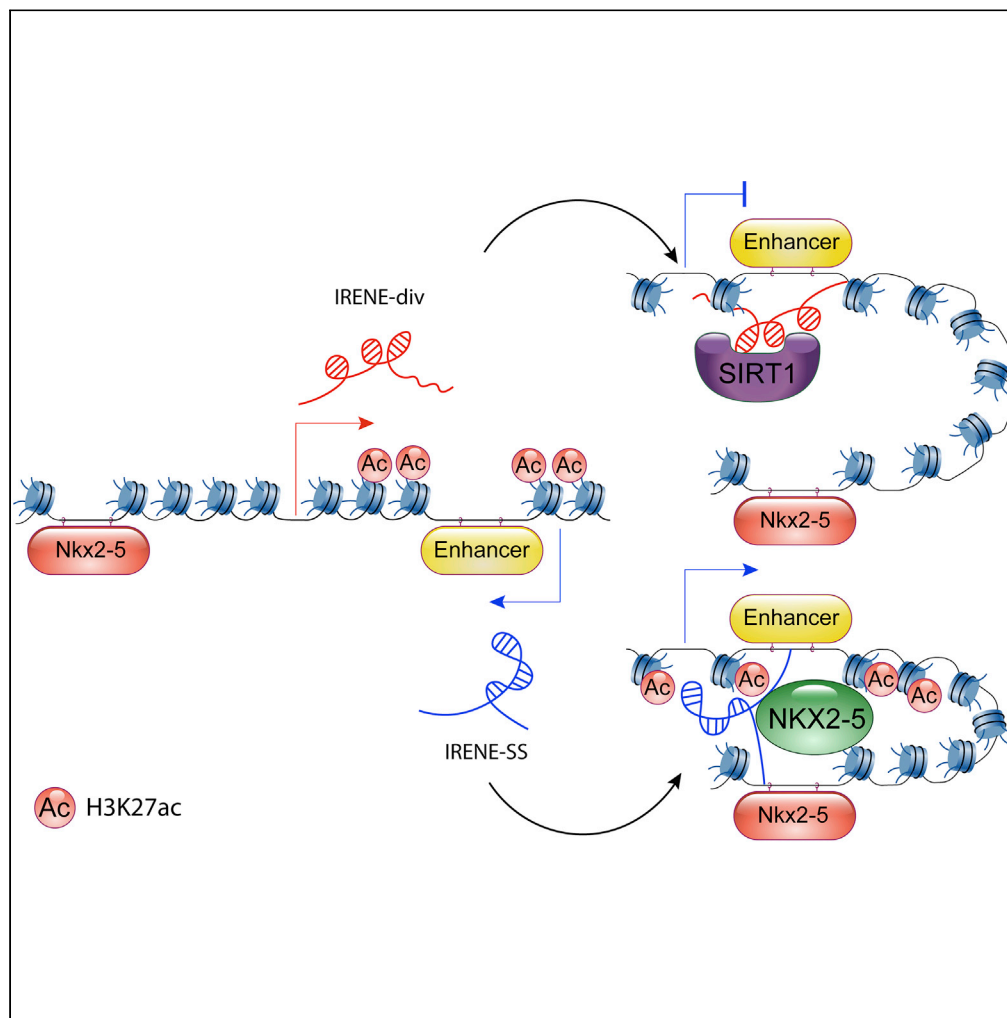## Article

Divergent Transcription of the *Nkx2-5* Locus Generates Two Enhancer RNAs with Opposing Functions

Irene Salamon,
Simone Serio,
Simona Bianco, ...,
Mario Nicodemi,
Roberto Papait,
Gianluigi
Condorelli

roberto.papait@uninsubria.it
(R.P.)
gianluigi.condorelli@hunimed.
eu (G.C.)

HIGHLIGHTS

Two eRNAs (*IRENE-SS*, *IRENE-div*) with opposing functions are found upstream of *Nkx2-5*

IRENE-SS works as a classical eRNA, acting as a transcriptional activator

IRENE-div acts unconventionally, functioning as a transcriptional repressor

*IRENE*s epigenetically control enhancer status and, subsequently, locus architecture

Article

Divergent Transcription of the *Nkx2-5* Locus Generates Two Enhancer RNAs with Opposing Functions

Irene Salamon,[1] Simone Serio,[1,2] Simona Bianco,[3] Christina Pagiatakis,[1] Silvia Crasto,[1,5] Andrea M. Chiariello,[3] Mattia Conte,[3] Paola Cattaneo,[1,5] Luca Fiorillo,[3] Arianna Felicetta,[1,2] Elisa di Pasquale,[1,5] Paolo Kunderfranco,[1] Mario Nicodemi,[3,4] Roberto Papait,[1,6,*] and Gianluigi Condorelli[1,2,5,7,*]

SUMMARY

Enhancer RNAs (eRNAs) are a subset of long noncoding RNA generated from genomic enhancers: they are thought to act as potent promoters of the expression of nearby genes through interaction with the transcriptional and epigenomic machineries. In the present work, we describe two eRNAs transcribed from the enhancer of *Nkx2-5*—a gene specifying a master cardiomyogenic lineage transcription factor (TF)—which we call Intergenic Regulatory Element Nkx2-5 Enhancers (*IRENEs*). The *IRENEs* are encoded, respectively, on the same strand (SS) and in the divergent direction (div) respect to the nearby gene. Of note, these two eRNAs have opposing roles in the regulation of *Nkx2-5*: *IRENE-SS* acts as a canonical promoter of transcription, whereas *IRENE-div* represses the activity of the enhancer through recruitment of the histone deacetylase sirtuin 1. Thus, we have identified an autoregulatory loop controlling expression of the master cardiac TF NKX2-5, in which one eRNA represses transcription.

INTRODUCTION

Cardiomyocyte (CM) homeostasis is critical for correct cardiac function: indeed, its impairment causes heart failure (Ross, 1983). Transcription factors (TFs) have long been known to orchestrate and control the pattern of gene expression that underpins homeostasis in the CM (Saadane et al., 1999; Luna-Zurita et al., 2016). More recently, long noncoding RNAs (lncRNAs), a class of transcripts >200 nucleotides in length but with a coding potential of <100 amino acids, have been described playing important roles in defining the transcriptional programs of heart development and pathophysiology (Papait et al., 2013b). For example, the lncRNAs *Fendrr* (FOXF1 adjacent non-coding developmental regulatory RNA) (Grote et al., 2013) and *Bvht* (braveheart) (Klattenhoff et al., 2013) control cardiac lineage commitment and embryonic heart development; the myosin heavy-chain-associated RNA transcript *Mhrt* (myheart) is also required for correct CM functioning, with its downregulation involved in pressure overload-induced cardiac remodeling (Han et al., 2014); *Chaer* (cardiac-hypertrophy-associated epigenetic regulator) promotes the transcription of hypertrophic genes blocking polycomb repressor complex 2 (PRC2) (Wang et al., 2016); and *MIAT* (myocardial infarction associated transcript) acts as a pro-fibrotic lncRNA in post-infarct myocardium (Qu et al., 2017).

Among the different types of lncRNAs controlling gene expression are enhancer RNAs (eRNAs), transcripts derived from enhancer regions of the genome in a manner dependent on the activation of those regulatory elements (Li et al., 2016). Enhancers generating abundant eRNAs, a characteristic of those comprised in super-enhancers—regions densely occupied by TFs and other master regulators—promote an expression of nearby genes higher than those that do not produce them (Kim et al., 2018). The control of gene activation by eRNAs is achieved via different mechanisms, including maintenance of chromatin accessibility for RNA polymerase II (Mousavi et al., 2013), and interaction with the co-activator complex Mediator (Lai et al., 2013) or with the cohesion complex (Hsieh et al., 2014) for enhancer-promoter loop maintenance. Despite the importance of eRNAs, their role in the heart remains poorly explored.

Here, we describe two eRNAs transcribed from the enhancer of NK2 homeobox 5 (*Nkx2-5*)—which encodes a master TF for cardiac development and pathophysiology (Pashmforoush et al., 2004; Saadane et al., 1999;

Bruneau et al., 2000; Delaughter et al., 2016)—with antagonistic regulatory effects on transcription of the target gene. These two eRNAs, which we have called Intergenic Regulatory Element *Nkx2-5* Enhancers (*IRENE*s), are encoded from the same strand (SS) and from the divergent direction (div). Of note, *IRENE-SS* acts as a canonical promoter of *Nkx2-5* transcription, whereas *IRENE-div* represses the activity of the enhancer by recruiting the histone deacetylase (HDAC) sirtuin 1 (SIRT1). Experiments on human induced pluripotent stem cells (hiPSCs) revealed that the two human homologs are regulated during cardiac differentiation *in vitro*. Our findings demonstrate the existence of an epigenetic mechanism involving an autoregulatory loop in which the transcription of a TF is controlled by the balance of eRNAs stemming from its own gene's enhancer.

RESULTS

The *Nkx2-5* Enhancer Encodes Two eRNAs

To study the role of eRNAs in the heart, we first identified in mouse cardiomyocytes (CMs) any potential eRNAs conserved in humans. To that end, we analyzed RNA sequencing (RNA-seq) datasets generated from CMs isolated from adult mouse left ventricle (Greco et al., 2016; Rosa-Garrido et al., 2017) (Figure 1A). We found 112,286 lncRNAs (FPKM >0.1) in mouse, 4,127 of which were expressed also in myocardial tissue from nine healthy human donors (Liu et al., 2019) and in hiPSC-derived CMs at day 15 of differentiation. Selecting only those lncRNAs located in intergenic regions, we found 2,318 associated with a nearby protein-coding gene, 132 of which were enriched in heart. After cross-checking these with three datasets of cardiac enhancers (Wamstad et al., 2012; Papait et al., 2013a; Dickel et al., 2016), 46 were deemed putative eRNAs (Figure S1A). This set of putative eRNAs was associated with 26 protein-coding genes related to muscle tissue morphogenesis, cardiac cell development, and apoptosis, as revealed by gene ontology (GO) enrichment analysis (Figure S1B).

We focused subsequent investigation on two putative eRNAs (*IRENE-SS* and *IRENE-div*) encoded, respectively, on the positive and the negative DNA strand of the enhancer involved in regulating expression of *Nkx2-5* (Lien et al., 1999) (Figure 1B). The eRNAs were validated by northern blotting and through the mapping of their transcription start sites (TSSs) using 5′ rapid amplification of cDNA ends (5′ RACE) analysis (Figure S2A).

To characterize *IRENE-SS* and *IRENE-div*, we first assessed their expression with respect to that of *Nkx2-5* in 13 different adult mouse tissues and four sorted cardiac cell populations (CMs, fibroblasts, endothelial cells, and immune cells). Quantitative PCR (qPCR) revealed that both had CM-specific expression (Figures S2B and S2C). Absolute quantification at three stages of heart maturation (E14.5, neonatal, and adult) revealed that they were more highly expressed at the neonatal stage, but at differing ratios with respect to *Nkx2-5* (Figure S3A).

The function of a lncRNA is tightly related to its intracellular localization (Chen, 2016; Fazal et al., 2019). Thus, we assessed the subcellular distribution of the eRNAs through cellular fractionation (Figure S3B) and RNA-FISH (Figure S3B): 60% of *IRENE-SS* was located in the nucleus, 45% of which bound to chromatin, whereas the localization of *IRENE-div* was dependent upon the stage of the transcript's biogenesis, with the mature form located prevalently in the cytoplasm and the immature form having a subcellular localization similar to that of *IRENE-SS*. These results were suggestive of potential involvement of both eRNAs in chromatin organization.

Evolutionary selection is an important challenge in the study of lncRNAs because some diverge rapidly (Amaral et al., 2008), in some cases with purifying selection at the sequence level and in other cases with selection only for transcription (Chen et al., 2016). Therefore, we evaluated human orthologs with the Transcript-Transcript Identity (TTI) score (Chen et al., 2016) for each transcript in the syntenic genomic locus: we found the TTI scores on TSSs to be 45% and 64% for *IRENE-SS* and *IRENE-div* (Figure S3C). Moreover, by validating the expression of the eRNAs and that of the target gene in hiPSCs at different days of CM differentiation, we found that the expression of both transcripts closely followed that of *Nkx2-5* (Figure S3D). Thus, the *Nkx2-5* enhancer transcribes two putative cardiac-specific eRNAs whose expression follows that of the nearby gene in mouse and humans.

IRENE-SS Activates and *IRENE-div* Represses Transcription of *Nkx2-5*

Since eRNAs regulate the expression of neighboring protein-coding genes (Kim et al., 2018), we investigated the role of the *IRENE* transcripts in regulating the expression of *Nkx2-5*. To that end, primary neonatal mouse CMs were transfected with gapmeRs synthesized to specifically silence either one of the two eRNAs and then *Nkx2-5* expression was analyzed 24 and 48 h after transfection. qPCR revealed that silencing *IRENE-SS* reduced *Nkx2-5* at the mRNA (−34%) and protein (−49%) levels, whereas silencing
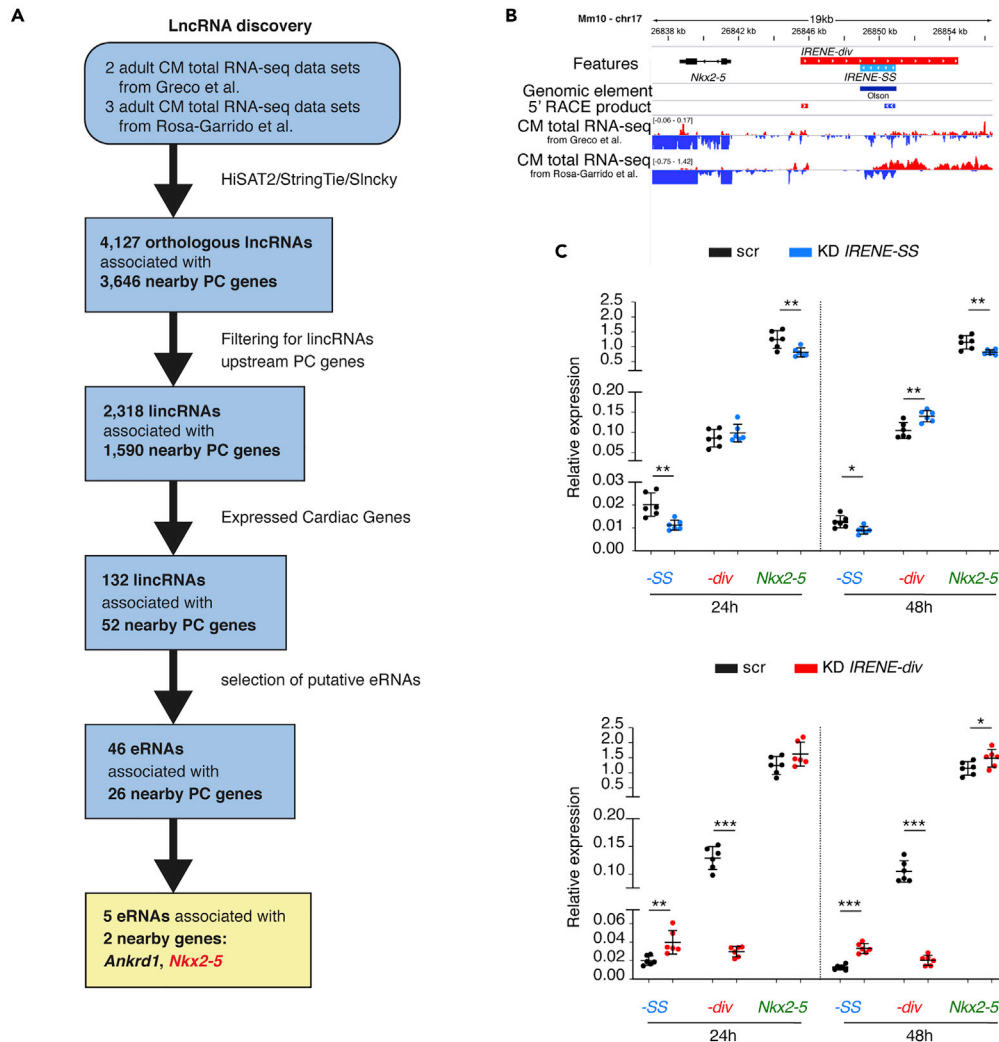
Figure 1. Identification of *IRENE-SS* and *IRENE-div* and Their Regulation of *Nkx2-5*

(A) Bioinformatics workflow for the discovery of conserved cardiac-enriched lncRNAs in adult mouse cardiomyocytes. PC, protein coding.

