paired or overlapped spots from two biological replicates. Scale bar = 5 μm.

F. qRT-PCR quantification of Charme and Charme target genes in GAP-scr- vs. GAP-2-transfected (2-day differentiated) myotubes. PCR data were normalized to GAPDH and represent mean ± SD of triplicates.

G. RNA Pol II (left) and H3K9ac (right) ChIP performed in GAP-scr- vs. GAP-2-transfected myotubes. The recovered chromatin was analyzed by qPCR in parallel with an intergenic region used to normalize the two (GAP-scr vs. GAP-2) conditions. Data were subtracted for background and are expressed as input percentage (% Input). Histograms represent the mean ± SEM of three biological replicates.

Data information: *P < 0.05, **P < 0.01, ***P < 0.001, unpaired Student’s t-test.

Source data are available online for this figure.

processes (Fig 2B). To note, we did not observe significant enrichment of any cell-cycle- and/or proliferation-related categories.

Interestingly, Charme depletion resulted in the downregulation of genes (i.e., Myh7 and Tnni2; Appendix Fig S2E) implicated in several types of cardiomyopathies (Fig 2C), such as familial hypertrophic (HCM), dilated (DCM), and arrhythmogenic right ventricular (ARVC) cardiomyopathies (Harvey & Leinwand, 2011; Becker et al., 2011; Maron et al., 2012).

pCharme is localized at specific chromosomal sites

To elucidate Charme mode of action, we proceeded with the identification of the sites of chromatin interaction in C2C12 cells. Chromatin isolation by RNA purification (ChIRP) analysis was carried out with two different pools of four alternating antisense biotinylated oligonucleotides (even and odd), together with four oligonucleotides against lacZ sequences (Appendix Fig S2F). Figure 2D (panel RNA) shows that a specific pull-down of pCharme isoform was obtained when compared to the lacZ control and the GAPDH pre-mRNA, while no enrichment was obtained for mCharme isoform. The copurified DNA was subjected to high-throughput sequencing (ChIRP-seq), and the reads from odd and even pull-downs were compared to those of the lacZ control (Table EV2). In agreement with the strong chromatin association, the major peak corresponded to the Charme locus. Several other genomic loci were identified (Table EV2); among them, the most prominent corresponded to a region (hereafter named as nctc) located on the same Charme chromosome (chr7:149,746,850–149,747,033) and containing several genes affected by Charme depletion (Appendix Fig S2G and Table EV1). The nctc peak was validated by PCR amplification on the initial ChIRP samples (Fig 2D, panel DNA) and reproduced in independent replicates. Interestingly, this region contains a shared core muscle enhancer (CME; Alzhano et al., 2010) required for muscle-specific activation of Igf2 and H19 (Eun et al., 2013), which are both decreased upon Charme knockdown (Table EV1). Interestingly, the knockout of part of this enhancer was previously shown to produce the loss of Igf2 expression and defects in muscle differentiation (Kaffer et al., 2001). Therefore, for subsequent analysis we concentrated on this specific locus.

The physical association of pCharme RNA with nctc (98 Mb far from Charme locus) was validated by FISH analysis with probes against Charme RNA (red) and BAC probes against the target loci (green). A detailed quantification of the signals indicated that in more than 60% of the nuclei Charme is in close proximity with the nctc region, with 2/3 of the signals showing partial overlapping (Fig 2E). No proximity was instead found when the unrelated Inc-31 locus (Ballarino et al., 2015), which was absent in the ChIRP-seq dataset, was used as negative control (Fig 2E).

Notably, the surrounding 2.5 Mb of the ChIRP peak in the nctc region contains relevant myogenic genes, such as Igf2, Tnni3, and Tnni2, that were downregulated upon Charme depletion in the RNA-seq dataset (Tables EV1 and EV2). qRT–PCR confirmed their downregulation upon Charme depletion both at the mRNA and pre-mRNA levels (Fig 2F). Along this line, ChIP analyses performed in conditions of Charme downregulation showed a strong decrease in the association of the RNA polymerase II (RNA Pol II) and in the acetylation of lysine K9 histone 3 (H3K9ac) of these genes (Fig 2G). Altogether, these data indicate that the effect of Charme on these target genes is exerted at the transcriptional level.

