Expanded View Figures

Figure EV1. Genomic features and expression analysis of the murine *Charme* locus.

A UCSC visualization showing the chromosome position and the genomic coordinates of *Charme* (highlighted in yellow) in the mm9 mouse genome. Red box indicates the genomic position of the previously annotated 5430431A17Rik transcript.

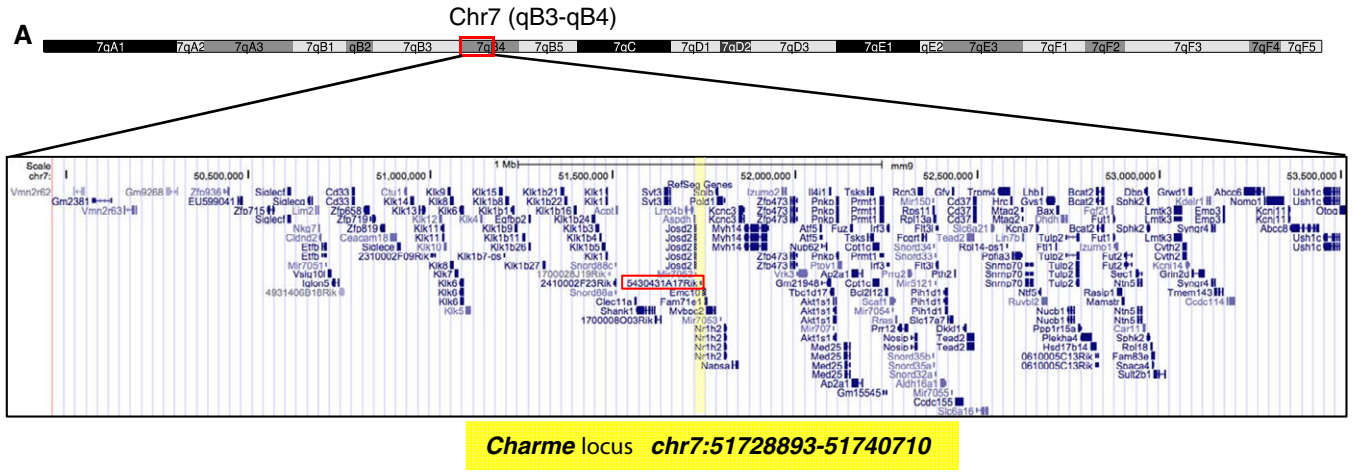
B Left: 5′-RACE analyses of *Charme* in proliferating (GM) and differentiating (DM) conditions. The position of the Racer primer is indicated. RNA treated to the same experimental regime, but with tobacco acid pyrophosphatase digestion omitted (−TAP), did not generate a product. The PCR products were sequenced, and the identified *Charme* transcription start site is reported (+1). Middle: Northern blot analysis of *Charme* on total RNA from proliferating (GM), differentiating (DM, GAP-scr), and *Charme*-depleted (DM, GAP-2) myotubes. The reduced level of the signal upon *Charme* depletion confirms the specificity of the observed band. 18S rRNA and 28S rRNA serve as a loading control. Right: RT–PCR quantification of *Charme* in cytoplasmic (Cyt), nuclear (Nu), nucleoplasmic (Np), or chromatin (Chr) fractions from differentiated myotubes. The quality of fractionation is shown in Fig 1B. Three biological replicates were analyzed, and a representative experiment is shown. −, RT-minus control.