(B) Schematic representation of *IRENE-div* (red) and *IRENE-SS* (blue) with genomic enhancer elements, transcription start sites (TSSs) reported as a 5' RACE product, and stranded cardiomyocyte total RNA-seq from Greco et al. (2016) and Rosa-Garrido et al. (2017).

(C) Relative expression of *IRENE-SS*, *IRENE-div*, and *Nkx2-5* after gapmeR silencing specific for either one of the two eRNAs compared with a negative control (scramble, *scr*) in primary cultures of neonatal cardiomyocytes at two different time points (24 and 48 h). In blue, knockdown (KD) of *IRENE-SS*; in red, KD of *IRENE-div*. Silencing of *IRENE-SS* decreased *Nkx2-5* mRNA; in contrast, silencing of *IRENE-div* increased *Nkx2-5* mRNA.

Data are represented as mean ± SD. Unpaired t test was used, *, p value <0.05; **, p value <0.01; ***, p value <0.001. See also Figures S1–S5.

IRENE-div caused an increase of *Nkx2-5* at both levels (mRNA, +31%; protein, +84%) (Figures 1C and S1C). Thus, the two eRNAs had opposite effects on *Nkx2-5* transcription: *IRENE-SS* functioned as a transcriptional activator, whereas *IRENE-div* acted as a transcriptional repressor.

To verify whether the biological effects of the eRNAs were mainly due to their ability to regulate *Nkx2-5* expression, we analyzed the impact of their silencing on global gene expression, determining whether the changes were related to modifications in the activity of TFs. To that end, we first compared the expression profile of neonatal CMs transfected with the gapmeRs versus that of CMs transfected with a scrambled control. Silencing of *IRENE-div* and *IRENE-SS* caused significant modulation of 2,754 and 2,542 genes,

respectively (Figure S4A). Then, we determined the enrichment of putative NKX2-5 targets versus targets of two other cardiac TFs: TBX5 and GATA4. Since these TFs regulate transcription through binding to enhancers as well as promoters (Akerberg et al., 2019), we considered as putative targets all genes mapping to chromatin immunoprecipitation (ChIP) peaks on both those regulatory elements, using a published ChIP-seq dataset (Luna-Zurita et al., 2016). The overlap between genes modulated by *IRENE-div* or *IRENE-SS* silencing and the datasets of putative targets of the three TFs analyzed revealed that the most enriched modulated genes were indeed NKX2-5 targets (Figure S4B): specifically, 43.6% (N = 1,108) and 43.8% (N = 1,205) of genes were modulated by depletion of *IRENE-SS* and *IRENE-div*, respectively. Moreover, GO analysis revealed that pathways related to cardiac muscle development and ion transmembrane transporter activity were specifically associated with genes repressed by the silencing of *IRENE-SS* and genes increased after the silencing of *IRENE-div* (Figure S4C), a finding coherent with our previous observations on the opposing effects of the eRNAs on *Nkx2-5*.

Then, to assess whether the effect of *IRENE-SS* on gene expression was specifically dependent upon its ability to regulate *Nkx2-5*, we tested if overexpression of NKX2-5 rescued correct gene expression in CMs with knocked-down *IRENE-SS*. To this end, neonatal CMs were transfected with a gapmeR targeting *IRENE-SS* and with a construct for NKX2-5 overexpression and then analyzed 24 h later with qRT-PCR for the expression of six NKX2-5-targeted genes identified by ChIP-seq. We found that overexpression of NKX2-5 was not sufficient to restore correct expression of the analyzed genes (Figure S5). The absence of a rescued phenotype could be explained by the fact that the eRNA acted on those genes independently of the level of NKX2-5, although we cannot exclude that the overexpression obtained was not sufficient to correct the transcriptional changes.

IRENE-SS Promotes Transcription by Recruiting NKX2-5 to Regulatory Sequences

Because *IRENE-SS* had two key characteristics of an eRNA—transcription from an enhancer and transcriptional promotion of the neighboring gene—we evaluated the impact of its silencing on the activity of the enhancer, analyzing the distribution of two histone modifications: H3K27ac, which marks active enhancers, and H3K27me3, which defines poised enhancers (Shlyueva et al., 2014). ChIP assay coupled with qPCR (ChIP-qRT-PCR) (see Figure 2A for the location of the primers used throughout this work) revealed that *IRENE-SS* silencing decreased H3K27ac on the enhancer, without interfering with H3K27me3 (Figure 2B).

Since NKX2-5 promotes its own transcription (Clark et al., 2013), and many eRNAs participate in enhancer-promoter loop formation (Wang et al., 2011; Arner et al., 2015; Sigova et al., 2015), we assessed whether *IRENE-SS* promoted the recruitment of NKX2-5 to its own promoter/enhancer regions. First, we determined if there was a physical association between *IRENE-SS* and NKX2-5. An RNA immunoprecipitation (RIP) assay on HL1 cells, an immortalized CM line, carried out with antibodies targeting NKX2-5 indicated that the TF bound *IRENE-SS* but not *IRENE-div* (Figure 2C). Moreover, *in vitro* RNA pull-down followed by immunoblotting using a biotinylated sense or antisense RNA fragment of *IRENE-SS* confirmed precipitation of NKX2-5 with the antisense RNA (Figure 2D). Then, we tested the binding of NKX2-5 protein to its own genomic regulatory regions (promoter and enhancer) in neonatal CMs after silencing of *IRENE-SS*: ChIP-qRT-PCR revealed decreased NKX2-5 binding after silencing (Figure 2E). Thus, *IRENE-SS* works as a canonical eRNA, promoting the transcription of *Nkx2-5* through the recruitment of NKX2-5 to its own promoter and enhancer.

IRENE-div Silences the *Nkx2-5* Enhancer by Recruiting SIRT1

The above results indicated that *IRENE-div* repressed *Nkx2-5* transcription. This finding was supported by the analysis of H3K27ac and H3K27me3 on the regulatory regions of the *Nkx2-5* gene after silencing of the eRNA with a gapmeR in neonatal CMs: indeed, ChIP-qRT-PCR revealed that *IRENE-div* silencing enriched H3K27ac at the enhancer of *Nkx2-5*, whereas the level of H3K27me3 did not change (Figure S6A).

Since *IRENE-div* bound chromatin and its silencing increased histone acetylation, we tested the hypothesis that the eRNA controlled the transcription of *Nkx2-5* through recruitment of HDACs. To that end, we examined in neonatal CMs and quiescent skeletal muscle satellite cells the genomic distribution of the enzymes as reported in published ChIP-seq datasets (Ai et al., 2017; Ryall et al., 2015). We found that the HDACs SIRT1 and HDAC2 were enriched at the *Nkx2-5* enhancer (Figure S6B). Binding of those HDACs to the enhancer of the TF's gene was then validated in HL1 cells with ChIP-qRT-PCR: as expected, antibodies targeting either SIRT1 or HDAC2 immunoprecipitated the *Nkx2-5* enhancer region (Figure S6C). However, RIP performed with antibodies against SIRT1 and HDAC2 showed significant enrichment of *IRENE-div* in the SIRT1 immunoprecipitate versus negative
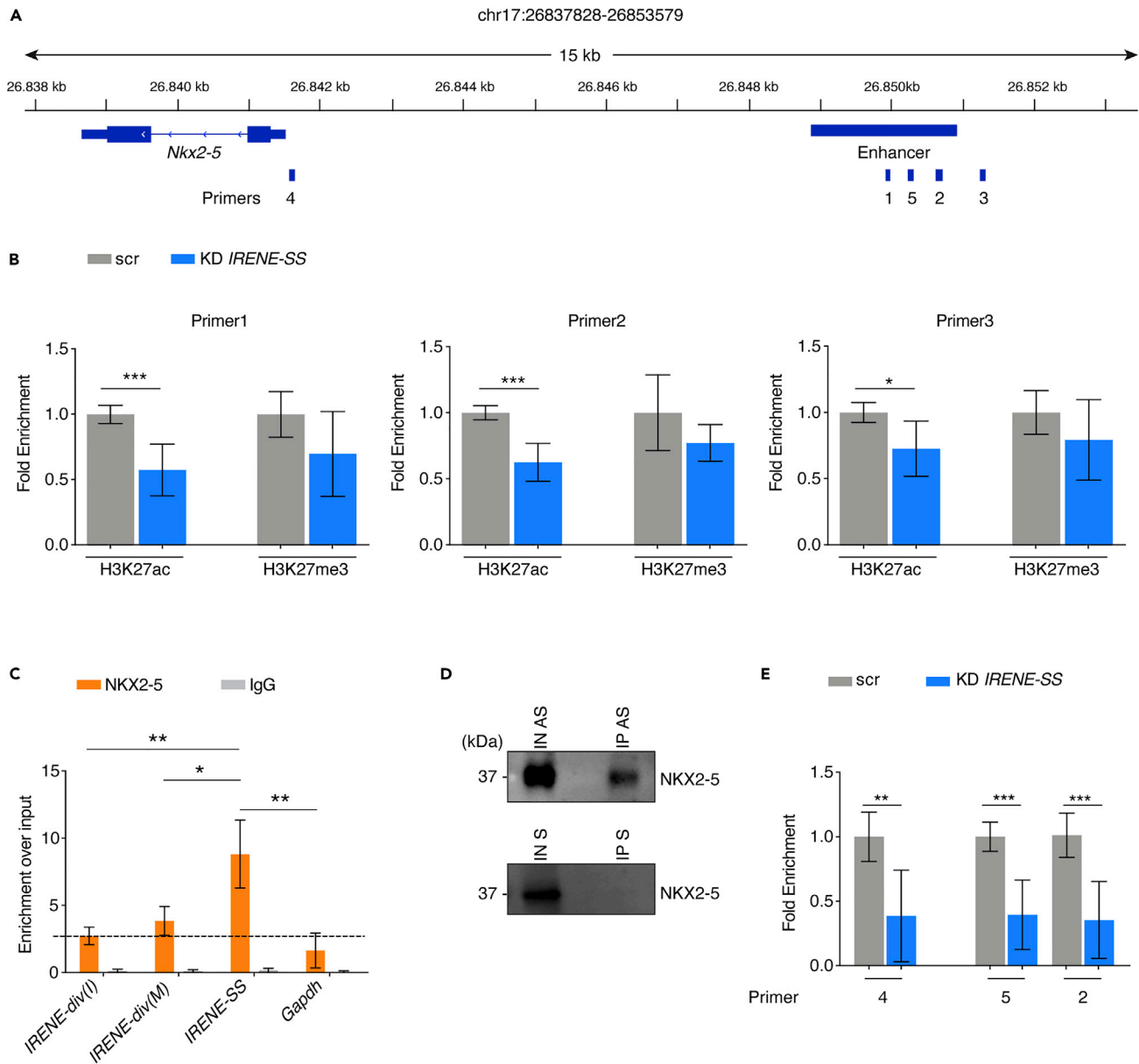
Figure 2. *IRENE-SS* Acts via the Binding and Recruitment of NKX2-5 to Its Own Regulatory Region

(A) Location of primers used in this work.

(B) ChIP-qRT-PCR was conducted on neonatal cardiomyocytes. Silencing of *IRENE-SS* was associated with reduced H3K27ac on the enhancer region (primers 1, 2, and 3) and no effect on H3K27me3, with the exception of primer 2. Values are expressed as fold enrichment over scrambled gapmeR (scr)-treated cells. N = 3 biological replicates (Student's unpaired t test).

(C) RNA immunoprecipitation (RIP) assay conducted on HL1 cells demonstrating biochemical interaction of NKX2-5 with *IRENE-SS* but not with *IRENE-div* or the control gene *Gapdh*. Interaction is expressed as enrichment over input. N = 3 different experiments (one-way ANOVA with Sidak post hoc test).

(D) *In vitro* RNA pull-down revealed an association of NKX2-5 with the RNA fragment of *IRENE-SS* after SDS-PAGE and immunoblotting. IN, input; IP, immunoprecipitate; AS, antisense; S, sense.

(E) ChIP-qRT-PCR performed on neonatal cardiomyocytes after *IRENE-SS* knockdown (KD), showing a reduction of NKX2-5 enrichment at its promoter (primer 4) and enhancer (primers 5 and 2). Values are expressed as fold enrichment over scr-treated cells of N = 3 biological replicates (Student's unpaired t test).