Due to the strong accumulation of pCharme at the sites of its own transcription, we tested whether the Charme genomic locus itself could be in contact with nctc. To this purpose, FISH analyses were performed with DNA BAC probes flanking the Charme gene (red) or overlapping the nctc region (green; Figs 3 and EV2A–C). In GM conditions, when Charme is not expressed, the two districts are clearly distinguished as independent spots (Fig 3A, left panel, and Fig EV2C, GM). Notably, upon Charme expression, the two regions get close (Fig EV2C, DM1 and DM1.5, and Fig EV2D) with a consistent percentage of chromosomes showing overlapping or proximal signals at days 1 and 1.5 of differentiation (Fig 3A, right panels, and Fig EV2C). As control, no proximity was found with the unrelated Inc-31 locus (Figs 3A and EV2C, lower panels). Notably, the pCharme-mediated long-range interaction was significantly reduced when Charme was downregulated by 50% through GAP-2 treatment (Figs 3B and EV2E). Overall, these in vitro data suggest an important role of pCharme in a long-range physical interaction between two chromatin districts distantly located, but present on the same chromosome.
Figure 3.
A further evidence for Charme acting in the nucleus at the sites of its own transcription derives from experiments where we tried to rescue the phenotype of Charme depletion with a cDNA construct expressing a GAPmer-resistant mCharme. This construct failed to recover the expression of myogenic markers (Fig EV2F) and produced a transcript with a cytoplasmic-restricted localization (Fig EV2G). The lack of rescue with cytoplasmic RNA further supports the finding that the active Charme species is the nuclear one; moreover, these results are in line with the accepted notion that for chromatin-associated lncRNAs acting in cis, rescue phenotypes are not obtained with exogenous gene overexpression (Goff & Rinn, 2015; Wang et al., 2015).

**Charme functional knockout in mice affects the myogenic process**

In consideration of the interesting function of Charme observed in vitro, the study of its function in vivo was strongly demanded. Therefore, the CRISPR/Cas9 methodology was used to generate a Charme null allele in mice. In order to avoid any structural alteration of the locus, we inserted at the beginning of exon 2 the strong poly(A) polyadenylation signal of the β-globin gene (Levitt et al., 1989) followed by two MAZ sites, known to promote efficient cleavage and transcriptional termination (Yonaha & Proudfoot, 2000; Gromak et al., 2006; Fig 4A). The correct genomic editing was confirmed in a founder mouse, and the Charme null allele was backcrossed for four generations to obtain homozygous animals (Fig EV3A). Charme−/− mice are viable, fertile, and produced in Mendelian ratios. Analysis of skeletal muscles of 1-month-old mutant and control animals indicated that the insertion of the poly(A)MAZ cassette resulted in a very effective depletion of both pCharme and mCharme isoforms (Fig 4B). The histological analysis of skeletal muscles from Charme−/− mice revealed a slight but significant reduction in the area of the fiber sections (Fig 4C), which is the most common parameter used to assess correct tissue morphology. Moreover, as a further evidence for delayed differentiation, a significant decrease in the mRNAs of myogenic markers myosin creatine kinase (MCK) and myosin heavy chain (MHC) was observed (Fig 4D). These features appeared consistent with the reduction in the fiber caliber of differentiating myotubes and the decrease in myoblast fusion observed in vitro in C2C12 cells (Fig 1G). Moreover, the Charme targets located in the nctc region identified in C2C12 cells were also specifically affected in murine muscle tissues (Fig 4E).