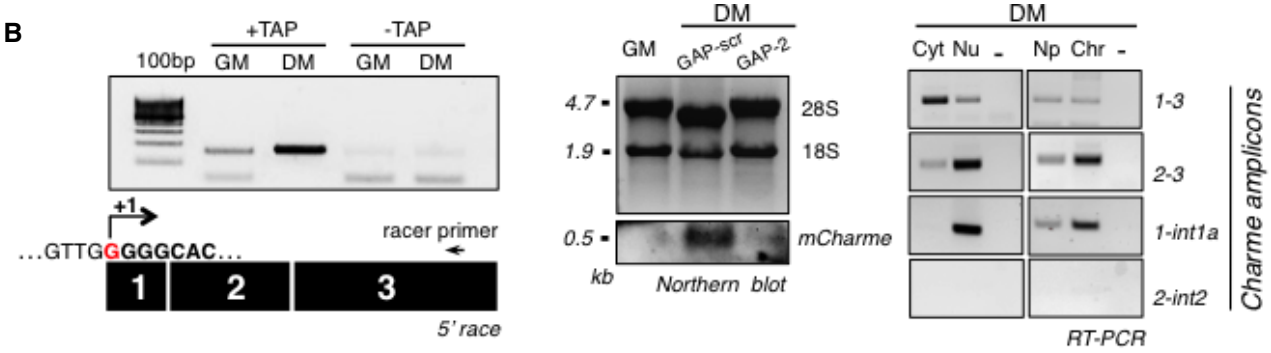
C RNA-seq coverage visualization of *Charme* locus during the time-course of C_2C_{12} differentiation. A quantification of the intronic reads is represented in the histogram aside. Normalized reads (NR) were obtained by dividing the total number of intronic reads for their respective lengths.

D RT–PCR validation of *Charme* intron 1 retention performed on RNA from cytoplasmic or nuclear samples.

E Assessment of *Charme* isoform stability. RNA half-lives were calculated upon 0, 3, and 6 h of actinomycin D (ActD) treatment; *GAPDH* and *pre-GAPDH* serve as controls. Data represent the means ± SD from three independent experiments.

Source data are available online for this figure.