Data are represented as mean ± SD. *p value <0.05; **, p value <0.01; ***, p value <0.001.
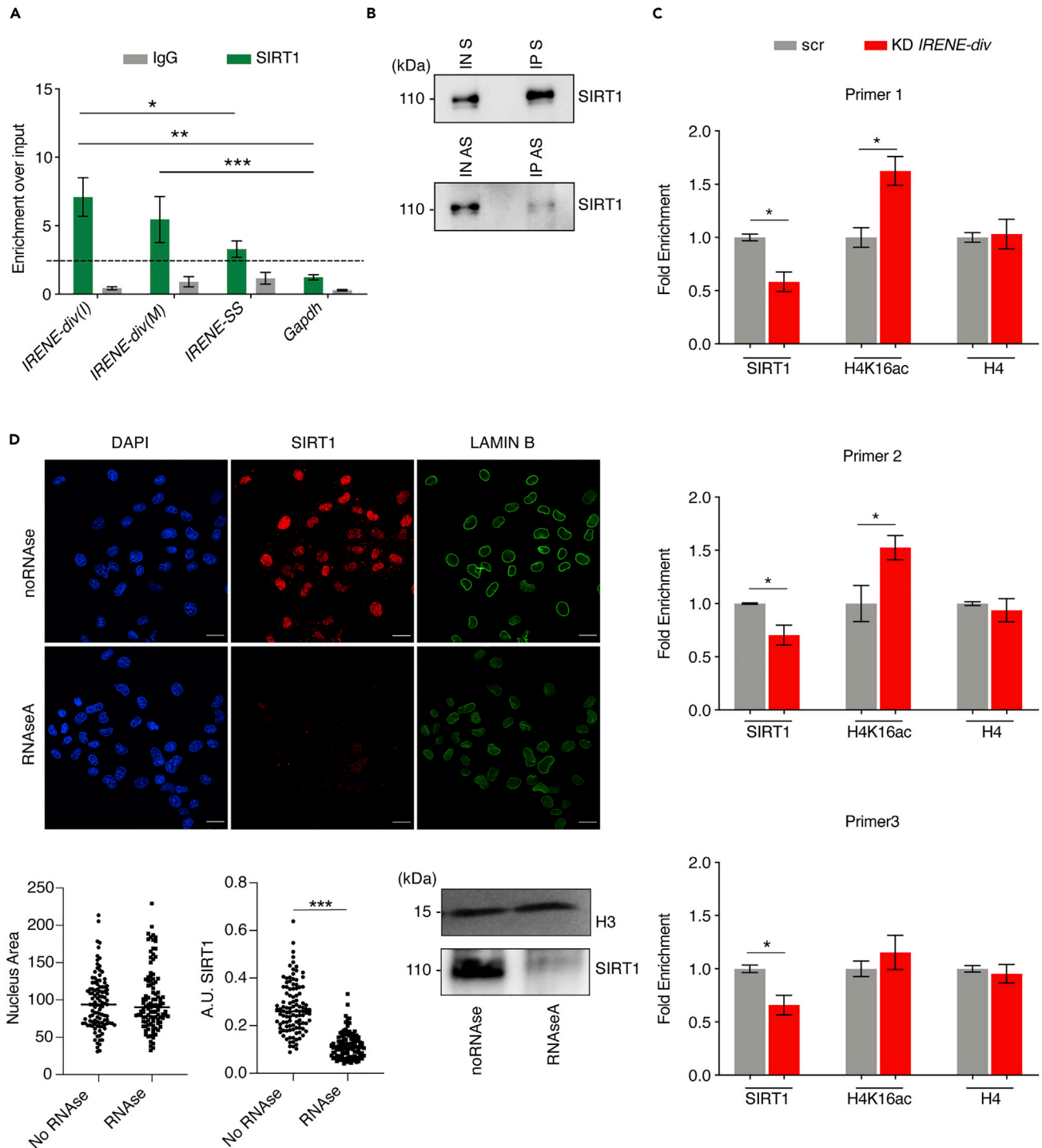
Figure 3. *IRENE-div* Acts as a Transcriptional Repressor through Recruitment of SIRT1

(A) RNA immunoprecipitation performed on HL1 cells, using antibodies against SIRT1, showing that immature (I) and mature (M) forms of *IRENE-div*, but not *IRENE-SS* or *Gapdh* mRNA, bind SIRT1. Interaction is expressed as enrichment over input in N = 3 different experiments (one-way ANOVA with Sidak post hoc test).

(B) *In vitro* RNA pull-down revealed an association of SIRT1 with the RNA fragment of *IRENE-div* after SDS-PAGE and immunoblotting. IN, input; IP, immunoprecipitate; AS, antisense; S, sense.

(C) ChIP-qRT-PCR conducted on HL1 cells after silencing of *IRENE-div* (red bars) compared with scrambled gapmeR-treated cells (gray bars). Antibodies against SIRT1, H4K16ac, and H4, and three primers (1, 2, and 3) designed on the genomic enhancer element were used. Knockdown (KD) of *IRENE-div*

Figure 3. *Continued*

decreased SIRT1, with a consequential increase in H4K16ac, the histone modification produced by that enzyme. H4 immunoprecipitation was used as control. Values are fold enrichment over scr-treated cells. N = 2 biological replicates (Student's unpaired t test).

(D) Immunofluorescence in HL1 cells treated with RNAse A. Untreated cells were used as control. Nuclei stained with DAPI (blue); nuclear periphery stained with lamin B (green); SIRT1 visualized in red (scale bars, 10 μm). Dot plots of the quantification of the immunofluorescence signal normalized for nuclear area revealed loss of the SIRT1 signal after RNAse A treatment, a finding confirmed by western blotting.

Data are represented as mean ± SD. *, p value <0.05; **, p value <0.01; ***, p value <0.001. See also Figure S6.

control (IgG immunoprecipitate), indicating that *IRENE-div* bound SIRT1 but not HDAC2 (Figures 3A and S6D). This interaction was not observed for *IRENE-SS*. RNA pull-down confirmed that SIRT1 precipitated with the sense RNA but not with an antisense one (Figure 3B).

Next, to verify whether the above interaction was necessary for recruitment of SIRT1 to the enhancer, we assessed the amount of the HDAC and one of its substrates, histone H4 acetylated at lysine 16 (H4K16ac), at this genomic region after silencing *IRENE-div* in HL1 cells. ChIP-qRT-PCR revealed that silencing decreased the level of the HDAC and increased H4K16ac deposition at the enhancer region (Figure 3C), suggesting that *IRENE-div* promoted the transcriptional repression of *Nkx2-5* through the recruitment of SIRT1. This finding was confirmed in primary cultures of neonatal mouse CMs (Figure S6E). Furthermore, when HL1 cells were treated with RNase A, which digests single-stranded RNA, SIRT1 became unbound from chromatin (Figure 3D), suggesting that RNA was needed to maintain this HDAC in the nucleus and that *IRENE-div* was at least partially responsible for local binding at the enhancer region under study.

Polymer Model Simulation Reveals Alteration in 3D Chromatin Conformation

The above findings indicated that *IRENE-SS* and *IRENE-div* evoked antagonistic effects on the activity of the *Nkx2-5* enhancer through modulation of the epigenetic signature at this regulatory region. However, chromatin architecture represents a higher order of organization with respect to the epigenetic landscape. Therefore, we tested the hypothesis that the eRNAs and the histone marks modulated by their silencing also modified the chromatin architecture at the enhancer. To that end, we modeled the 3D structure of the *Nkx2-5* locus, using a computational inference approach based on the Strings & Binders (SBS) polymer model of chromatin (Bianco et al., 2018). In an SBS model, a filament of chromatin is represented as a coarse-grained chain of beads with binding sites for diffusing molecular binders (Figure 4A). A detailed description of the theoretical model is reported in the Methods section.

Specifically, we focused on a 200-kb locus around *Nkx2-5* (chr17:26740000–2694000, mm10) and used published Hi-C data on mouse CMs at 5-kb resolution (Rosa-Garrido et al., 2017) to derive a 3D model of the locus. The Hi-C experimental contact frequency matrix was well reproduced by the model contact matrix, as indicated by their high Pearson's (r) and distance-corrected (r') correlations (Figure 4B). 3D reconstruction of the genomic region revealed, in general, a spatial proximity of the enhancer to the promoter through a loop formation (Figure 4C).

The derived polymer model for the *Nkx2-5* locus included 15 different types of binding sites (genomic locations given in Figure 4D). To get insights into their molecular nature, we evaluated their correlation with available ChIP-seq peaks of histone marks (H3K4me3, H3K9ac, H3K27ac, H3K79me2, and H3K27me3 in Papait et al., 2013a) (Papait et al., 2013a) and the CCCTC-binding factor (CTCF) (Rosa-Garrido et al., 2017). We found that each binding site type had a specific, combinatorial pattern of positive and negative correlations (Figure 4D). Since silencing of *IRENE* altered acetylation at the enhancer, we tested *in silico* the impact of the eRNAs on locus folding by rendering inert the binding sites associated with H3K27ac in the two 5-kb bins containing the ncRNAs/enhancer region. Then, we re-computed the model contact matrix and compared it with the experimental control Hi-C. Notably, the correlation decreased (Δr' = r'(wt)-r'(silenced)) in the silenced model versus wild-type in the *Nkx2-5* topologically associating domain (green region), with Δr' significantly different from that in a control model in which we silenced all other possible bin pairs within the region (Figure 4E), excluding bins containing sites for CTCF, a protein known to affect chromatin architecture (Fudenberg et al., 2016). These findings indicated there was a rearrangement of locus architecture after silencing of the eRNAs. In particular, the model predicted a loss of contact of ~30% between the *Nkx2-5* gene and its enhancer compared with the wild-type (Figure 4F).

DISCUSSION

eRNA was identified as a class of lncRNA able to promote transcription (Kim et al., 2010). Here, we show that two eRNAs—namely, *IRENE-SS* and *IRENE-div*, both encoded from the same enhancer region
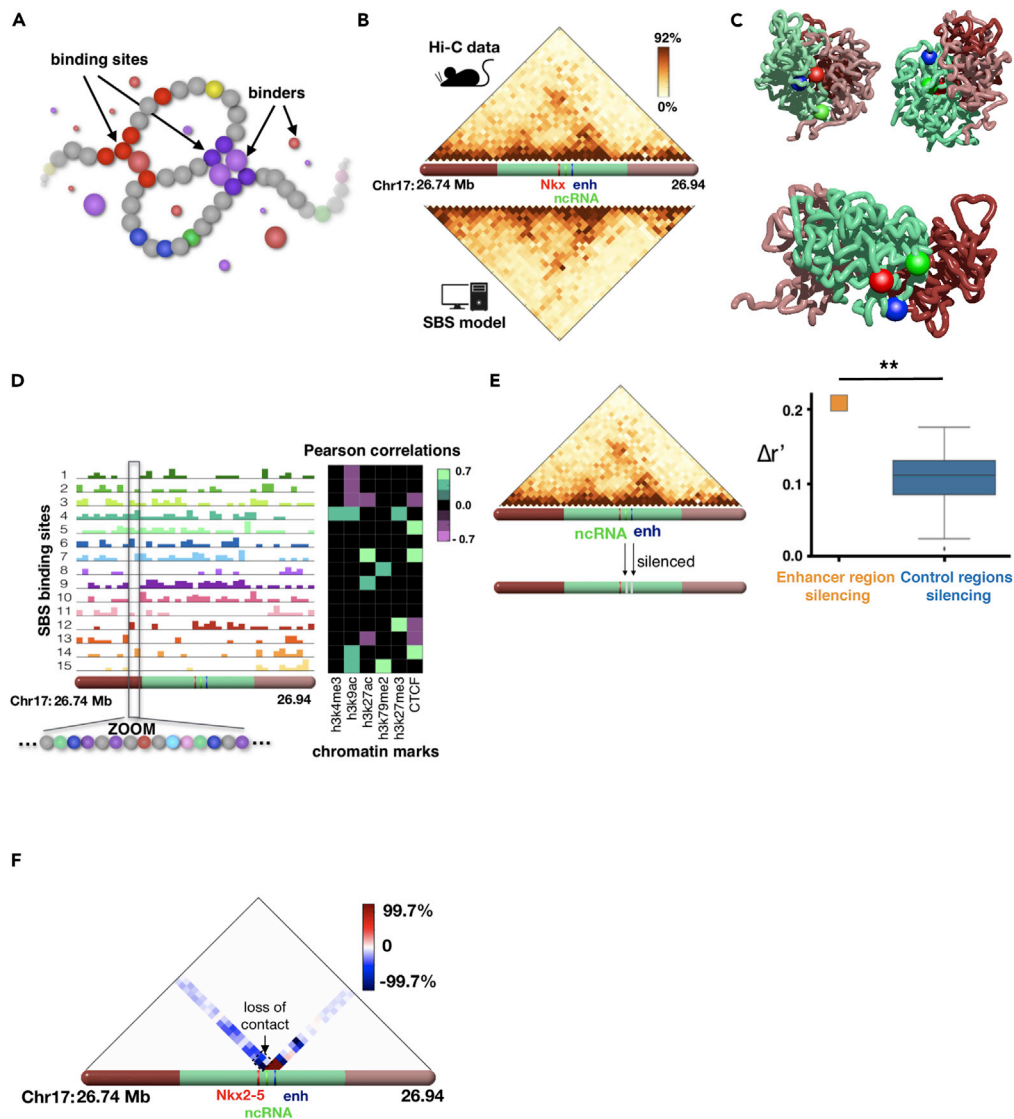
Figure 4. Polymer Modeling of the *Nkx2-5* Locus Predicts that eRNA Silencing Impacts 3D Chromatin Conformation

(A) Representation of the Strings & Binders polymer model, with the chromatin filament represented as a string of beads, some of which are binding sites for diffusing molecular binders. Different types of binding sites interacting with specific cognate binders are represented with different colors; non-interacting, inert beads are represented in gray.

(B) Hi-C contact matrix (top triangle) of the *Nkx2-5* locus is well reproduced by the SBS-derived contact matrix (bottom triangle) (Pearson correlation coefficient, r = 0.97; distance-corrected correlation, r' = 0.75). In the bar in the middle, the color scheme used is that of the 3D polymer representation given in (C).

(C) Examples of SBS-derived 3D structures of the locus, with relevant elements indicated by colored beads: *Nkx2-5* gene (red), enhancer (blue), and eRNA (green). The 3D structures highlight that promoter and enhancer tend to colocalize, albeit with an intrinsic structural variability in the general architecture.

(D) Genomic location and abundance of the different types of binding sites inferred for the *Nkx2-5* locus. The different types are represented in different colors (left) and are found to correlate with specific combinations of chromatin marks (right).

(E) Impact of eRNA silencing on 3D architecture at the locus. The binding sites associated with acetylation were made inert (gray) at the eRNA and enhancer regions (left diagram). The distance-corrected correlation with Hi-C of the silenced model contact matrix (orange box) is significantly lower than the wild-type (blue box). **, p value <0.01.

(F) Subtraction map between the silenced and wild-type model contact matrices. Red and blue indicate gain and loss of contacts upon *in silico* silencing of acetylated binding sites at the enhancer/ncRNA region. Note that most of the contact changes are located in the *Nkx2-5* topologically associating domain (TAD; green). The dashed box highlights the predicted loss of contacts between *Nkx2-5* and the enhancer.

upstream from the TSS of the *Nkx2-5* gene—have opposing effects on the regulation of the nearby gene: *IRENE-SS* acts as a classical eRNA promoting the transcription of its target gene, whereas *IRENE-div* is an unusual eRNA that functions as a transcriptional repressor. Here, we describe the differential molecular effects generated by two eRNAs transcribed from opposite strands of a genomic enhancer, and of how the activity of a gene's enhancer, namely, that of *Nkx2-5*, depends on the balance of the two eRNAs; indeed, specific disruption of each one of the eRNAs produced contrasting effects on the transcription of *Nkx2-5*.