**Charme−/− mice exhibit an altered cardiac phenotype**

Since Charme also displays a cardiac restricted expression (Ballarino et al., 2015), we analyzed the effects of Charme knock-out in heart. A peculiar phenotype was detected when analyzing the heart of Charme−/− mice where the expression of both pCharme and mCharme isoforms was almost abolished by the poly(A) site insertion (Fig 5A). Morphometric analyses, performed with hematoxylin/eosin staining in both adult (Fig 5B) and neonatal (Fig EV3B) cardiac sections, showed in all the mice tested an increase in the heart area (Ha), a definite decrease in ventricular chamber dimensions (Lva and Rva), and an increase in the thickness of the ventricular free walls (Lvw and Rvw) and interventricular septum (IVS; Fig EV3C). Moreover, as expected from this altered morphology, higher blood pressures (DBP and SBP) and heart-to-body weight ratio (Hw/Bw) were also detected in Charme−− mice (Fig EV3D). Notably, compared to wild-type littermates, which have a life expectancy of 24–38 months, none of the available Charme−− mice (n = 5, P < 0.001) survived for more than 1 year, suggesting a midlife lethality caused by Charme ablation. In analogy to skeletal muscles, the histological analysis performed on Charme−− cardiac tissues revealed a reduction in cell area (Fig 5C) together with the downregulation of terminal differentiation markers (Fig 5D). To note, the expression of Igf2, the only gene of the nctc region expressed in the cardiac tissue, was also found to be downregulated. Notably, in line with the in vitro data, DNA/DNA FISH experiments performed in neonatal hearts proved that the co-localization between Charme and the nctc regions is also occurring in vivo while it is impaired in Charme−−− animals (Fig 5E). Interestingly, the proximity was also found to persist in adult hearts (Fig EV3E).

**Identification of a functionally conserved human Charme transcript**

An evolutionary conservation analysis performed by examining RNA-seq (Legrini et al., 2017) and FANTOM5 (Noguchi et al., 2017) data revealed the existence of a Charme orthologue transcript (hs-Charme) produced by the human syntenic locus (Fig EV4A), with a structure very similar to the murine counterpart (Figs 6A and EV4B–D, and Table EV3). Notably, the hs-Charme locus displays an overall ~ 45% sequence identity (Fig EV4E and Appendix Table S1), which is a high value for a lncRNA genomic region (Pang et al., 2006). Similarly to the murine counterpart (Ballarino et al., 2015), the expression of the early myogenic transcriptional factor MyoD was found to precede that of hs-Charme (Fig EV5A, left panel; and Table EV3) and a specific peak of MyoD ChIP-seq (MacQuarrie et al., 2013) was identified, as in mouse, at the level of MyoD canonical E-boxes upstream to the TSS (Fig EV5B and C). Altogether, these data suggest a common regulatory control of MyoD on mouse and human Charme RNA expression. Indeed, both hs-pCharme and hs-mCharme were found expressed in human primary myoblasts upon induction of differentiation; moreover, both transcripts resulted in delayed expression in myoblasts from Duchenne muscular dystrophy patients (Figs 6B and EV5A, right panel; and Table EV3), where MyoD expression is strongly impaired and cells are blocked in a more...
Figure 4. Generation and skeletal muscle characterization of Charme<sup>−/−</sup> mice.

A Top: Charme locus is shown together with the position of sgRNA1 and sgRNA2 divergent single guides. Bottom: design of the single-stranded oligodeoxynucleotide (ssODN) containing the two homology arms (gray line) and the 100-nt-long poly(A)/MAZ cassette.

B qRT–PCR quantification of mCharme and pCharme expression in skeletal homogenates from Charme<sup>+/+</sup> and Charme<sup>−/−</sup> mice. The number of mice tested is the same for the two groups, and it is indicated in the white bars. Data were normalized to GAPDH mRNA and represent mean ± SEM.

C Dystrophin staining analyzed by immunofluorescence on Charme<sup>+/+</sup> and Charme<sup>−/−</sup> skeletal (gastrocnemius) biopsies. The mean ± SD of cross-sectional area (µm²) is reported in the histograms besides. In total, N = 667 and 1,039 fibers per condition were analyzed. Original magnification = ×10. The number of mice tested is the same for the two groups, and it is indicated in the white bars. Scale bar = 100 µm.

D qRT–PCR quantification of MHC and MCK transcripts in skeletal (gastrocnemius) homogenates from Charme<sup>+/+</sup> and Charme<sup>−/−</sup> mice. The number of mice tested is the same for the two groups, and it is indicated in the white bars. Data were normalized to GAPDH mRNA and represent mean ± SEM.