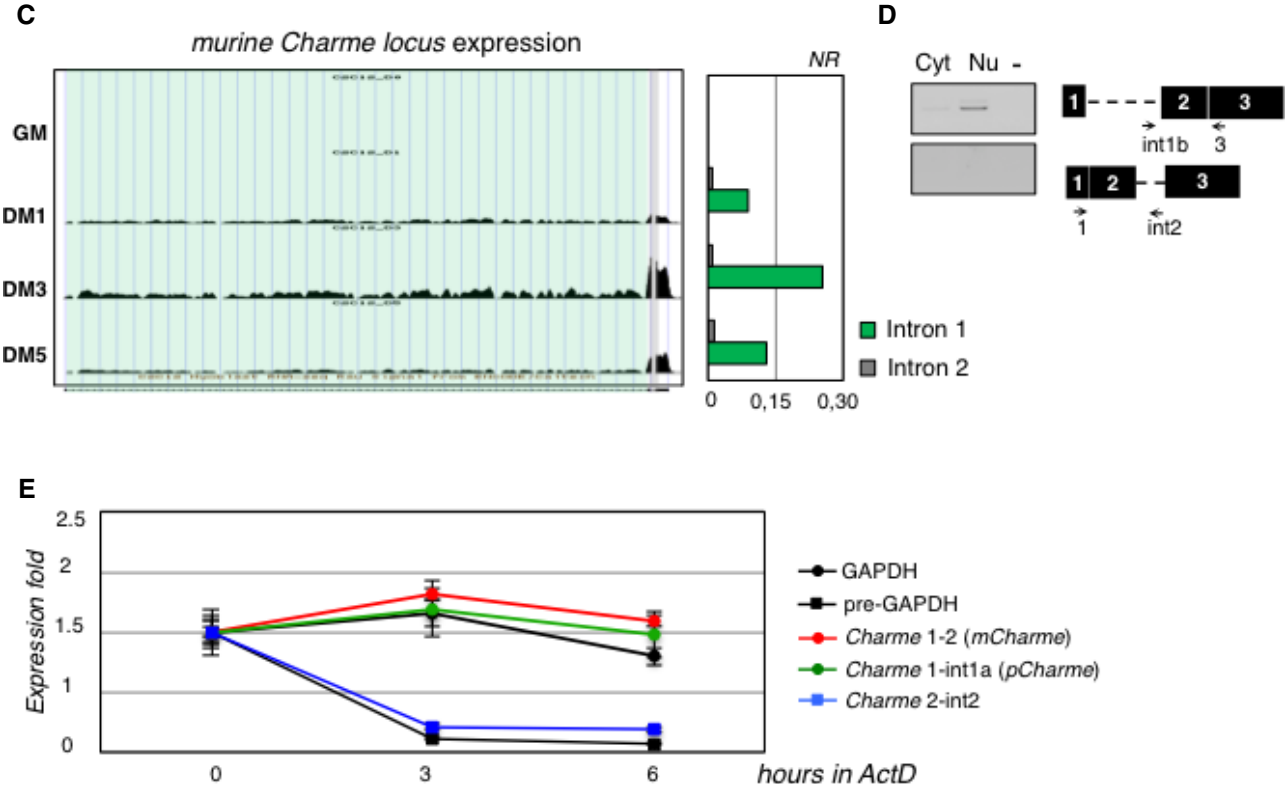


Figure EV1.

Figure EV2. Chromatin contacts and functional characterization of the murine *Charme* transcript.

A Examples of proximal *Charme*/*nctc* DNA FISH spots on the same Z-section. Red and green signals correspond to *Charme* or *nctc* chromatin regions, respectively. White boxes show binarized images of single focal planes containing long-splitted, paired, or overlapped patterns. On top of the panels, the 3D interallelic distance interval (ND) for each category is shown. ND values were obtained by normalization to the diameter of the nuclei. The number of green/red DNA spots per nucleus (n = 3) reflects the C_2C_{12} aneuploidy with three copies of chromosome 7. Scale bar = 5 μm.

B Plotting of Z-stack intensity distribution for each spot in the insets of (A). Gaussian fit curve (full line) for the data of fluorescence intensity (filled circles) in each channel is plotted along Z-planes (Z-step = 200 nm). The strong correspondence of green and red channel distribution indicates near co-localization and co-planarity of the signals.

C Full field view of the DNA FISH studies on the nuclear distribution of *Charme* and *nctc* (top) or *Charme* and *lnc-31* (bottom) loci. Scale bar = 5 μm.

D Normalized 3D distances between *Charme* and *nctc* loci at the indicated time points (GM, DM 1–1.5 days). Interallelic distances were normalized to nuclei diameter.

E Normalized 3D distances between *Charme* and *nctc* loci in GAP-scr or GAP-2 treated cells at the indicated differentiation times (DM 1–1.5 days). Interallelic distances were normalized to nuclei diameter. Mean ± SD ND values are shown. $^{**}P < 0.01$, $^{***}P < 0.001$, unpaired Student's t-test.

F Quantitative real-time RT–PCR analyses of *mCharme*, *MHC*, and *MCK* mRNAs in C_2C_{12} cells transfected with GAP-scr or GAP-1 in combination with a *mCharme* mutant (*mCharme*-mut) in the GAP-1 targeting site. Transfections with the empty vector were used as negative control. PCR data were normalized to *GAPDH* mRNA.

G sqRT–PCR quantification of *mCharme* in cytoplasmic (Cyt) and nuclear (Nu) fractions from C_2C_{12} cells transfected with *mCharme*-mut. The quality of fractionation was tested with mature (*GAPDH*) and precursor (*pre-GAPDH*) RNAs. −, RT-minus control.

Source data are available online for this figure.