The locus of the enhancer under study was originally described by E.N. Olson and colleagues: they demonstrated the overlapping of its activity with *Nkx2-5* expression during cardiogenesis and found that it is composed of three parts having different roles in the regulation of *Nkx2-5* gene expression (Lien et al., 1999). More in detail, a series of genomic deletions were used to define two activating regions separated by a negative regulatory segment driving the transcription of *Nkx2-5* during cardiogenic specification. Our findings elucidate the mechanism of action of this enhancer. The upstream portion of the tripartite enhancer encodes *IRENE-SS*. This lncRNA positively regulates *Nkx2-5* expression by promoting, on the one hand, the acetylation of H3K27 needed for chromatin looping, permitting the interaction between the enhancer and the promoter, and on the other the recruitment of NKX2-5 to both the enhancer and the promoter. Our findings also support the conclusion that the auto-regulatory loop of *Nkx2-5* is mediated by eRNAs generated in the proximity of the gene's locus and that the binding of NKX2-5 protein to chromatin depends not only on the presence of a specific binding motif but also on its direct interaction with the eRNAs. Although it has been already reported that RNAs stabilize TFs on regulatory elements of a gene (Sigova et al., 2015), our findings describe how this stabilization takes place at the *Nkx2-5* locus.

RNA FISH experiments identified several nuclear sites where these lncRNAs were present, leading us to not exclude the possibility that *IRENE-SS* and *IRENE-div* acted also in *trans*. For *IRENE-SS*, we can speculate that, aside from recruiting NKX2-5 to its own gene's enhancer, it is needed also for proper localization of the TF at other regulatory regions it targets. However, overexpression of NKX2-5 in neonatal CMs knocked-down for *IRENE-SS* did not restore the correct expression of selected analyzed target genes. The absence of direct phenotypic rescue suggests that *IRENE-SS* could regulate gene expression through mechanisms other than the regulation of *Nkx2-5* expression. Indeed, *IRENE-SS* could recruit NKX2-5 to the *Nkx2-5* enhancer as well as to other regulatory regions targeted by the TF. However, we cannot exclude that the level and timing of NKX2-5 overexpression we used were not ideal to obtain a rescued phenotype. Moreover, we were not able to determine how the direct effect on chromatin architecture induced by altered enhancer acetylation status impacted additional, indirect transcriptional changes. Further investigations are needed to resolve these aspects.

In contrast, *IRENE-div* acts as a transcriptional repressor, recruiting the HDAC SIRT1. Indeed, the silencing of *IRENE-div* was associated with increased acetylation of the enhancer and decreased binding of SIRT1 to this regulatory region. SIRT1 has been shown to regulate NKX2-5 at the protein level (Tang et al., 2016). Our findings support also a role for this HDAC as a silencer of *Nkx2-5*'s enhancer. This, together with the nuclear solubilization of SIRT1 after digestion with RNase A, suggests a role for RNA in stabilizing the binding of HDACs to chromatin. However, the mechanism by which *IRENE-div* anchors SIRT1 to chromatin remains to be defined.

The importance of TF balance and cooperation for correct heart development and CM homeostasis is already known (Luna-Zurita et al., 2016; Akerberg et al., 2019). Our work strongly supports the notion that eRNAs interacting with critical TFs also contribute to the complex network of TF activity necessary for the CM specification program. Of note, we have demonstrated not only the existence of the two eRNAs in human cells but also that they are expressed in conjunction with that of the nearby *Nkx2-5* gene during hiPSC differentiation toward the CM lineage *in vitro*. Although further investigation is needed to establish the molecular role in human CMs, the relatively high sequence identity and conservation of the genomic locus between human and mouse suggests the existence of similar functional orthologs in humans.

Limitations of the Study

The absence of an animal model is a limitation of our study. Indeed, we focused our efforts to understanding the molecular mechanism of the two transcripts at the cellular level. An *in vivo* study could generate other important information on the physiological and pathophysiological importance of these two eRNAs in the heart. Further study is also needed to document the function of the human orthologs in order to verify and confirm the relevance of the transcripts in human cardiomyocyte homeostasis.

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gianluigi Condorelli (gianluigi.condorelli@hunimed.eu).

Materials Availability

All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and Code Availability

The published article includes all datasets generated or analyzed during this study. Original data have been deposited to Mendeley Data: https://doi.org/10.17632/25cnynmzx6.1. The accession numbers for the RNA-seq data reported in this paper are NCBI Gene Expression Omnibus GSE143930 and GSE143929.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101539.

AUTHOR CONTRIBUTIONS

Conceptualization: I.S., R.P., and G.C.; Methodology: I.S., S.B., C.P., E.d.P., S.C., A.F., and P.C.; Software: S.S., S.B., A.M.C., L.F., and M.C.; Formal Analyses: I.S. and S.S; Investigation: I.S. and C.P.; Writing – Original Draft: I.S. and R.P; Writing – Review & Editing: I.S., R.P., and G.C; Supervision: R.P. and G.C.; Funding Acquisition: G.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Ai, S., Peng, Y., Li, C., Gu, F., Yu, X., Yue, Y., Ma, Q., Chen, J., Lin, Z., Zhou, P., et al. (2017). EED orchestration of heart maturation through interaction with HDACs is H3K27me3-independent. Elife 6, e24570.

Akerberg, B.N., Gu, F., Vandusen, N.J., Zhang, X., Dong, R., Li, K., Zhang, B., Zhou, B., Sethi, I., Ma, Q., et al. (2019). A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. Nat. Commun. 10, 4907.

Amaral, P.P., Dinger, M.E., Mercer, T.R., and Mattick, J.S. (2008). The eukaryotic genome as an RNA machine. Science 319, 1787–1789.

Arner, E., Daub, C.O., Vitting-Seerup, K., Andersson, R., Lilje, B., Drablos, F., Lennartsson, A., Ronnerblad, M., Hrydziuszko, O., Vitezic, M., et al. (2015). Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. Science 347, 1010–1014.

Bianco, S., Lupianez, D.G., Chiariello, A.M., Annunziatella, C., Kraft, K., Schopflin, R., Wittler, L., Andrey, G., Vingron, M., Pombo, A., et al. (2018). Polymer physics predicts the effects of structural variants on chromatin architecture. Nat. Genet. 50, 662–667.

Bruneau, B.G., Bao, Z.Z., Tanaka, M., Schott, J.J., Izumo, S., Cepko, C.L., Seidman, J.G., and Seidman, C.E. (2000). Cardiac expression of the ventricle-specific homeobox gene Irx4 is modulated by Nkx2-5 and dHand. Dev. Biol. 217, 266–277.

Chen, J., Shishkin, A.A., Zhu, X., Kadri, S., Maza, I., Guttman, M., Hanna, J.H., Regev, A., and Garber, M. (2016). Evolutionary analysis across mammals reveals distinct classes of long non-coding RNAs. Genome Biol. 17, 19.

Chen, L.L. (2016). Linking long noncoding RNA localization and function. Trends Biochem. Sci. 41, 761–772.

Clark, C.D., Zhang, B., Lee, B., Evans, S.I., Lassar, A.B., and Lee, K.H. (2013). Evolutionary conservation of Nkx2.5 autoregulation in the second heart field. Dev. Biol. 374, 198–209.

Delaughter, D.M., Bick, A.G., Wakimoto, H., Mckean, D., Gorham, J.M., Kathiriya, I.S., Hinson, J.T., Homsy, J., Gray, J., Pu, W., et al. (2016). Single-cell resolution of temporal gene

expression during heart development. Dev. Cell *39*, 480–490.

Dickel, D.E., Barozzi, I., Zhu, Y., Fukuda-Yuzawa, Y., Osterwalder, M., Mannion, B.J., May, D., Spurrell, C.H., Plajzer-Frick, I., Pickle, C.S., et al. (2016). Genome-wide compendium and functional assessment of in vivo heart enhancers. Nat. Commun. *7*, 12923.

Fazal, F.M., Han, S., Parker, K.R., Kaewsapsak, P., Xu, J., Boettiger, A.N., Chang, H.Y., and Ting, A.Y. (2019). Atlas of subcellular RNA localization revealed by APEX-seq. Cell *178*, 473–490 e26.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of chromosomal domains by loop extrusion. Cell Rep. *15*, 2038–2049.

Greco, C.M., Kunderfranco, P., Rubino, M., Larcher, V., Carullo, P., Anselmo, A., Kurz, K., Carell, T., Angius, A., Latronico, M.V., et al. (2016). DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. Nat. Commun. *7*, 12418.

Grote, P., Wittler, L., Hendrix, D., Koch, F., Wahrisch, S., Beisaw, A., Macura, K., Blass, G., Kellis, M., Werber, M., et al. (2013). The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. Dev. Cell *24*, 206–214.

Han, P., Li, W., Lin, C.H., Yang, J., Shang, C., Nuernberg, S.T., Jin, K.K., Xu, W., Lin, C.Y., Lin, C.J., et al. (2014). A long noncoding RNA protects the heart from pathological hypertrophy. Nature *514*, 102–106.

Hsieh, C.L., Fei, T., Chen, Y., Li, T., Gao, Y., Wang, X., Sun, T., Sweeney, C.J., Lee, G.S., Chen, S., et al. (2014). Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. Proc. Natl. Acad. Sci. U S A *111*, 7319–7324.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182–187.

Kim, Y.J., Xie, P., Cao, L., Zhang, M.Q., and Kim, T.H. (2018). Global transcriptional activity dynamics reveal functional enhancer RNAs. Genome Res. *28*, 1799–1811.

Klattenhoff, C.A., Scheuermann, J.C., Surface, L.E., Bradley, R.K., Fields, P.A., Steinhauser, M.L., Ding, H., Butty, V.L., Torrey, L., Haas, S., et al. (2013). Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell *152*, 570–583.

Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature *494*, 497–501.

Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat. Rev. Genet. *17*, 207–223.

Lien, C.L., Wu, C., Mercer, B., Webb, R., Richardson, J.A., and Olson, E.N. (1999). Control of early cardiac-specific transcription of Nkx2-5 by a GATA-dependent enhancer. Development *126*, 75–84.

Liu, X., Ma, Y., Yin, K., Li, W., Chen, W., Zhang, Y., Zhu, C., Li, T., Han, B., Wang, S., et al. (2019). Long non-coding and coding RNA profiling using strand-specific RNA-seq in human hypertrophic cardiomyopathy. Sci. Data *6*, 90.

Luna-Zurita, L., Stirnimann, C.U., Glatt, S., Kaynak, B.L., Thomas, S., Baudin, F., Samee, M.A., He, D., Small, E.M., Mileikovsky, M., et al. (2016). Complex interdependence regulates heterotypic transcription factor distribution and coordinates cardiogenesis. Cell *164*, 999–1014.

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

Papait, R., Cattaneo, P., Kunderfranco, P., Greco, C., Carullo, P., Guffanti, A., Vigano, V., Stirparo, G.G., Latronico, M.V., Hasenfuss, G., et al. (2013a). Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. Proc. Natl. Acad. Sci. U S A *110*, 20164–20169.

Papait, R., Kunderfranco, P., Stirparo, G.G., Latronico, M.V., and Condorelli, G. (2013b). Long noncoding RNA: a new player of heart failure? J. Cardiovasc. Transl. Res. *6*, 876–883.

Pashmforoush, M., Lu, J.T., Chen, H., Amand, T.S., Kondo, R., Pradervand, S., Evans, S.M., Clark, B., Feramisco, J.R., Giles, W., et al. (2004). Nkx2-5 pathways and congenital heart disease; loss of ventricular myocyte lineage specification leads to progressive cardiomyopathy and complete heart block. Cell *117*, 373–386.

Qu, X., Du, Y., Shu, Y., Gao, M., Sun, F., Luo, S., Yang, T., Zhan, L., Yuan, Y., Chu, W., et al. (2017). MIAT is a pro-fibrotic long non-coding RNA governing cardiac fibrosis in post-infarct myocardium. Sci. Rep. *7*, 42657.

Rosa-Garrido, M., Chapski, D.J., Schmitt, A.D., Kimball, T.H., Karbassi, E., Monte, E., Balderas, E., Pellegrini, M., Shih, T.T., Soehalim, E., et al. (2017). High-resolution mapping of chromatin conformation in cardiac myocytes reveals structural remodeling of the epigenome in heart failure. Circulation *136*, 1613–1625.

Ross, J., Jr. (1983). Cardiac function and myocardial contractility: a perspective. J. Am. Coll. Cardiol. *1*, 52–62.

Ryall, J.G., Dell'orso, S., Derfoul, A., Juan, A., Zare, H., Feng, X., Clermont, D., Koulnis, M., Gutierrez-Cruz, G., Fulco, M., et al. (2015). The NAD(+)-dependent SIRT1 deacetylase translates a metabolic switch into regulatory epigenetics in skeletal muscle stem cells. Cell Stem Cell *16*, 171–183.

Saadane, N., Alpert, L., and Chalifour, L.E. (1999). Expression of immediate early genes, GATA-4, and Nkx-2.5 in adrenergic-induced cardiac hypertrophy and during regression in adult mice. Br. J. Pharmacol. *127*, 1165–1176.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. *15*, 272–286.

Sigova, A.A., Abraham, B.J., Ji, X., Molinie, B., Hannett, N.M., Guo, Y.E., Jangi, M., Giallourakis, C.C., Sharp, P.A., and Young, R.A. (2015). Transcription factor trapping by RNA in gene regulatory elements. Science *350*, 978–981.

Tang, X., Ma, H., Han, L., Zheng, W., Lu, Y.B., Chen, X.F., Liang, S.T., Wei, G.H., Zhang, Z.Q., Chen, H.Z., et al. (2016). SIRT1 deacetylates the cardiac transcription factor Nkx2.5 and inhibits its transcriptional activity. Sci. Rep. *6*, 36576.

Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. Cell *151*, 206–220.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120–124.

Wang, Z., Zhang, X.J., Ji, Y.X., Zhang, P., Deng, K.Q., Gong, J., Ren, S., Wang, X., Chen, I., Wang, H., et al. (2016). The long noncoding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy. Nat. Med. *22*, 1131–1139.