E qRT–PCR quantification of Tnnt3, Tnni2, and Igf2 transcripts in skeletal (gastrocnemius) homogenates from Charme<sup>+/+</sup> and Charme<sup>−/−</sup> mice. Data were normalized to GAPDH mRNA and represent mean ± SEM. The number of mice tested is the same for the two groups, and it is indicated in the white bars.

Data information: *P < 0.05, **P < 0.01, ***P < 0.001, unpaired Student’s t-test.

Source data are available online for this figure.
Figure 5.
Data support the notion of an important conserved function of Charme and its depletion affects the same subset of myogenic genes. These results can be considered as discrete regulatory units that are regulated by a conserved mechanism across species. The identification of these conserved regions in both human and mouse lines provides insights into the evolutionary conservation of the Charme lncRNA functions in myogenesis.

**Discussion**

In this study, we show the characterization of Charme, a chromatin-associated lncRNA induced upon myogenic differentiation. In vitro Charme knockdown causes defects in myoblast fusion and downregulation of several crucial myogenic genes. Among them, several have been previously shown to be involved in familial hypertrophic cardiomyopathy (HCM), dilated (DCM), and arrhythmogenic right ventricular (ARVC) cardiomyopathies (Maron et al., 2001; Alzhanov & Rotwein, 2016). The nctc region also contains genes involved in muscle activity and metabolism, such as muscle troponins (Tnnt2 and Tntnt3). Interestingly, all these genes are affected by Charme depletion, indicating a direct connection between the lncRNA expression and the muscle-specific transcriptional output of the region. We also show that Charme, retained at the sites of its own transcription as an incompletely spliced form, is required for the maintenance of the long-distance chromosomal interaction with the nctc locus and for consolidating the expression of the myogenic genes therein included. These data are in line with recent studies indicating the importance of the tridimensional chromatin folding and with the evidence that the eukaryotic genome is organized in a hierarchy of well-defined, highly compartmentalized, and extremely dynamic structures within the nucleus, where chromosome territories are segmented into functional compartments (Rowley & Corces, 2016). These regions can be considered as discrete regulatory units where genes far apart in terms of linear distance but close in space are activated or repressed in a coordinated fashion during cell differentiation (Zaug & Ceci, 1980; Guerrier-Takaoka et al., 1983).

Notably, Charme depletion in vivo produced similar molecular and morphological phenotypes as the one observed in vitro. In fact, particularly interesting because it is localized in cis with respect to the Charme locus and harbors a shared core muscle enhancer (CME; Alzhanov et al., 2010), previously shown to control muscle-specific activation of the neighboring Igf2 and H19 promoters (Eun et al., 2013). Moreover, the knockout of part of the nctc region was performed by several groups and resulted in the loss of Igf2 expression and in defective myogenesis (Kaffer et al., 2001; Alzhanov & Rotwein, 2016). The nctc region also contains genes involved in muscle activity and metabolism, such as muscle troponins (Tnnt2 and Tntnt3). Interestingly, all these genes are affected by Charme depletion, indicating a direct connection between the lncRNA expression and the muscle-specific transcriptional output of the region. We also show that Charme, retained at the sites of its own transcription as an incompletely spliced form, is required for the maintenance of the long-distance chromosomal interaction with the nctc locus and for consolidating the expression of the myogenic genes therein included. These data are in line with recent studies indicating the importance of the tridimensional chromatin folding and with the evidence that the eukaryotic genome is organized in a hierarchy of well-defined, highly compartmentalized, and extremely dynamic structures within the nucleus, where chromosome territories are segmented into functional compartments (Rowley & Corces, 2016). These regions can be considered as discrete regulatory units where genes far apart in terms of linear distance but close in space are activated or repressed in a coordinated fashion during cell differentiation (Zaug & Ceci, 1980; Guerrier-Takaoka et al., 1983).
Figure 6.
the same genes affected in C2C12 cells were downregulated in the skeletal muscles of Charme−/− mice and the decrease in fiber calibers observed in adult muscles correlated with the reduction in the caliber of differentiating myotubes observed in vitro. Charme−/− mice showed also a clear cardiac phenotype, characterized by a conspicuous decrease in fiber size, and resulting in an overall heart remodeling and alteration of morphological and functional parameters. Also in this district, Charme was shown to contribute to the physical proximity with the nctc genomic region and to affect the cardiac expression of Igf2. Since Charme depletion provides a clear heart phenotype already at neonatal stages, the observed remodeling of the cardiac tissue must represent outcomes of defects starting early during embryonic development and persisting in adult life. The reduction in life span observed in adult mice indicated the persistence of effects of Charme depletion even at very advanced ages. It is important to note that this is the first time that such a cardiac phenotype has been directly ascribed to a lncRNA ablation in the absence of any stressful condition (Guttman et al., 2010; Devaux et al., 2015). Moreover, the chronic in vivo phenotype, where the overall alteration of tissue homeostasis and cardiac structure leads to late lethality, correlates well with the role of Charme as a fine-tuning regulator.