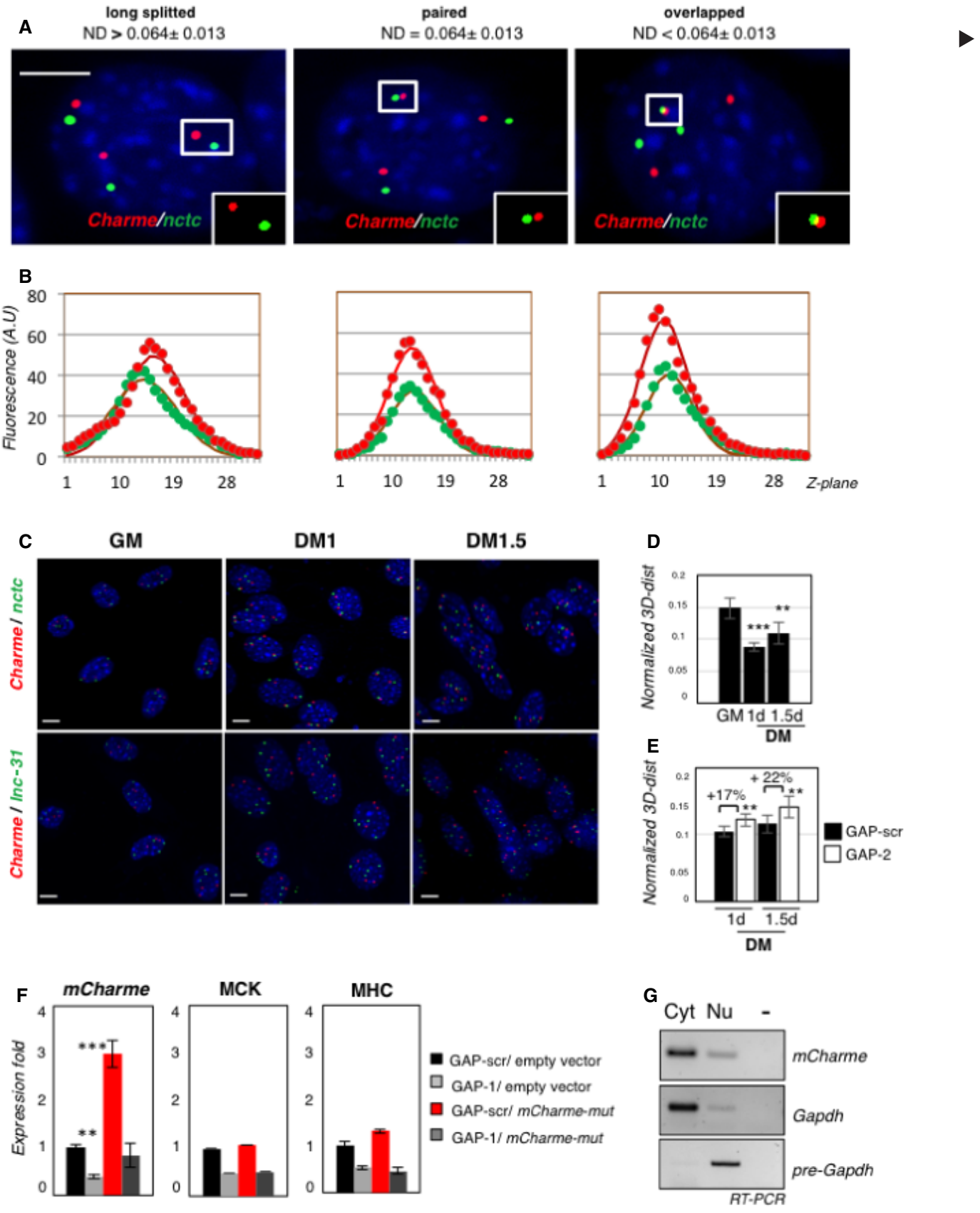


Figure EV2.

Figure EV3. Generation and functional characterization of *Charme*$^{-/-}$ mice.

A DNA extracted from F0-generation mice was analyzed by PCR amplification for genotyping. The position of PCR primers is indicated in Fig 1A. Electrophoretic analyses of PCR products and multiple sequence alignment (performed by Muscle-3.8) indicated the insertion of the entire poly(A)/2×MAZ cassette on a single allele in the #79, #80, and #82 pups (red). The #79 was selected as founder for successive breeding.