Supplemental Information

Divergent Transcription of the *Nkx2-5*

Locus Generates Two Enhancer

RNAs with Opposing Functions

Irene Salamon, Simone Serio, Simona Bianco, Christina Pagiatakis, Silvia Crasto, Andrea M. Chiariello, Mattia Conte, Paola Cattaneo, Luca Fiorillo, Arianna Felicetta, Elisa di Pasquale, Paolo Kunderfranco, Mario Nicodemi, Roberto Papait, and Gianluigi Condorelli
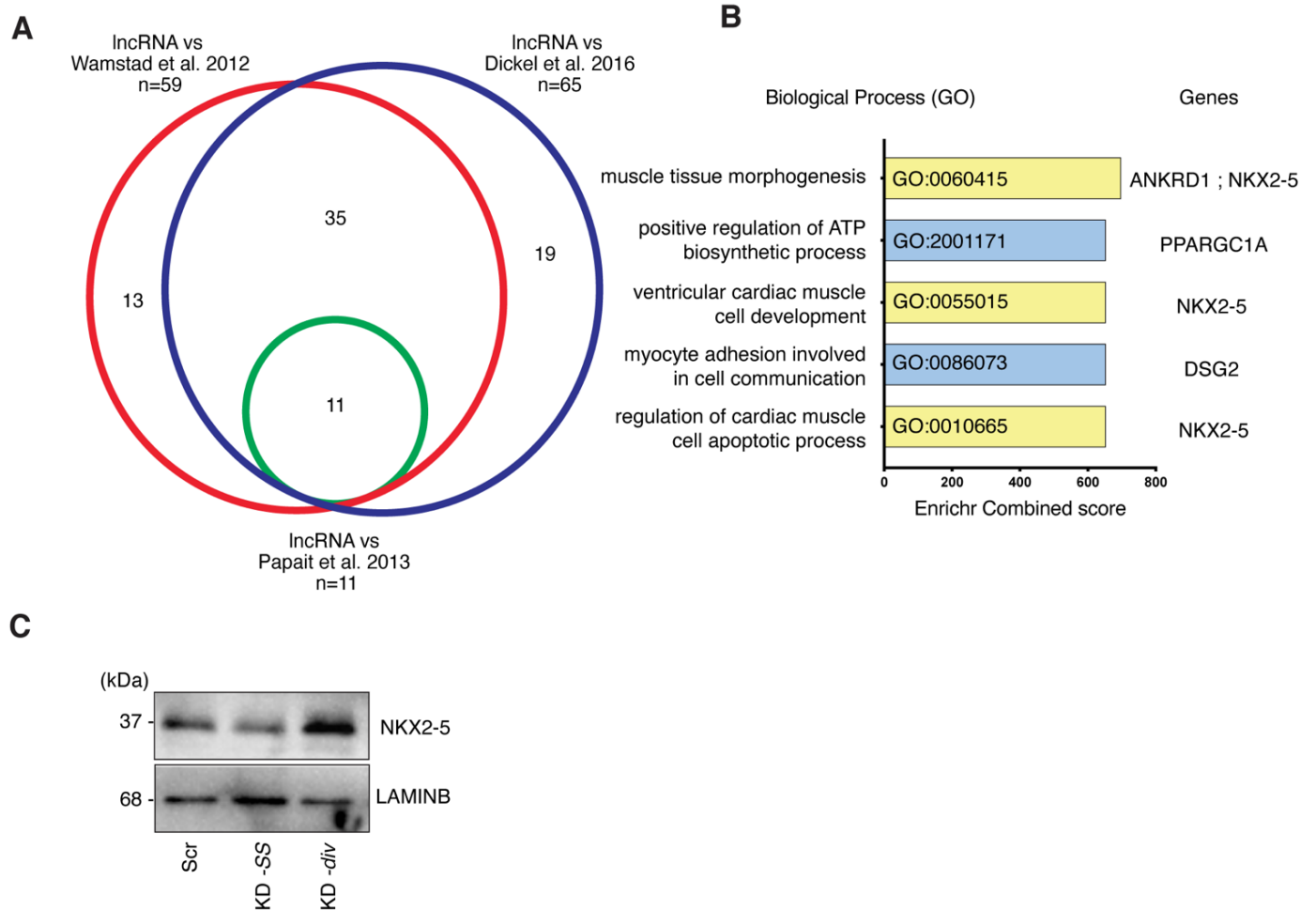
SUPPELEMENTAL FIGURES

Figure S1. Identification of *IRENE*-SS and *IRENE*-div Upstream of *Nkx2-5*, Related to Figure 1
(A) Venn diagram of the overlap between the identified lncRNAs and three enhancer datasets (Papait et al., 2013a; Wamstad et al., 2012; Dickel et al., 2016).
(B) Enriched biological processes obtained through gene ontology (GO) analysis of the 26 protein-coding genes associated with 46 putative eRNAs. *NKX2-5* is associated with muscle tissue morphogenesis, ventricular cardiac muscle cell development, and regulation of cardiac muscle cell apoptotic processes.
(C) Western blot showing downregulation of NKX2-5 after knockdown (KD) of *IRENE-SS*. Cells were collected 48 h after gapmeR transfection. Scr, scrambled gapmeR; KD-*SS*, gapmeR targeting *IRENE-SS*; KD-*div*, gapmeR targeting *IRENE-div*.
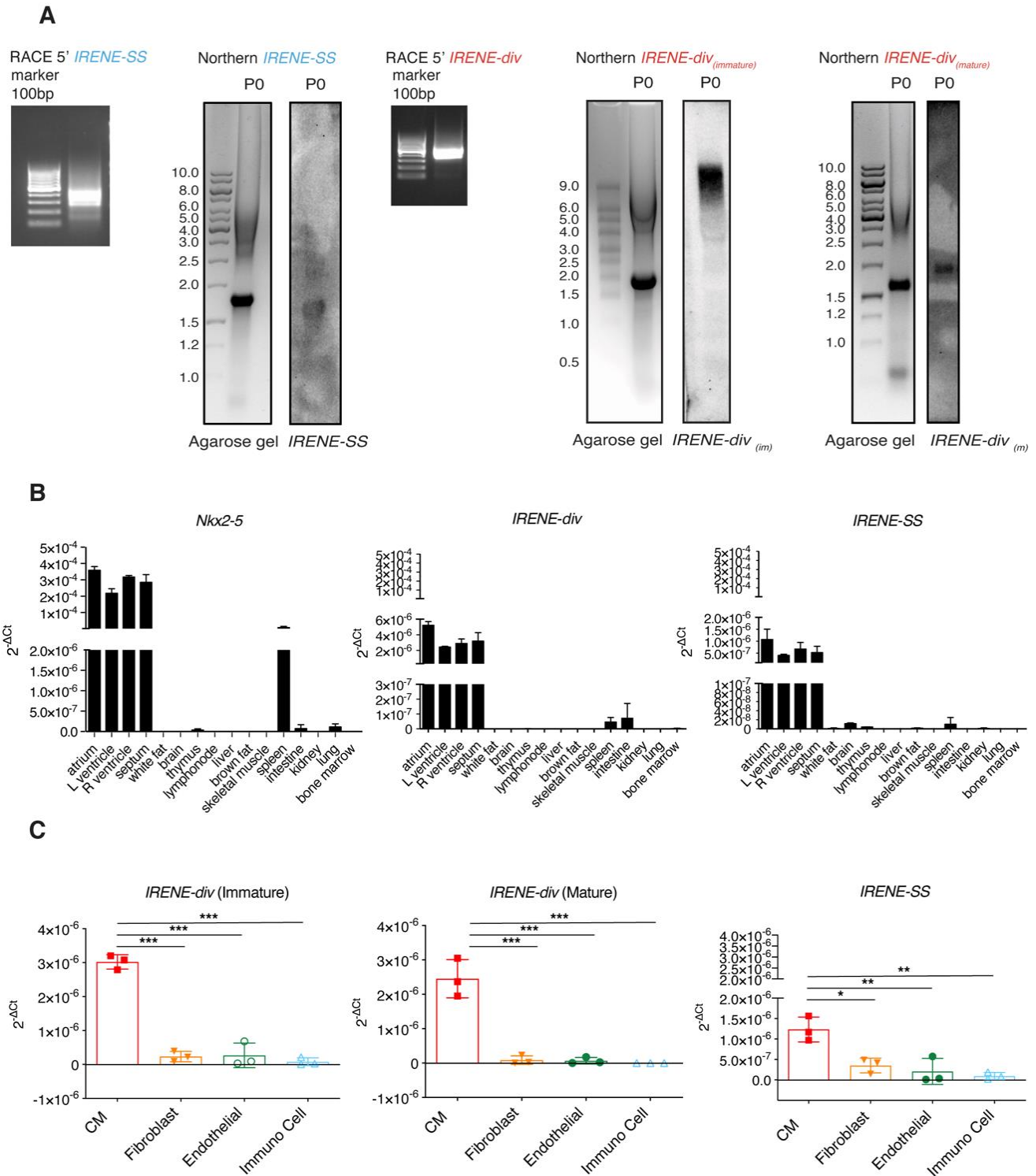
Figure S2. Characterization of *IRENE*-SS and *IRENE*-div Upstream of *Nkx2-5* in Mice, Related to Figure 1

(A) 5' RACE product and Northern blotting for *IRENE-SS* and *IRENE-div,* the latter in its two forms (immature/unspliced and mature).

(B) qRT-PCR analysis of *Nkx2-5*, *IRENE-div*, and *IRENE-SS* in 13 different mouse tissues revealed enrichment in four different cardiac districts.

(C) *IRENE* transcript expression in sorted adult heart cell populations (cardiomyocytes, fibroblasts, endothelia and immune cells), revealing cardiomyocyte-specific enrichment. CM, cardiomyocyte.

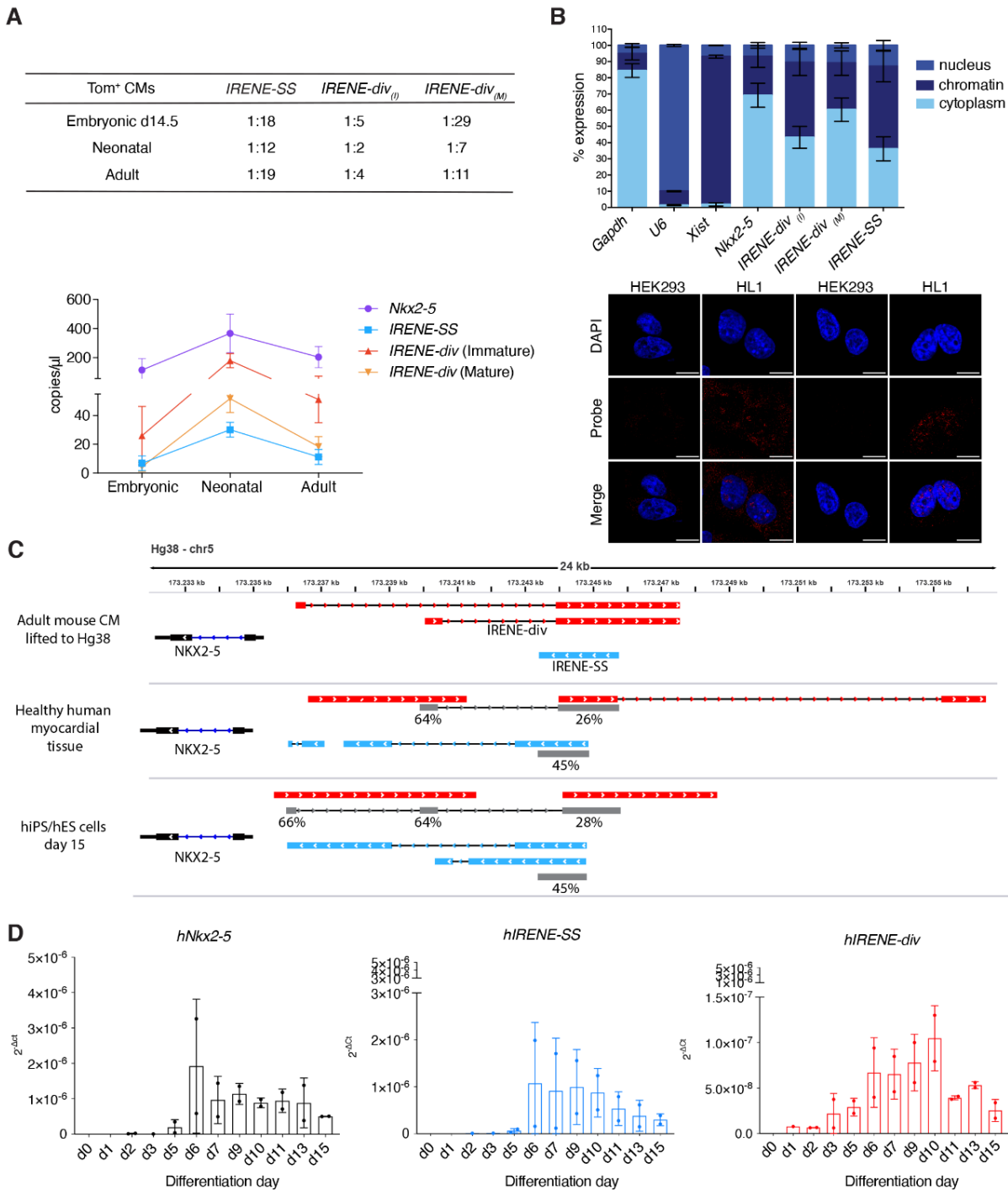Data represented as mean ± SD. *, P-value <0.05; **, P-value <0.01; ***, P-value <0.001.

Figure S3. Expression of *IRENE-SS* and IRENE-*div* Upstream of *Nkx2-5* in Mice and Humans, Related to Figure 1

(A) Expression of *Nkx2-5* and *IRENE* transcripts measured by ddPCR at 3 different cardiomyocyte maturation stages: embryonic day 14.5, neonatal day p1, and 8-week-old adult mice. The table indicates the ratio between the copies of *IRENE* and the nearby gene (*Nkx2-5*) at the three different stages, and the graph shows expression reported as copies/µl.

(B) HL1 cell fractionation (nucleus, chromatin, and cytoplasm, as indicated by the different shades) (top) and RNA-FISH (bottom) assessing intracellular localization of *IRENE* transcripts: *IRENE-SS* and immature *IRENE-div* were 60% in the nucleus and mostly chromatin bound, whereas mature *IRENE-div* was 60% in the cytoplasm. *Gapdh*, *U6,* and *Xist* were used as RNA controls for the 3 different compartments. Bars = 10 µm.

(C) Transcript-Transcript Identity scores showing the percentages of identity between mouse and human transcripts.