Since Charme is interacting with several other regions besides the nctc locus, it is important to underline that these other domains will likely contribute to the overall phenotype observed in vitro and in vivo. For the nctc locus, which has been prioritized in this study, a clear correlation between the expression of the myogenic genes therein included and Charme activity has been assessed (Fig 6F). Further characterization of the other Charme targets will be necessary to complete the genome-wide clarification of its function in skeletal and cardiac muscles.

Materials and Methods

Cell culture conditions and transfection

C2C12 murine myoblasts were cultured as previously described (Agostini et al., 2013) and differentiated in 0.5% fetal bovine serum (FBS). Human control healthy and DMD myoblasts from the Telethon Neuromuscular Biobank were cultured as previously described (Faggioli et al., 2008) and differentiated with human skeletal muscle differentiation medium (PromoCell). RNAi experiments were performed with 75 nM of LNA GAPmers (Exiqon) using Lipofectamine 2000 (Invitrogen). GAPmers were designed against Charme exon 2 (GAP-2) or EX2/EX3 splice junction (GAP-2/3) sequences using the Exiqon web tool. Sequences are reported in Table EV4.

RNA analyses

Total RNA from myoblasts/myotubes was isolated using QIAzol Lysis Reagent (Qiagen)/chloroform extraction followed by spin-column purification (RNeasy Mini Kit; Qiagen) and quantification by NanoDrop (Thermo Scientific). Northern blot analysis was performed on 12 µg of total RNA from mouse myoblasts/myotubes. Total RNA was denatured in glyoxal loading dye (Ambion) for 30 min at 50°C and loaded on 1.2% agarose gel. Electrophoresis was carried out for 2 h at 60 V. RNA was transferred overnight on Hybond-N+ membrane (GE Healthcare) by capillarity in 10× SSC. The transferred RNA was UV-cross-linked (1,200 × 100 µl/cm²), and the membrane was washed in 50 mM Tris (pH 8) for 20 min at 45°C. Pre-hybridization and hybridization were performed in NortherMax buffer (Ambion) at 68°C (30 min and overnight, respectively). For the hybridization, 500 ng of DIG-labeled probes (DIG-RNA Labeling Mix; Roche) was produced by in vitro transcription with T7 RNA polymerase (Promega) using a PCR-amplified Charme template (see Table EV4 for oligo sequences). The membrane was washed twice in 2× SSC/0.1% SDS and twice in 0.2× SSC/0.1% SDS at 68°C. The membrane was finally processed for DIG detection using the DIG Luminescence Detection Kit (Roche), according to the manufacturer’s instructions. Nucleoplasm/chromatin/cytoplasmic fractionation was performed using the NE-PER Kit (Pierce-Thermo Scientific) following the manufacturer’s specifications. 5′-RACE was performed using FirstChoice RLM-RACE kit (Ambion) on 5 µg of total RNA and by PCR amplification with Charme 4 and Charme 3 reverse primers. For semiquantitative and quantitative PCR analyses, RNA (0.5–1.0 µg) was reverse-transcribed using SuperScript III RT (Life Technologies) according to the manufacturer’s instructions and amplified by PCR using MyTaq (Bioline) enzyme. Real-Time PCRs (qPCRs) were performed using QuantiTect SYBR Green (Qiagen).