B Hematoxylin and eosin staining of transverse sections of the myocardium in 2-day-old *Charme*$^{+/+}$ (left) and *Charme*$^{-/-}$ (right) mice. Left (LV) and right (RV) ventricles are indicated. Scale bar = 0.5 mm.

C Morphometric analyses of heart area (Ha), left and right ventricle areas (LVa and RVa), left and right ventricle walls (LVw and RVw), and interventricular septum (IVS) in *Charme*$^{+/+}$ and *Charme*$^{-/-}$ mice.

D Systolic (SBP) and diastolic (DBP) blood pressures and heart/body weight ratio (Hw/Bw) in *Charme*$^{+/+}$ and *Charme*$^{-/-}$ mice. Values represent mean ± SEM of replicates. The numbers of mice tested for each group are indicated in the white bars. Mean ± SD ND values are shown. $^{**}P < 0.01$, $^{***}P < 0.001$, unpaired Student's *t*-test.

E DNA/DNA FISH in adult cardiac tissues. *Charme*$^{+/+}$ adult cardiac tissues exhibit spatial proximity of *Charme* and *nctc* genomic regions. The percentages of chromosome 7 showing paired and overlapped signals are 12% and 7%, respectively, in a total of 412 nuclei analyzed. Scale bar = 10 μm.

Source data are available online for this figure.