(D) *Nkx2-5*, *IRENE-SS*, and *IRENE-div* expression in human iPSCs at 12 different time points during cardiomyocyte differentiation (N=2).
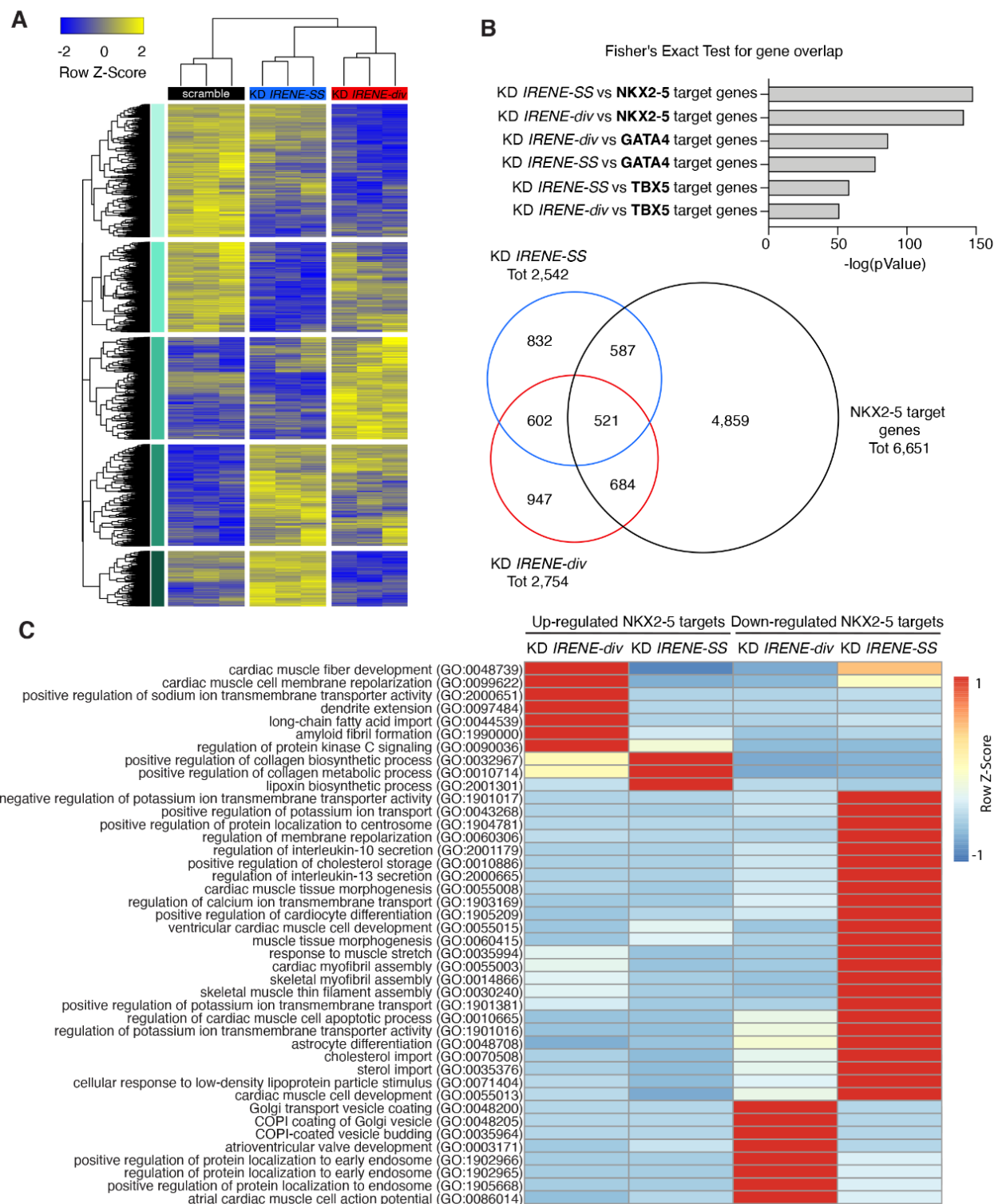
Data represented as mean ± SD.

Figure S4. Differential Regulation of *Nkx2-5* by *IRENE-SS* and *IRENE-div*, Related to Figure 1

(A) Heatmap of differentially expressed genes for each gapmeR treatment (FDR ≤0.1; log CPM ≥0).

(B) Bar chart of significant overlaps between genes associated with three TFs (NKX2-5, GATA4, and TBX5) and those modulated upon knockdown of either *IRENE-SS* or *IRENE-div*. Many more modulated genes were targets of NKX2-5 rather than of the other two TFs evaluated. The Venn diagram gives the number of modulated genes that were NKX2-5 targets.

(C) Enriched biological process gene ontology (GO) terms for genes up- and down-regulated by silencing *IRENE-SS* or *IRENE-div*, respectively. The analysis reveals that highlighted terms are oppositely regulated by the silencing of *IRENE-SS* or *IRENE-div*, confirming the opposing effects mediated by the two eRNAs on NKX2-5 target genes.
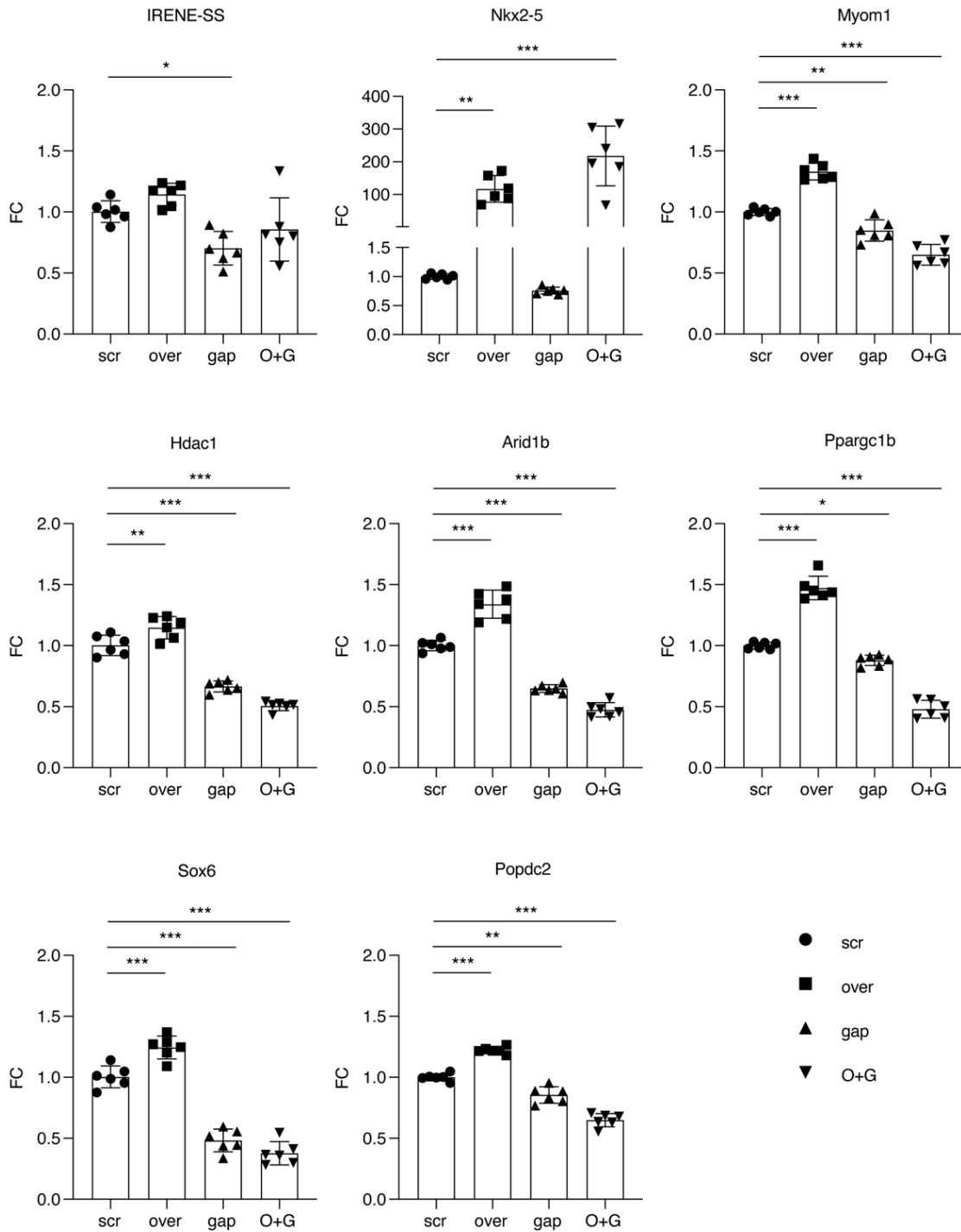
Figure S5. *IRENE-SS* Acts on NKX2-5 Target Genes Independently of NKX2-5 Level, Related to Figure 1
For the six NKX2-5 target genes analysed, overexpression of NKX2-5 increased expression, whereas knockdown of *IRENE-SS* caused a decrease. However, the overexpression of NKX2-5 in neonatal cardiomyocytes with knocked-down *IRENE-SS* did not restore correct expression of the genes. Scr, scramble-treated cells (transfected with a control vector and the control gapmeR); over, cells treated with an NKX2-5 overexpression vector; gap, cells treated with gapmeR targeting *IRENE-SS;* O+G, cells treated with the combination of overexpression vector and gapmeR targeting *IRENE-SS.*
Data represented as mean ± SD. *, P-value <0.05; **, P-value <0.01; ***, P-value <0.001 (one-way ANOVA with Sidak post hoc test).

Figure S6. *IRENE-div* Recruits Histone Deacetylase to the *Nkx2-5* Enhancer and its Reduced Impact on Acetylation Status, Related to Figure 3

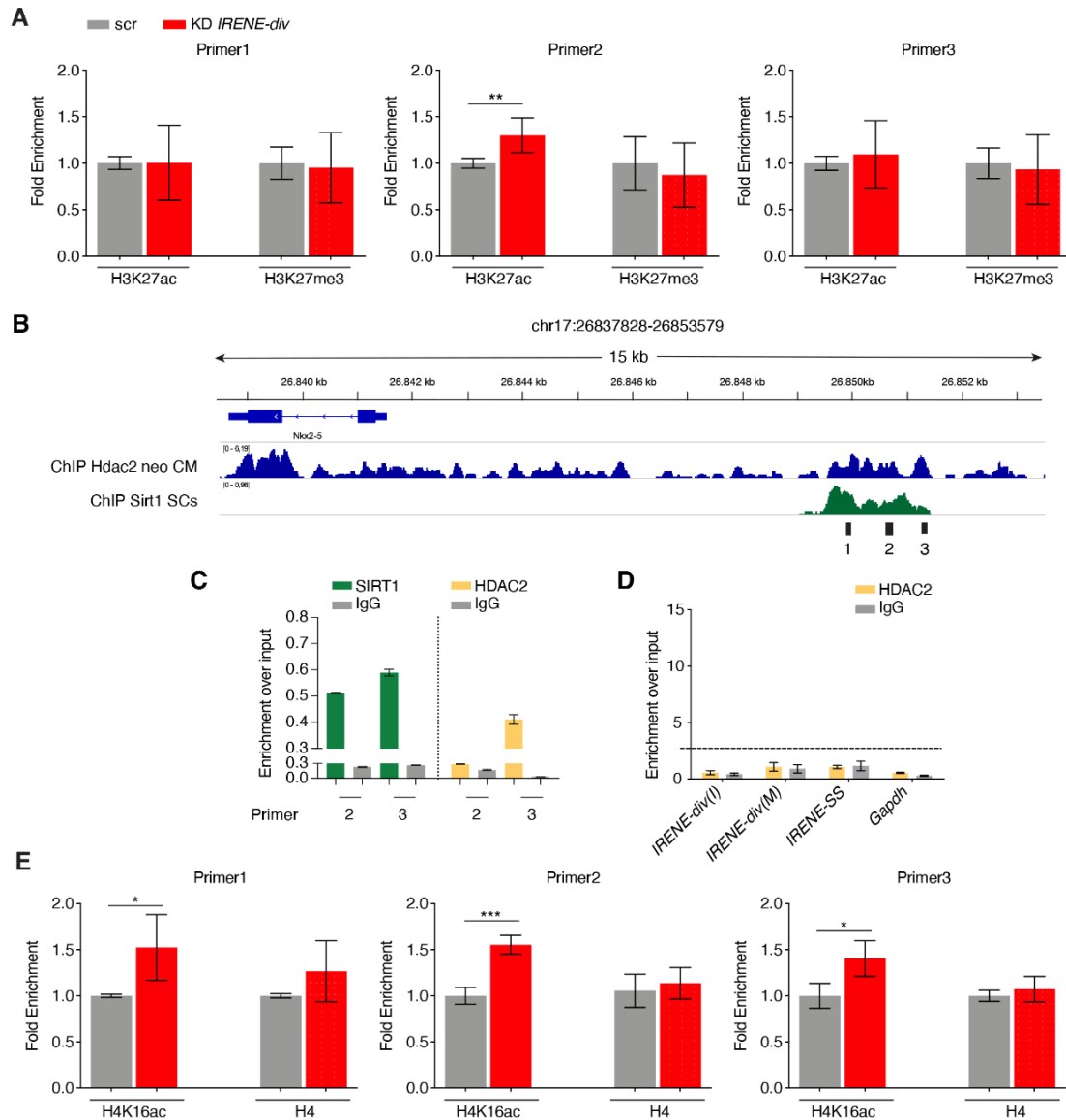(A) ChIP−qRT-PCR conducted on neonatal cardiomyocytes revealed an enrichment of H3K27ac on the enhancer region (significant with primer 2) upon *IRENE-div* silencing, with no effect on methylation (H3K27me3). Values are expressed as fold enrichment on scr-treated cells. N=3 biological replicates.

(B) Cardiomyocyte ChIP-seq tracks of HDAC2 and SIRT1 from published datasets on neonatal CMs (neo CM) and skeletal muscle satellite cells (SCs) (Ai et al., 2017; Plimpton, 1995), showing peaks at the enhancer locus.

(C) Validation of ChIP-seq peaks from published datasets of two deacetylases − SIRT1 (green) and HDAC2 (yellow) − through ChIP−qRT-PCR performed on HL1 cells using primers 2 and 3 on the enhancer element.

(D) RIP performed on HL1 cells, using antibodies against HDAC2, showing that immature and mature forms of *IRENE-div* do not bind HDAC2. Interaction is expressed as enrichment over input in N=3 different experiments (one-way ANOVA with Sidak post hoc test). I, immature form; M, mature form.

(E) ChIP−qRT-PCR conducted on neonatal cardiomyocytes after silencing of *IRENE-div* (red bar) compared with scrambled gapmeR-treated cells (grey bar). We used the antibodies against H4K16ac and H4, and three primers (1, 2, and 3) designed on genomic enhancer elements. Knockdown of *IRENE-div* led to a decrease in the enrichment of SIRT1 and a consequential increase in H4K16ac, the histone modification mark induced by SIRT1, whereas no change was detected in the control (H4 immunoprecipitation). Values are expressed as fold enrichment over scr-treated cells. N=2 biological replicates (Student's unpaired t-test).

Data represented as mean ± SD. *, P-value <0.05; **, P-value <0.01; ***, P-value <0.001.