RNA-seq and functional enrichment analysis

RNA-seq data were analyzed using RAP (D’Antonio et al., 2015) using FASTQC (www.bioinformatics.babraham.ac.uk), NGS QC Toolkit (Patel & Jain, 2012), Bowtie 2 (Langmead & Salzberg, 2012), TopHat2 (Kim et al., 2013), and Cufflinks and Cuffdiff (Trapnell et al., 2012, 2013). Default parameters were used for the analyses except for library type (fr-firststrand), mate inner distance (set to 0.5, and variation of gene expression in the same direction. Functional enrichment analysis was performed by FIDEA (D’Andrea et al., 2013) using the genes expressed in the scramble condition as background.

RNA-seq data from GSE70389 (Legnini et al., 2017) were analyzed as described above. The read coverage and the Cufflinks tracks were loaded on IGV genome browser (Robinson et al., 2011; Thorvaldsdóttir et al., 2013); see also Fig EV4B.

Chromatin isolation by RNA purification (ChIRP) sequencing

ChIRP-seq was performed according to the procedure described by Chang (Chu et al., 2011) using nuclear extract from 10² differentiated myotubes/each sample. Charme pull-down was performed with different sets of oligos (Appendix Fig S2F) or lacZ as control. Input and purified DNA were sequenced with HiSeq 2500 (40 million single-end, 50-base pair-long reads/sample). Reads were preprocessed with Trimomatic (Bolger et al., 2014) and Cutadapt (Martin, 2011) to trim the 5′ and 3′ adapters and to remove the first five nucleotides, whose average quality score was lower than that of the other nucleotides. Low-complexity reads were then removed
using the DUST algorithm included in the PRINSEQ suite (Fatica & Bozzone, 2014; Morlando et al., 2014; complexity score threshold = 7). The procedure adopted by Chu and colleagues (Chu et al., 2011) was used to find the EVEN and ODD common peaks (mm9 assembly, MACS14 bw = 300 (Zhang et al., 2008), fold enrichment/input > 2, average coverage > 10, and Pearson correlation > 0.3). BEDTools (Quinlan & Hall, 2010) was used to identify the overlapping lacZ peaks (obtained with MACS14 using the input as background). Peaks with fold enrichment/input 25 times lower than that of the overlapping lacZ peaks and with more than 50% of the length covered by repeats (Chen, 2002; http://www.repeatmasker.org) were excluded. HOMER (Heinz et al., 2010) was employed to annotate peaks using GENCODE M4 (Harrow et al., 2012). WindowBed tool was used to find all downregulated genes which are closer than 2.5 Mb to Charme ChiRP peaks.

Immunofluorescence and DNA/RNA FISH

MHC staining and dystrophin staining were performed as previously described (Cazzella et al., 2012). Samples were imaged on inverted microscope Zeiss AxioObserver A1 equipped with AxioCam MRm R camera and Plan-Neofluar EC 10 ×/0.3 M27 and LD 20 ×/0.4 Korr objectives. Images were acquired with AxioVision Rel.4.8 software.

Fluorescent in situ hybridization (FISH) was carried out on cells that adhered to collagen (Corning) using fluorescent (Fluorescein or Cy3)-conjugated nick-translated BAC DNA (DNA FISH) or synthetic DNA oligonucleotides (RNA FISH). Three-dimensional (3D) DNA FISH on C2C12 was performed with some modifications of an already described procedure (Cremer et al., 2008). Briefly, cells were fixed in 4% PFA (20 min at 4°C) and pre-processed by freeze-thawing permeabilization to ensure the preservation of nuclear structures. After denaturation, FISH probes were hybridized O.N. at 37°C and samples washed, DAPI-stained, and mounted in ProcLong Diamond reagent (Thermo Fisher Scientific). For DNA FISH on histological muscle sections, after permeabilization with Triton buffer (0.5% Triton X-100/PBS), a mild pepsin digestion (pepsin 0.1% /0.1 M HCl) was applied before denaturation in order to reduce the autofluorescence deriving from protein components. For RNA FISH, fixed cells were permeabilized in 0.5% Triton (10 min at 4°C) and incubated O.N. at 37°C with fluorescent DNA probes diluted in hybridization buffer (10% dextran sulfate, 2× SSC, 10% formamide, 2 mM ribonucleoside-vanadyl complex). The specificity of RNA FISH was tested by pre-incubating the cells with 200 µg/ml RNaseA for 1 h at 37°C, or by GAPnmer transfection. Sequential hybridizations were used for DNA/RNA FISH. Samples were imaged on inverted microscope (Olympus IX73) equipped with a Confocal Imager (CREST X-LIGHT) spinning disk, a CoolSNAP MYO CCD camera (Photometrics), and a Lumencor Spectra X LED illumination. Images were acquired with 60× NA1.35 oil objective (UPLAN-SApo) and MetaMorph (Molecular Devices) using 300 (Cy3), 500 (Fluorescein), or 100 (DAPI) ms as exposure time.