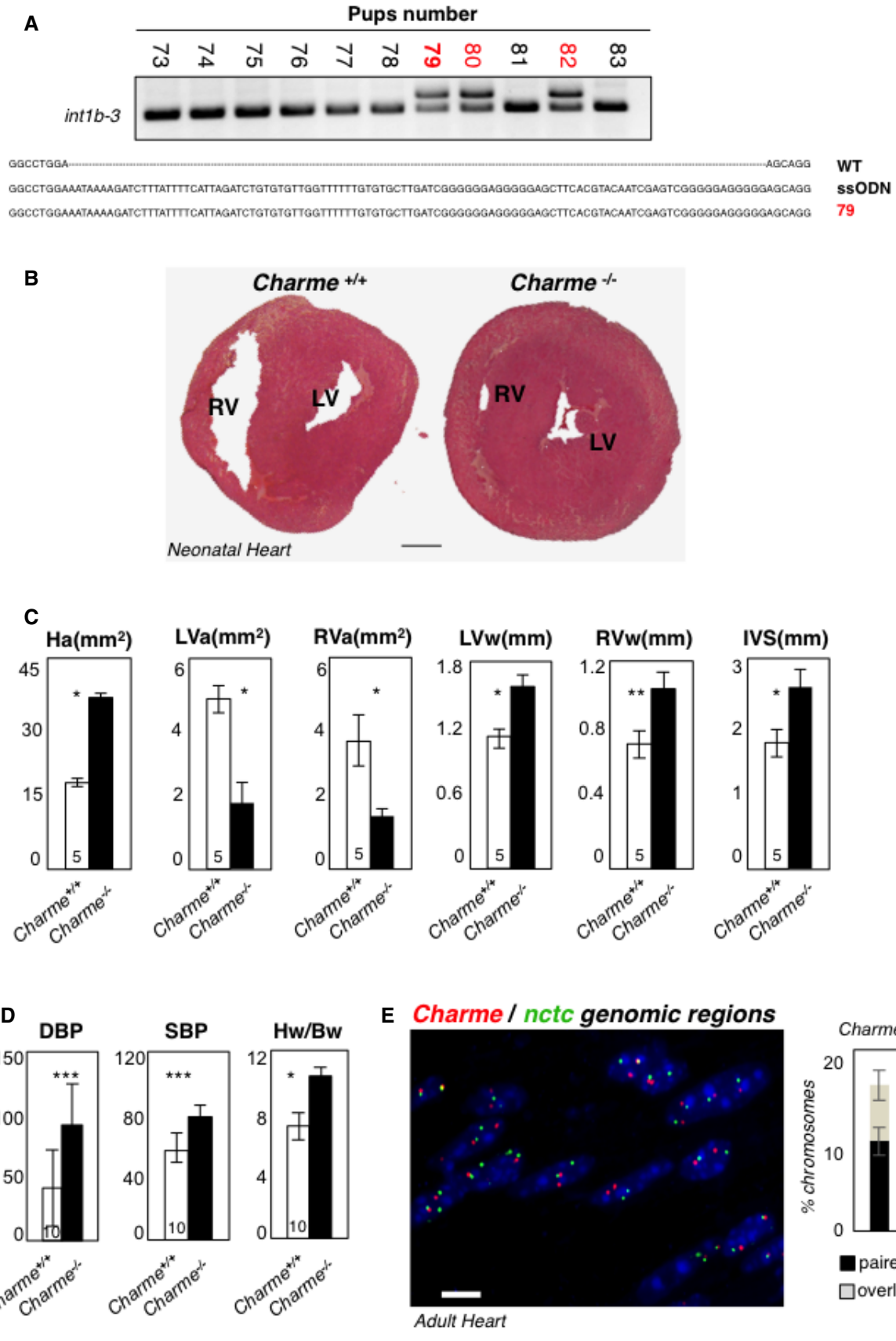


Figure EV3.

Figure EV4. Genomic features and expression analysis of the human *Charme* locus.

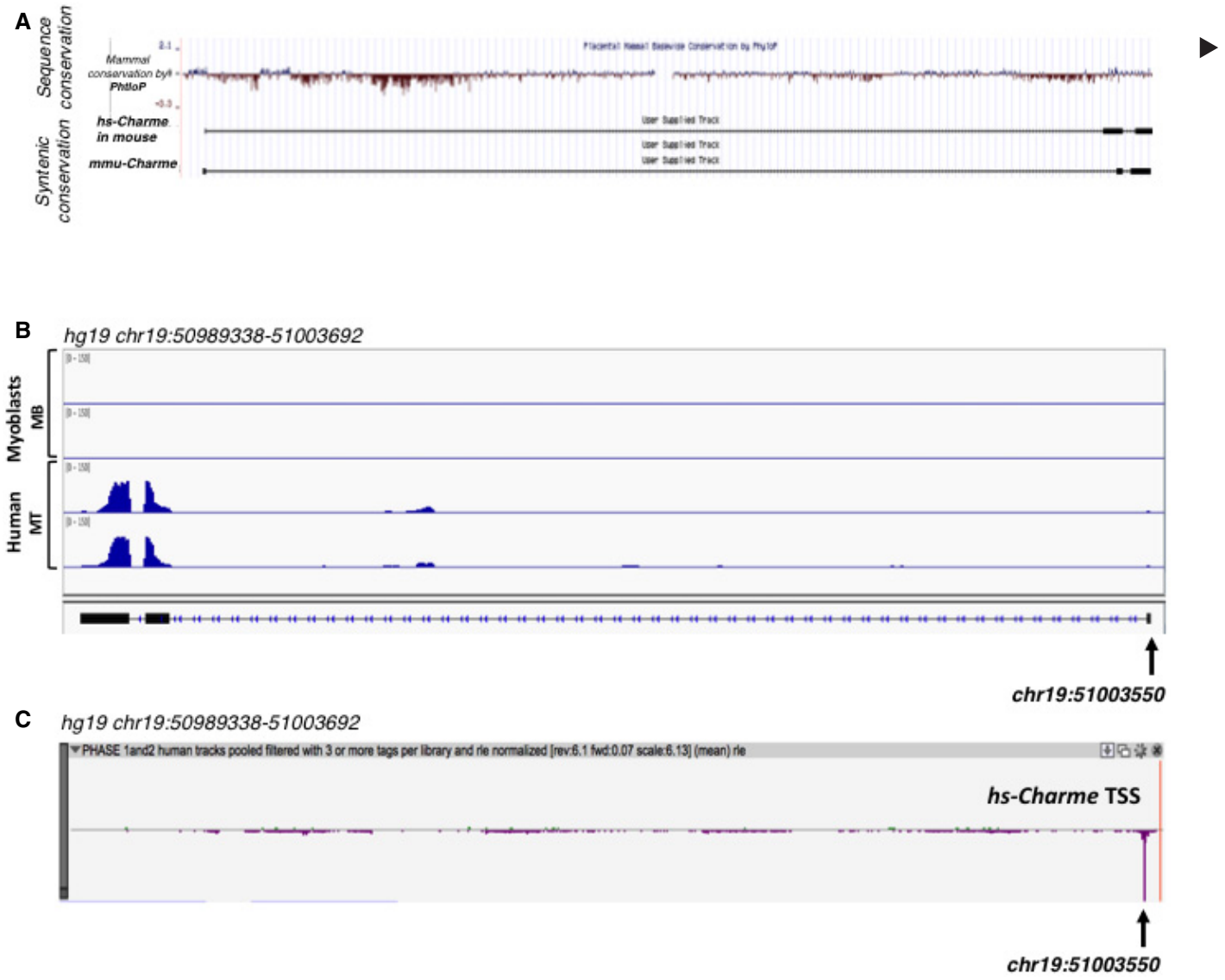
A UCSC visualization of the merged between human (hg19, top) and mouse (mm9, bottom) *Charme* genomic coordinates in the mm9 mouse genome. The analysis shows the mammal genomic sequence conservation (PhyloP) and the *Charme* transcript synteny between mouse and human.