TRANSPARENT METHODS

Annotation, Identification, and Conservation of lncRNAs and eRNAs

Adult mouse cardiomyocyte RNA-seq data (from GSE66847 and GSE96690), human RUES2 cells, iPSC-derived CMs at day 15 of differentiation, and myocardial tissue from nine healthy humans (from GSE130036) were mapped using HISAT2 (Kim et al., 2019) and aligned against GRCm38 and GRCh38 for mouse and human samples, respectively. BAM files with aligned paired-end reads were split into two strand-specific BAM files using SAMtools (Li et al., 2009). The properly paired reads originating from each strand were selected using SAM Flags 99 and 147 for one strand, 83 and 163 for the other strand. Forward (Fw) and reverse (Rw) strand was assigned according to the dUTP method used for sequencing. The transcript assembly was performed on each strand-specific BAM file, using StringTie (Pertea et al., 2015) with parameters -s 2 -g 100 for lncRNA identification. The potential transcripts identified were unified using StringTie –merge with parameters -T 0 -F 0, according to each set of replicated biological data. Transcripts with FPKM <0.1 were subsequently excluded using FPKM_count.py module from RSeQC software (Wang et al., 2012). Slncky (Chen et al., 2016), with default parameters, was used to predict the orthologous transcripts in human samples (hiPSC-derived CMs and healthy human myocardial tissue). Only orthologues identified in both datasets were considered. Three published enhancer datasets (Papait et al., 2013a; Wamstad et al., 2012; Dickel et al., 2016) were used to select the putative eRNAs, identified as lncRNAs that overlapped with enhancer genomic regions. The overlap between genomic regions was performed with IntersectBed (Quinlan and Hall, 2010). The list of 26 protein-coding genes associated with the eRNAs was then analysed using the online tool Enrichr (Kuleshov et al., 2016) to identify GO enrichments. The presented combined score (c) represents the multiplication of the p-value (Fisher exact test) and the z-score (z) of the deviation from the expected rank ($c=\log(p) \cdot z$; for more details, see: http://amp.pharm.mssm.edu/Enrichr/.

Mouse RNA-sequencing

Three biological replicates of neonatal primary CM cultures for each specific gapmeR treatment were profiled. DNase I-treated RNA analysed on the Tape Station 4200 (Agilent) showed an RNA integrity number (RIN) of 9.7−9.8. SMART-seq v4 Ultra Low Input RNA Kits (Takara Bio) were used for cDNA amplification. Indexed sequencing libraries were generated using Nextera XT DNA Library Preparation Kits (Illumina) following the manufacturer's instructions. Electrophoresis of the libraries on a LabChip GX instrument showed average library size of 500−555 bp. Single-end multiplexed libraries were sequenced using a NextSeq500 instrument. 19−34 million reads/sample were obtained. Single-end reads of 75 bp were aligned to the GENCODE *Mus musculus* reference genome (build GRCm38/mm10) using STAR v2.7.2b (Dobin et al., 2013). Raw read counts were normalized with TMM implemented in edgeR (Robinson et al., 2010); differential expression analysis of read counts was performed using glmQLFit and glmQLFTest functions. Significant differential genes were chosen based on an FDR ≤0.1 and logCPM ≥0. Hierarchical clustering of significantly modulated genes was performed using hclust and dist R functions (R Core Team, 2019). Clustering was performed with complete linkage method and Euclidean distance to generate a heatmap using pheatmap (Kolde, 2019).

Human RNA-sequencing

One biological replicate of human RUES2 cells and one of iPSC-derived CMs at day 15 of differentiation were profiled. DNase I-treated RNA analysed on the LabChip GXT (PerkinElmer) had RIN >7. Indexed sequencing libraries were generated using TruSeq Stranded Total RNA LT with Ribo-Zero Human/Mouse/Rat (Illumina), following the manufacturer's instructions. Electrophoresis of the libraries on a LabChip GXT instrument showed an average library size of 311bp. Paired-end multiplexed libraries were sequenced using the NextSeq500 instrument. 96−104 million reads for each sample were obtained.

Nkx2-5, *Tbx5*, and *Gata4* Gene Analysis

Nkx2-5, *Tbx5*, and *Gata4* average footprints obtained from *in vitro* differentiated mouse CMs (GSE72223) were used. Only ChIP-seq peaks that fell on gene promoters (−3kb/+2kb from the TSS) and/or enhancer genomic regions were selected. The NKX2-5, TBX5, and GATA4 target gene sets were identified using rGreat (Gu, 2019), using "oneClosest" rule and "adv_oneDistance = 500" to associate the selected regions to the nearby gene. Fisher test R function was used to test the significant overlap between modulated genes obtained from RNA-seq of the *IRENE-SS* and *IRENE-div* knockdown and the TF's associated targets identified. GO enrichment analysis was performed with enrichR (Jawaid, 2019) and, using pheatmap, we plotted the raw z-score computed by centring and scaling the "enrichR combined scores". RNA-seq data has been deposited in the NCBI GEO database (GSE143930, GSE143929).

Primary Cardiomyocyte Culture

Primary CMs were isolated from 1-day-old pups with the Neonatal Heart Dissociation Kit (Miltenyi). Cells were cultured in M-199, DMEM, 1% penicillin/streptomycin (Pen-Strep, 10000 U/ml), 1% ultraglutamine, 10% horse serum, and 5% foetal bovine serum (FBS).

HL1 Cell Culture

HL-1 cells, a mouse atrial CM immortalized cell line, were grown in Claycomb Medium (Sigma-Aldrich) supplemented with 10% FBS (Sigma-Aldrich), 1% Pen-Strep, 1% ultraglutamine (200 mM), and 1 mM norepinephrine (Sigma-Aldrich) on gelatin/fibronectin pre-coated flasks. After reaching full confluence, cells were split according to Dr. Claycomb's instructions (Claycomb et al., 1998).

GapmeR Silencing and NKX2-5 Overexpression

Transfections were carried out using standard Lipofectamine 2000 (ThermoFisher) protocols. GapmeR silencing was performed using 100 nM of scrambled, anti-*IRENE-SS*, or anti-*IRENE-div* sequences (QIAGEN). Cells were harvested after 24 h and 48 h for qRT-PCR, 48 h for Western blotting, and 24 h for ChIP−qRT-PCR. The antisense LNA GapmeR standard sequences were: scr negative control A 5'-3': AACACGTCTATACGC; anti-*IRENE-SS* 5'-3': GAAGGTGATAATTTGA; and anti-*IRENE-div* 5'-3': TTCGGGTCAAGAGTGC.

Overexpression was conducted using 250 ng of Nkx2-5 overexpression pcDNA3 vector. The rescue experiment was performed combining the transfection of 100 nM of gapmeR anti-*IRENE-SS* and 250 ng of overexpression vector in neonatal cardiomyocytes or the same amount of scrambled gapmeR and empty control vector. Cells were collected after 24 h and processing for qRT-PCR.

Human Induced Pluripotent Stem Cell-derived Cardiomyocytes

An erythroid body-based protocol was used to generate functional cardiomyocyte-like cells from feeder-free female and male hiPSCs. Human samples were collected after written informed consent by the patients. The expression of markers typical of pluripotent cells and karyotype was used to control cell lines, as described elsewhere (Di Pasquale et al., 2013). hiPSCs were differentiated using the monolayer differentiation protocol as described elsewhere (Salvarani et al., 2019). In brief, iPSCs were treated with 12 µM CHIR99021 and 5 µM IWR and cultured until they started to spontaneously beat.

The following twelve different stages of hiPSC-CMs differentiation were used for the experiments: days 0, 1, 2, 3, 5, 6, 7, 9, 10, 11, 13, and 15. Two sets of differentiation experiments were used and merged. RNA was extracted, retro-transcribed, and used for qRT-PCR. The primer sequences for the gene expression analyses were:

hNKX2-5:   Fw GAGCCGAAAAGAAAGCCTGAA,   Rw CACCGACACGTCTCACTCAG;
hIRENE-div: Fw GACATCACCCTCCCATTCCT,   Rw AGAGTCATCTCCAGGGCTACC;
hIRENE-SS: Fw GGCAGACCCCACACTTCA,   Rw TTCCACGCTTGTTCAGTGC.

Sorting of Pure Cardiomyocytes and Heart Cell Populations

To isolate pure CM populations, we used a constitutively active CM-specific tdTomato reporter mouse (αMHC Cre/+;R26tdTomato) that expresses Cre recombinase protein under control of the alpha myosin heavy chain (αMHC) promoter (Abel et al., 1999) in addition to a loxP site-dependent tdTomato reporter within the ubiquitous *Rosa26* gene locus, in the C57BL/6J background (Larcher et al., 2018). Cells were sorted to isolate CMs (Tom+), fibroblasts (CD45−/PDGFα+), endothelial cells (CD45−/CD31+), and leukocytes (CD45+). All experiments were performed according to the 2010/63/EU Directive and approved by the ethics committee of Humanitas Research Hospital.

RNA Preparation

RNA was extracted using PureZOL isolation reagent (BioRad), following the manufacturer's instructions, for mouse tissues, Northern blotting, and RIP. The Zymo Extraction Kit (EuroClone), plus DNAse treatment on columns, was used for RNA-seq, qRT-PCR, and ddPCR.

Quantitative RT-PCR

RNA was retro-transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) for gene expression, and the SuperScript Vilo cDNA Synthesis Kit (Invitrogen) for fractionation and RIP. qRT-PCR was performed using the SYBR Green PCR Master Mix (Applied Biosystems). Primers used were:

IRENE-SS: Fw CGTTTATGCTGTAGCGCTGT,   Rw CTTGGGGCTGTGATTGATTT;
IRENE-div (immature): Fw AGCATGAAGGAGCGTGACTT,   Rw GTGCAGGTACAGGGGTTGTT;
IRENE-div (mature):   Fw AGGACCTTGTTACTGGCACTG, Rw TGTCTGTCTTTCTGTCTGGAGAA.

Hprt was used to normalize Ct values, and the $2^{-\Delta Ct}$ method and fold change were conducted to analyse data.

Droplet Digital PCR

The QX200 ddPCR system (BioRad) and an EvaGreen-based assay were used. The PCR reaction mix was loaded into 96-well plates in an automated droplet generator. The plate was sealed, placed in the thermal cycler, and after the PCR reaction put into the droplet reader, experimental details set using QuantaSoft (BioRad), and the reading started. Data analysis was performed using QuantaSoft and values expressed as copies/µl.

RNA-Fluorescent in-situ hybridization (FISH)

Custom Stellaris FISH probes were designed against *IRENE-SS* and *IRENE-div* by utilizing the Stellaris RNA FISH Probe Designer (Biosearch Technologies) available online at www.biosearchtech.com/stellarisdesigner (version 4.2). HL1 and HEK293 cells were hybridized with the *IRENE-SS* and *IRENE-div* Stellaris RNA-FISH probe set labelled with TAMRA-C9 and QUASAR 670 dye, respectively (Biosearch Technologies), following the manufacturer's instructions available online at www.biosearchtech.com/stellarisprotocols. Briefly, the cells were fixed with 3.7% formaldehyde and permeabilized 1 h at 4°C with 70% ethanol. The coverslips were washed and hybridized overnight at 37°C in Hybridization Buffer with 2μl of specific probe (stock solution: 12.5 μM). The coverslips were then washed, stained with DAPI, and mounted on slides using one drop of FluorSave Reagent (Millipore).

Cell Fractionation

HL1 cells were collected in phosphate-buffered saline (PBS) and lysed in Cytoplasm Buffer (50 mM Tris-HCl [pH 8.0], 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40) supplemented with RNAsin (Promega) for 10 min on ice. After centrifugation at 500 rpm for 5 min, the supernatant was collected as the cytoplasmic fraction. The remaining pellet was washed in 1X PBS and resuspended in Nucleoplasm Buffer (50 mM Tris-HCl [pH 8.0], 500 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40) supplemented with RNAsin for 10 min on ice. Chromatin was pelleted at maximum speed for 5 min. Supernatant was collected as the nucleoplasm and the chromatin pelleted in 1X PBS. RNA from cytoplasmic and nucleoplasm fractions were precipitated using $C_2H_3NaO_2$. RNA was extracted using the Zymo Kit supplemented with DNAse. The Super Script VILO cDNA Synthesis Kit (Invitrogen) was used to retro-transcribe RNA, with ddPCR to quantify target genes.

Chromatin Immunoprecipitation

For ChIP−qRT-PCR, cells were crosslinked in 1% formaldehyde and lysed to prepare nuclear extracts. Chromatin was sonicated to an average length of 200–300 bp using a Bioruptor Pico (Diagenode), clarified, precleared for 1 h, and immunoprecipitation performed at 4°C overnight with protein G Dynabeads (Invitrogen) coupled with the following antibodies (2 μg for histone modification, 10 μg for transcription factor and HDAC): anti-H3K27ac (ab4729, Abcam), anti-H3K27me3 (07-449, Millipore), anti-H4 (ab10158, Abcam), anti-H4K16ac (17-10101, Millipore), anti-NKX2-5 (sc-8697, Santa Cruz), anti-SIRT1 (ab12193, Abcam), anti-HDAC2 (ab7029, Abcam). The washes were made in Buffer I (1 mM EDTA, 0.1% SDS, 10 mM Tris HCl [pH 8.0], 140 mM NaCl, 0.1% sodium deoxyxholate, 1% Triton X-100), Buffer II (1 mM EDTA, 0.1% SDS, 10 mM Tris HCl [pH 8.0], 500 mM NaCl, 0.1% sodium deoxyxholate, 1% Triton X-100), Buffer III (1 mM EDTA, 10 mM Tris HCl [pH 8.0], 250 mM LiCl, 0.5% sodium deoxycholate, 0.5% NP-40), and Tris-EDTA (pH 8.0). Crosslinking was reverted overnight in Elution Buffer (5 mM EDTA, 10 mM TrisHCl [pH 8.0], 300 mM NaCl, 0.5% SDS) shaking at 65°C. Proteinase K treatment for 2 h shaking at 45°C. DNA was extract with phenol:chloroform extraction. Primer sequences for ChIP−qRT-PCR were:

ChIP_1: Fw GCCATTCGCTCCTAGCTTC,　　Rw GGGCAGTGACTCTTGGACTC;
ChIP_2: Fw CCTCGCTCCAGTCAAACTTC,　Rw CCACACAGCAGAGGGTTTGT;
ChIP_3: Fw ACTTTGGAAGGGGAGAGGAG, Rw GGAAGTGGAGCAGTTGAAGC;
ChIP_4: Fw TCTCTGCCCTGGATGTGG,　　 Rw CCCCTACTCCAGCCTGCTC;
ChIP_5: Fw CTGTCCTTCGCGTTCCTG,　　 Rw AGTGACCCGCTCCATGTC;
Ctrl:　　 Fw GTAGATGGGATGGGTCATGG, Rw TGCTTTGAGGTTGCTTTGTG.

Western Blotting

Cells were washed with PBS and lysed in NP-40 lysis buffer (50 mM Tris, 150 mM NaCl, 0.5% NP-40, 2 mM EDTA, 100 mM NaF, 10 mM sodium pyrophosphate) containing Protease Inhibitor Cocktail (Roche) and 1 mM phenylmethylsulfonyl fluoride. Protein concentrations were determined by the Bradford assay and 30 μg of total protein was resolved on SDS-PAGE and blotted onto nitrocellulose (BioRad) for 1 h. Non-specific binding sites were blocked using 5% milk in PBS-Tween. Membranes were incubated with primary antibodies overnight at 4°C in 5% bovine serum albumin in PBS. Primary antibodies were: anti-Nkx2-5 (1:500; Santa Cruz), anti-lamin B (1:500; Santa Cruz), anti-tubulin (1:2000, Abcam), and anti-Sirt1 (1:1000, Millipore). Membranes were incubated with the HRP-conjugated secondary antibody (donkey anti-goat-HRP, goat anti-rabbit-HRP, goat anti-mouse-HRP, 1:2000 in 5% milk-PBS-T) for 1 h at room temperature. Protein was detected with Enhanced Chemiluminescence Western Blotting Substrate (Immobilon).