Immunofluorescence and RNA and DNA FISH post-acquisition analyses

The fusion index (F.I.) indicates the percentage of nuclei in multinucleated myotubes (defined by at least two nuclei) divided by total number of nuclei in a microscop field; the average number of nuclei inside individual myotubes (Nu/MT) was determined by dividing the total number of nuclei by the number of myotubes; and the fraction of mononucleated MHC-positive cells (MonoMHC+) was calculated as the mean percentage of nuclei in MHC-positive cells divided by total number of nuclei in each field. The number of nuclei was estimated by the average of nuclei counted in three independent and randomly chosen microscope fields. Dystrophin staining was used to evidence the border of the muscle fibers. The single-channel dystrophin images were merged in dystrophin/DAPI images with ImageJ software and then used to quantify the fiber number manually. The fiber cross-sectional area analysis was performed on dystrophin immunofluorescence by using a semi-automatic tool (Briquet et al., 2004; Mula et al., 2013). The immunofluorescence images were binarized to obtain a mask of fiber boundaries and then applied to originals for the automatic measure of the areas. Statistical analysis of fiber cross-sectional area (CSA) was performed using unpaired t-test, and the differences between means were considered significant at P ≤ 0.05.

For DNA/RNA FISH staining analysis, the Z-stack confocal microscopy images were taken automatically (0.2-µm path) and merged with maximum-intensity projection or processed for image analysis. A Laplacian-of-Gaussian filter was applied to 3D volume after image acquisition to enhance the spots signal over background (Raj et al., 2008). Intensity threshold was manually adjusted using MetaMorph or ImageJ software. Signal from cells hybridized without probes was considered as background.

Spatial proximity analysis of long-split, paired, or overlapped patterns in DNA FISH was performed by computing the 3D interal-lelic distance on Z-stack acquisitions using ImageJ (Spot Distance; Schober et al., 2008). Changes in the C2C12 (elliptic) shape of nuclei occurring along differentiation were taken into account as in Cremer et al. (1984) before the calculation of the Normalized Distances mean (ND; ND = Interallelic Distances/d, d = major axis + minor axis/2). Values above or below the paired ND range were attributed to long-splitted or overlapped patterns, respectively (Fig EV2A). The co-planarity and the co-localization of the DNA and RNA spots were established by recording the main grayscale value (expressed as arbitrary units) along Z-planes using ImageJ (Plot Z-axis profile). A Gaussian function (determination coefficient > 0.9) was used to fit the fluorescent intensities and to extrapolate the main Z-plane value, which corresponds to the maximum fluorescence value of green and red channels in DNA FISH staining (Fig EV2B).

The percentage of nuclei showing spatial proximity of genomic loci was calculated by manual screening on binarized images on a total of about 500 nuclei showing at least one pair of spots with paired or overlapped patterns. To determine spot intensity (fluorescence intensity) in RNA FISH analysis, we measured the mean gray value of nonbinarized images within a selection that circumscribes the spot; this value was divided by the diameter of the spot (d = 2*(3/4πV/π1/3)) by applying the formula of a sphere with the same volume of the spot.