B IGV (Robinson *et al*, 2011; Thorvaldsdóttir *et al*, 2013) visualization of the reads coverage from a RNA-seq experiment performed in human myoblasts (Legnini *et al*, 2017; GSE70389) in proliferating (MB) and differentiating (MT) conditions; the arrow shows the TSS coordinate. The aligned reads were assembled into the transcript whose exon/intron structure is shown below.

C Zenbu (Severin *et al*, 2014) visualization of the *hs-Charme* TSS peak from FANTOM5 (Phase 1 and 2) CAGE datasets across 1,829 samples; the arrow shows the TSS coordinate which corresponds to the one identified by RNA-seq as reported in (B). The genomic coordinates are reported below.

D The diagram shows the *hs-Charme* expression in muscle vs. non-muscle tissues. The entire dataset was exported from FANTOM5 (Phase 1 and 2) CAGE datasets across 1,829 samples (see also Table EV3, sheet 1).

E Table represents the values obtained by analyzing the local sequence alignment between the human and murine *Charme* transcripts. Data were produced by using the implementation of the Smith–Waterman algorithm available at http://www.ebi.ac.uk/Tools/psa/emboss_water/. See also Table EV4.



	ex1ex2ex3	intron1	intron2
Length	608	15256	139
Sequence identity	253/608 (41.6%)	6821/15267 (44,7%)	75/139 (54,0%)
Gaps	300/608 (49,3%)	6745/15267 (44,2%)	38/139 (27,3%)
Score	562.5	13649.5	343,5

Figure EV4.

Figure EV5. MyoD is an upstream transcriptional regulator of both murine and human *Charme* transcripts.

A *hs-Charme* and MyoD expression during a time-course of *in vitro* differentiation of WT (left panel) and DMD (right panel) myoblast into myotubes. The expression values were exported from FANTOM5 (Phase 1 and 2) CAGE datasets and are reported in Table EV3 (sheet 2). Data represent the relative logarithmic expression (rle) of the tag-per-million (TPM) values of the TSS.

B UCSC visualization of MyoD ChIP-seq binding peaks in myoblast (MB) and myotubes (MT) from murine C_2C_{12} cells.

C UCSC visualization of MyoD ChIP-seq binding peaks in human myotubes (MT). The identity of each track is indicated along the left side, while the genomic coordinates are indicated below. E-boxes are highlighted in yellow.