Northern Blotting

RNA was extracted using PureZOL isolation reagent (BioRad) and mixed in a 1:1 ratio with RNA sample loading buffer, boiled for 6 min at 70°C, and run on 1% agarose gel (1X MOPS buffer, 1% formaldehyde) in 1X MOPS buffer at 70 V for 3 h. The gel was subsequently transferred to a positively charged nylon membrane in 20X SSC buffer by standard capillary transfer overnight. RNA was crosslinked to the membrane using UV fixation at 200 mJ for 3 min. Membrane was pre-hybridized in NorthernMax Pre-hybridization/Hybridization Buffer for 3 h at 60°C. Membranes were incubated overnight with DIG-conjugated probes diluted in Hybridization Buffer at 60°C. Membranes were then washed in Northern blot low

stringency buffer (3 times, 10 min/wash) at 60°C, followed by washes in Northern blot high stringency buffer (3 times, 10 min/wash) at 60°C, and one wash in 1X PBS-Tween at room temperature for 5 min. Membranes were blocked in 1% blocking solution in 1X PBS-T (Roche) for 30 min at room temperature and then incubated with α-DIG-HRP secondary antibody (1:2000) for 1 h at room temperature. Membranes were washed 3 times in 1x PBS-Tween at room temperature, and RNA visualized using enhanced chemiluminescence.

RNA Immunoprecipitation

For RIP, HL1 cells were crosslinked using 150 mJ UV irradiation, washed in 1X PBS, scraped, and collected by centrifugation at 1000 rpm for 5 min. The cell pellet was lysed in RIP lysis buffer (25 mM Tris-HCl [pH 7.4], 150 mM KCl, 5 mM EDTA, 0.5% NP-40, 0.5 mM DTT, protease inhibitors, and RNA inhibitor) on ice for 30 min. Lysate was centrifuged at 15,000 rpm for 10 min to remove cell debris, and incubated for 1 h (rotating, 4°C) with 50 µL protein-G magnetic beads (washed 3 times in 1X PBS) for pre-clearing. Pre-cleared lysate was then incubated overnight with previously conjugated antibody-protein-G-bead complexes (beads incubated with 10 µg primary antibody overnight at 4°C, rotating). Bead-antibody-lysate complexes were washed three times in 1X PBS, and RNA eluted with 300 µl PureZOL isolation reagent (BioRad), followed by classical RNA extraction.

RNA Pull-down

RNA pull-down was performed as described elsewhere (Yamazaki et al., 2018). Briefly, a portion of *IRENE-SS* or *IRENE-div* was cloned into the pBlueScript II KS (+) vector, and sense or antisense RNAs were synthetized using MAXIscript (Thermo Fisher Scientific) with T7 or T3 RNA polymerase, respectively, and DNase I. RNAs were precipitated with ammonium acetate/ethanol and subsequently biotinylated using the Pierce RNA 3′ End Biotinylation Kit (Thermo Fisher Scientific). HL1 cell extract (CE) was prepared using Cell Extraction Buffer (Thermo Fisher Scientific) following the manufacturer's instructions. The CE was precleared with 50 µl of Dynabeads M-280 Streptavidin (Thermo Fisher Scientific) washed with RNA Pulldown Buffer (1X PBS, 0.1% Triton X-100, 0.6 mM PMSF, Roche Complete Protease Inhibitor Cocktail) and incubated with rotation at 4°C for 1 h. 1 µg of *in vitro*-transcribed RNAs was heated at 90°C for 2 min and then placed on ice for 2 min. An equal volume (10 µl) of RNA structure buffer (20 mM Tris-HCl [pH 7.4], 0.2 M KCl, 20 mM MgCl$_2$) was added and placed at room temperature for 20 min. RNA was then mixed with washed 50µl Dynabeads M-280–Streptavidin and rotated at 4°C for 1 h. The conjugated beads were then washed five times with cold RNA pulldown buffer and incubated with precleared CE in rotation at 4°C for 3 h. Finally, proteins were eluted at 95°C for 5 min in Laemmli sample buffer and load on SDS-PAGE gels.

Immunofluorescence

Cells were fixed in 4% paraformaldehyde, permeabilized with 0.3% triton in PBS, blocked, and incubated with a primary rabbit polyclonal anti-SIRT1 (Millipore) and mouse monoclonal anti-LAMIN B1 (B10, Santa Cruz), both diluted 1:200, and detected using Alexa Fluor 594- or 647-conjugated secondary antibodies (1:500; Life Technology). Cells were visualized using confocal fluorescence microscopy.

3D Modelling

Strings & Binders modelling of the *Nkx2-5* locus: Three-dimensional (3D) conformations of the *Nkx2-5* locus were reconstructed using the SBS polymer chromatin model (Barbieri et al., 2012), which quantitatively describes the biological scenario where chromatin contacts are mediated by proteins, such as TFs, which can bind multiple chromatin sites simultaneously and, thus, bridge them together. Chromatin filaments are represented as self-avoiding chains of beads, some of which are binding sites for specific, diffusing particles named binders. If the binder concentration c and the bead-binder interaction energy E_{int} are above threshold, stable contacts can be formed and the system undergoes coil-globule phase transition. In general, binding sites can be of different types, each interacting with its specific, cognate binders. Schematically, different binding sites can be visualized as different colours, as shown in Figure 4A. Bead−binder interaction of the same colour drives the folding of the chain. The SBS model for the *Nkx2-5* locus, i.e. the optimal number of different types of binding sites and their arrangement along the polymer, were inferred with the PRISMR learning algorithm (Bianco et al., 2018). Briefly, PRISMR uses as its only input a contact map of the locus of interest, such as Hi-C, and iteratively finds the polymer that minimizes the distance between the input and model contact map. Here, we employed Hi-C data from (Rosa-Garrido et al., 2017) at 5kb resolution to model a 200kb region around the *Nkx2-5* gene (mm10, chr17:26740000-26945000) in mouse CMs. The optimal polymer model inferred by our procedure was made up of 615 beads and included 15 different binding site types (Figure 4D).

Polymer model simulations: An ensemble of equilibrium 3D conformations of our polymer model was generated by massive parallel molecular dynamics (MD) simulations performed with the freely available LAMMPS software (Plimpton, 1995). The SBS polymer is initialized in a self-avoiding walk (SAW) conformation, with binders randomly placed in the surrounding environment. The simulation is performed in a cubic box with linear side as long as at least twice the gyration radius of the polymer SAW conformation. Periodic boundary conditions were implemented to reduce finite size effects. Starting from its

initial conformation, the system (i.e. beads and binders) evolves by Langevin dynamics, with standard interaction potentials described in classical polymer physics studies (Kremer and Grest, 1990) and widely employed in the field (Annunziatella et al., 2018).

The main system parameters are the binder concentration c and binding interaction energy E_{int}. The values for c and E_{int} were set to ensure the coil-globule transition (Chiariello et al., 2016), in order to achieve thermodynamically stable conformations. In particular, we employed c ~0.02% (expressed in volume fraction) and $E_{int} = 8.1 K_B T$, in line with previous studies (Barbieri et al., 2017; Chiariello et al., 2016). Starting from the SAW conformations, the system evolved up to 2×10^8 simulation time-steps, so to reach thermodynamic equilibrium where the polymer is completely folded. The complete ensemble of configurations consists of equilibrium structures obtained from 10^2 independent simulations. Figure 4C shows a typical model conformation of the *Nkx2-5* locus obtained from an MD simulation.

Contact maps and correlation metrics: From the generated ensemble of 3D polymers, we computed the average contact map for the SBS model (Figure 4B) as the probability of contact between each pair of polymer beads i and j, by considering them in contact if their Euclidean distance $r_{i,j}$ was below a threshold (Chiariello et al., 2016). To compare Hi-C and model contact maps of the *Nkx2-5* locus, we employed the Pearson correlation coefficient, r, and the distance-corrected correlation, r' (Bianco et al., 2018). r' is introduced to avoid the trivial contribution of the decaying trend of the contact probability as a function of the genomic distance. It is computed as the Pearson correlation coefficient between experimental and model matrices obtained by replacing each matrix element (i,j) with its relative variation with respect to the average contact probability at the genomic distance |i-j|.

Comparison of the model binding sites tracks with epigenetic marks: To investigate the biological nature of the model's inferred binding sites, we compared them with available ChIP-seq peaks of histone marks (Papait et al., 2013b) and CTCF (Rosa-Garrido et al., 2017). To this aim, we binned the ChIP-seq tracks at 5kb resolution and computed Pearson correlation coefficients between each binding site and epigenetic track pair (Bianco et al., 2018) (Figure 4D). Next, we tested the statistical significance of the obtained correlations by comparing them against a random control model. We built the control model as a distribution of correlations between ChIP-seq data and randomized binding sites tracks (10^3 independent realizations), obtained by bootstrapping the original binding sites positions (Bianco et al., 2018). We considered as significant the correlations falling below the 10th percentile (negative correlations) and above the 90th percentile (positive correlations) of the random control distribution. Resulting significant correlations are given in Figure 4D.

Modelling the impact of the eRNA on folding: To test *in silico* the impact of the eRNAs on 3D folding, we implemented the observed decrease in acetylation after *IRENE-SS* silencing. To this aim, the wild-type (wt) model binding sites which correlate with H3K27ac, i.e. types 7 (light blue) and 9 (purple) of Figure 4D were made inert (i.e. converted in non-interacting, grey beads) in the two 5kb windows containing the eRNAs and the enhancer (chr7:2674105000-2674115000). Next, the model contact matrix was re-computed and compared with the wt Hi-C by measuring distance-corrected correlation r' (silenced). To speed up computation, we used the mean-field approximation of the model contact matrices described in (Bianco et al., 2018). Notably, after silencing we found a significant correlation decrease, $\Delta r' = r'(wt)-r'(silenced)$, in the *Nkx2-5* TAD region (green area in the 3D colour scheme of Figure 4E), statistically higher than the correlation decrease found in a control model built by silencing in turn all other possible bins pairs (105 pairs in total) within the region, excluding bins containing CTCF sites (Figure 4E, empirical P-value $<1/105 \approx 0.01$). A subtraction map between the silenced and wt model contact matrices (silenced - wt) was also produced to highlight the contact changes around the silenced region (Figure 4F). In particular, the model predicts a reduction of the contact frequency between the gene and enhancer region (dashed box in Figure 4F) of ~30% with respect to the wild type.

Statistical Analysis
Unpaired t-test was used for two-group comparisons, and ordinary one-way ANOVA for multiple comparisons. Statistical analyses were performed with GraphPad Prism 8 software on at least two or three independent experiments, and two-sided P-value calculated (*, $P<0.05$; **, $P<0.01$; ***$P<0.001$). Data are represented as mean ± standard deviation (SD).

SUPPLEMENTAL REFERENCES

Abel, E. D., Kaulbach, H. C., Tian, R., Hopkins, J. C., Duffy, J., Doetschman, T., Minnemann, T., Boers, M. E., Hadro, E., Oberste-Berghaus, C., et al. (1999). Cardiac hypertrophy with preserved contractile function after selective deletion of GLUT4 from the heart. J Clin Invest 104, 1703-14.

Annunziatella, C., Chiariello, A. M. a. M., Esposito, A., Bianco, S., Fiorillo, L. and Nicodemi, M. (2018). Molecular Dynamics simulations of the Strings and Binders Switch model of chromatin. Methods 142, 81-88.

Barbieri, M., Chotalia, M., Fraser, J., Lavitas, L. M., Dostie, J., Pombo, A. and Nicodemi, M. (2012). Complexity of chromatin folding is captured by the strings and binders switch model. Proc Natl Acad Sci U S A 109, 16173-16178.

Barbieri, M., Xie, S. Q., Torlai Triglia, E., Chiariello, A. M., Bianco, S., De Santiago, I., Branco, M. R., Rueda, D., Nicodemi, M. and Pombo, A. (2017). Active and poised promoter states drive folding of the extended HoxB locus in mouse embryonic stem cells. Nat Struct Mol Biol 24, 515-524.

Chiariello, A. M., Annunziatella, C., Bianco, S., Esposito, A. and Nicodemi, M. (2016). Polymer physics of chromosome large-scale 3D organisation. Sci Rep 11, 559.

Claycomb, W. C., Lanson, N. A., Jr., Stallworth, B. S., Egeland, D. B., Delcarpio, J. B., Bahinski, A. and Izzo, N. J., Jr. (1998). HL-1 cells: a cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte. Proc Natl Acad Sci U S A 95, 2979-84.

Di Pasquale, E., Song, B. and Condorelli, G. (2013). Generation of human cardiomyocytes: a differentiation protocol from feeder-free human induced pluripotent stem cells. *J Vis Exp*. 2013/07/16 ed.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21.

Gu, Z. (2019). rGREAT: Client for GREAT Analysis.

Jawaid, W. (2019). enrichR: Provides an R Interface to 'Enrichr'. R package version 2.1 ed.

Kim, D., Paggi, J. M., Park, C., Bennett, C. and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37, 907-915.

Kolde, R. (2019). pheatmap: Pretty Heatmaps. R package version 1.0.12 ed.

Kremer, K. and Grest, G. S. (1990). Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. J Chem Phys 92, 5057-5086.

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 44, W90-7.

Larcher, V., Kunderfranco, P., Vacchiano, M., Carullo, P., Erreni, M., Salamon, I., Colombo, F. S., Lugli, E., Mazzola, M., Anselmo, A., et al. (2018). An autofluorescence-based method for the isolation of highly purified ventricular cardiomyocytes. Cardiovasc Res 114, 409-416.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-9.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T. and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33, 290-5.

Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. J Comput Phys 117, 1-19.

Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-2.

R Core Team (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Robinson, M. D., Mccarthy, D. J. and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-40.

Salvarani, N., Crasto, S., Miragoli, M., Bertero, A., Paulis, M., Kunderfranco, P., Serio, S., Forni, A., Lucarelli, C., Dal Ferro, M., et al. (2019). The K219T-Lamin mutation induces conduction defects through epigenetic inhibition

of SCN5A in human cardiac laminopathy. Nat Commun 10, 2267.

Wang, L., Wang, S. and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics 28, 2184-5.

Yamazaki, T., Souquere, S., Chujo, T., Kobelke, S., Chong, Y. S., Fox, A. H., Bond, C. S., Nakagawa, S., Pierron, G. and Hirose, T. (2018). Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. Mol Cell 70, 1038-1053 e7.