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed on chromatin extracts according to the manufacturer’s protocol (MAgnify ChIP; Life Technologies) by O.N. incubation with 2 µg of immobilized monoclonal anti-RNA
Pol II (Millipore), anti-acetyl-histone H3 (Lys9; Millipore), or rabbit IgG antibodies. A standard curve was generated for each primer pair testing 5-point dilutions of input sample. Fold enrichment was quantified using qRTPCR (SYBR Green; Qiagen) and calculated as a percentage of input chromatin (% Imp). Data from GAP-sr vs. GAP-2 conditions were normalized to an unrelated genomic region and represent the mean of three experiments ± SEM.

**Generation of Charme−/− mice**

Charme−/− mice were generated in the C57BL/6 background using a CRISPR genome-editing system (Yang et al, 2014). A pair of divergent single guide RNAs (sgRNAs) inside exon 2 was designed using the http://crisp.mit.edu resource and cloned into the pX330 vector (Addgene). The sgRNAs were amplified together with the T7 promoter to generate templates for in vitro transduction. For homologous recombination, a ssODN (200 nt) containing two homology arms (50 nt each), a short poly(A) site (49 nt), and two MAZ sites (Levitt et al, 1989; Yonaha & Proudfoot, 2000) was synthesized by IDT. In vitro transcribed Cas9 mRNA and sgRNAs were injected into the cytoplasm of fertilized eggs of C57BL/6 mice in M2 medium (Sigma-Aldrich, MO, USA). Successful knockdown was validated by PCR analyses (see Table EV4 for oligo sequences). The wild-type allele yielded an amplicon of 235 bp, whereas the mutated allele yielded an amplicon of 736 bp (FP7/2007-2013) to D'O.C.; and Sapienza University (prot. RM1715S7C8176C1) to M.B.

**Offspring genotyping**

DNA extraction from tail biopsies was performed using RBC Real Genomics DNA Extraction Kit (RBC Bioscience) according to the manufacturer’s protocol. The poly(A)/MAZ insertion was detected by PCR amplifications with oligo pairs int1b-3 (Table EV4).

**Histological analysis**

Mice were sacrificed at different ages, and the morphological changes in the skeletal and cardiac muscles were observed under light microscopy. Muscle tissues were fixed in 4% formaldehyde, and sections were stained with hematoxylin and eosin.

**Study approval**

Animals were treated in respect of housing, nutrition, and care according to the guidelines of Good Laboratory Practice (GLP). MB and RT have successfully completed the Module1 as co-coordinated by the Administration of Animal Facilities in the premises of EMBL. The EMBL course meets the requirements of European legislation on basic training of researchers.

**Data availability**

Data that support the findings of this study have been deposited in GEO with the primary accession code GSE94498 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE94498).

**Expanded View** for this article is available online.

**Acknowledgements**

The authors acknowledge O. Sthandier and M. Marchioni for technical support, L. Biscegie for support in ChIRP analysis, and the EuroBioBank and Telethon Network of Genetic Biobanks (GB12001F) for providing biological samples. This work was partially supported by grants from ERC-2013 (AdG 340172–MUNCODD), Telethon (GGP16213), Human Frontiers Science Program Award RGPO009/2014, Parent Project Italia, AFM-Telethon (17835), Epigen-Epigenomics Flagship Project, and AriSLA full grant 2014 “ARCI” to I.B.; Telethon GGP14066 to A.M.; Fondazione Roma to I.B. and A.M.; ERC grant agreement no. GA 310206 (FP7/2007-2013) to D'O.C.; and Sapienza University (prot. RM1715S7C8176C1) to M.B.

**Author contributions**

IB conceived the study, wrote the manuscript, and secured funding; MB designed the study, wrote the manuscript, and conducted ChIRP and ChIP analyses; ACi performed most of the experiments and the computational analysis of RNA-seq; CC and DO performed the ChIP analyses; FD performed the Northern blot analyses; ACo performed the computational analysis of ChIRP-seq; MM provided expertise on myogenic field; and CN and AM provided expertise and feedback on the experiments in mice.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**References**


Becker JR, Deo RC, Werdich AA, Panâková D, Coy S, MacRae CA (2011) Human cardiomyopathy mutations induce myocyte hyperplasia and activate...
hypertrophic pathways during cardiogenesis in zebrafish. Dis Model Mech 4: 400 – 410
Charme LncRNA functions in myogenesis  Monica Ballarino et al


License: This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.