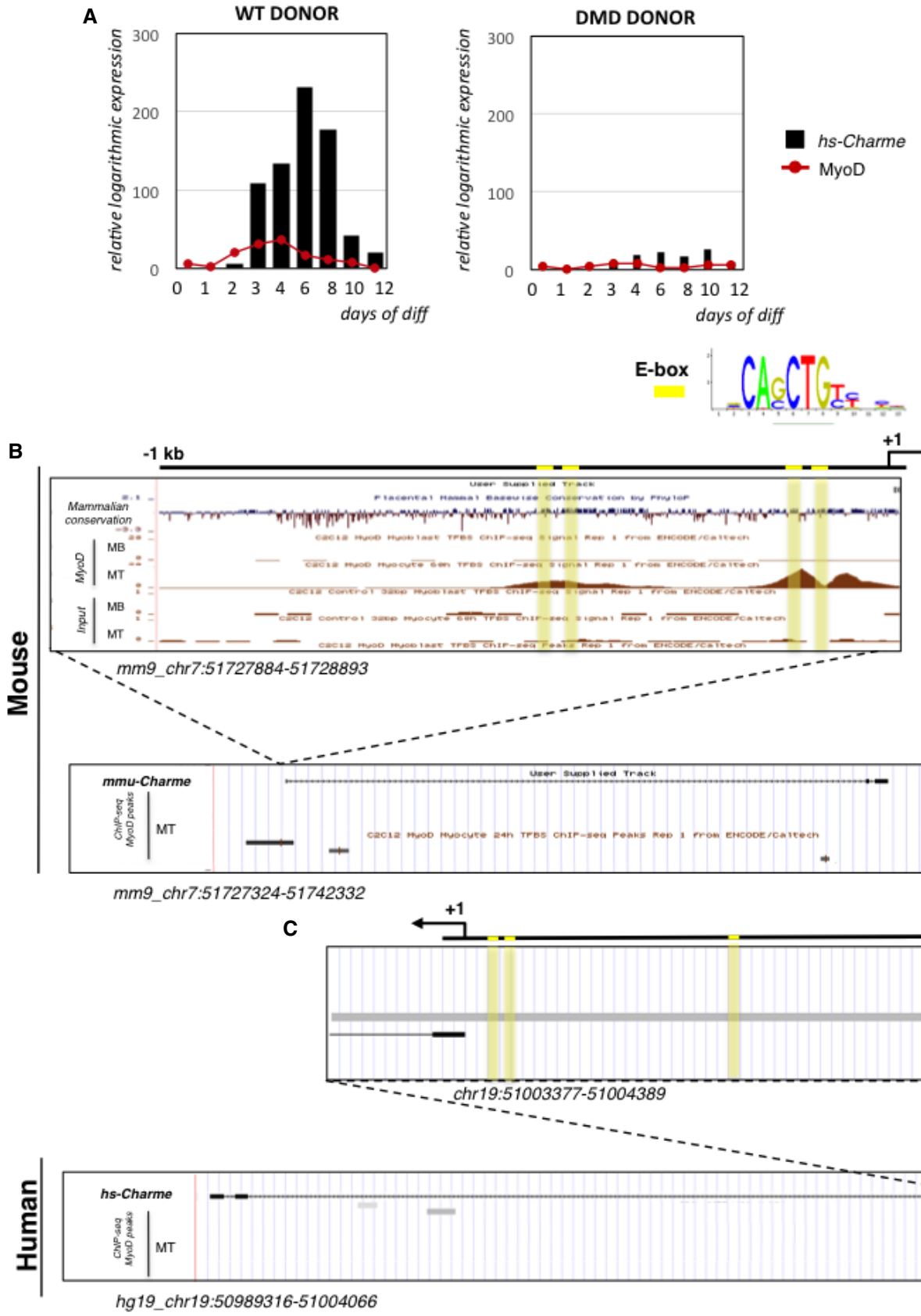


Figure EV5.

