















DOT1L regulates chamber-specific transcriptional networks during cardiogenesis and mediates postnatal cell cycle withdrawal

Received: 2 November 2022

Accepted: 17 November 2022

Published online: 02 December 2022

 Check for updates


Paola Cattaneo ^{1,2,3,4,5} , Michael G. B. Hayes ⁶, Nina Baumgarten ^{3,4}, Dennis Hecker ^{3,4}, Sofia Peruzzo ^{3,4,5}, Galip S. Aslan^{3,4,5,7}, Paolo Kunderfranco ⁸, Veronica Larcher^{3,4,5}, Lunfeng Zhang¹, Riccardo Contu⁶, Gregory Fonseca⁹, Simone Spinozzi ⁶, Ju Chen⁶, Gianluigi Condorelli ^{2,8,10}, Stefanie Dimmeler ^{3,4,5}, Marcel H. Schulz ^{3,4,5}, Sven Heinz ⁶, Nuno Guimarães-Camboa^{1,3,4,5} & Sylvia M. Evans ^{1,6} 

Mechanisms by which specific histone modifications regulate distinct gene networks remain little understood. We investigated how H3K79me2, a modification catalyzed by DOT1L and previously considered a general transcriptional activation mark, regulates gene expression during cardiogenesis. Embryonic cardiomyocyte ablation of *Dot1l* revealed that H3K79me2 does not act as a general transcriptional activator, but rather regulates highly specific transcriptional networks at two critical cardiogenic junctures: embryonic cardiogenesis, where it was particularly important for left ventricle-specific genes, and postnatal cardiomyocyte cell cycle withdrawal, with *Dot1l* mutants having more mononuclear cardiomyocytes and prolonged cardiomyocyte cell cycle activity. Mechanistic analyses revealed that H3K79me2 in two distinct domains, gene bodies and regulatory elements, synergized to promote expression of genes activated by DOT1L. Surprisingly, H3K79me2 in specific regulatory elements also contributed to silencing genes usually not expressed in cardiomyocytes. These results reveal mechanisms by which DOT1L successively regulates left ventricle specification and cardiomyocyte cell cycle withdrawal.

Epigenetic enzymes play critical roles in organogenesis by defining chromatin structures required for cell type-specific transcriptional networks¹. Yet, mechanisms by which genome-wide histone modifications result in activation or repression of specific genes remain little understood. DOT1L is the only epigenetic enzyme catalyzing

methylation of lysine 79 of histone 3 (H3K79me)^{2,3}. Originally identified in yeast due to its role in maintenance of telomeric regions², DOT1L has been extensively studied in MLL-rearranged leukemia, where it is considered an emerging therapeutic target⁴. Initial genome-wide studies suggested that, rather than being associated with specific gene

¹Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 92093 La Jolla, CA, USA. ²Institute of Genetic and Biomedical Research (IRGB), Milan Unit, National Research Council of Italy, 20138 Milan, Italy. ³Institute of Cardiovascular Regeneration, Center of Molecular Medicine, Goethe University, 60590 Frankfurt am Main, Germany. ⁴German Center for Cardiovascular Research (DZHK), Partner Site Rhein Main, 60590 Frankfurt am Main, Germany. ⁵Cardiopulmonary Institute, Goethe University Frankfurt, 60590 Frankfurt am Main, Germany. ⁶Department of Medicine, University of California San Diego, 92093 La Jolla, CA, USA. ⁷Faculty of Biological Sciences, Goethe University, 60590 Frankfurt am Main, Germany. ⁸IRCCS Humanitas Research Hospital, 20089 Rozzano (MI), Italy. ⁹Department of Medicine, Meakins-Christie Laboratories, McGill University, H4A 3J1 Montreal, QC, Canada. ¹⁰Department of Biomedical Sciences, Humanitas University, 20090 Pieve Emanuele (MI), Italy.

 e-mail: paola.cattaneo@irgb.cnr.it; syevans@health.ucsd.edu

expression programs, H3K79me2/3 is ubiquitously found in all transcribed loci, leading to a model of H3K79me as a generic mark of active genes⁵.

Recent studies have revealed that, in addition to gene body H3K79me, DOT1L can regulate expression of its targets via methylation of regulatory regions⁶. We have previously shown that, in vitro, DOT1L is required for proper differentiation of embryonic stem cells into cardiomyocytes (CMs)⁷, however it is currently unclear whether this enzyme is necessary for normal cardiogenesis in vivo. In mice, global ablation of *Dot1l* results in embryonic lethality between E9.5 and E10.5 with mutants displaying yolk sac angiogenic defects, growth impairment and cardiac dilation⁸. Selective *Dot1l* ablation in CMs using *αMHC-Cre* results in an adult lethal phenotype with perturbations in dystrophin expression⁹. However, owing to the nature of this knockout (using a Cre that is not active during early embryogenesis^{10,11}), it is not known whether DOT1L plays any role in cardiogenesis. Furthermore, our molecular understanding of the functioning of this enzyme in cardiogenesis in vivo is limited by the absence of RNA-seq or ChIP-seq datasets that would allow for unbiased identification of direct targets of DOT1L in embryonic or neonatal CMs.

The mammalian heart develops from two major populations of cardiogenic progenitors – the first and second heart fields. Cells from the first heart field give rise to the left ventricle, whereas cells from the second heart field give rise to the right ventricle and most of the atria¹². A set of asymmetrically expressed transcription factors are essential for the establishment of chamber-specific transcriptional programs that regulate cardiac patterning and chamber maturation^{13–16}. Amongst these, *Irx4* (expressed only in ventricular CMs but absent from the atria) is essential to establish a ventricular identity in lieu of an atrial phenotype¹⁷. *Hand1* (expressed in the left ventricle throughout development and in the cardiac conduction system post-birth¹⁸) and *Hand2* (restricted to the right ventricle in early patterning and subsequently expressed in endocardium and myocardium of both ventricles) play a major role in defining systemic and pulmonary ventricular identity, respectively^{19–21}. Epigenetic mechanisms contributing to the tightly regulated expression of chamber-specific genes remain mostly unexplored, however, it is known that the H3K4 methyltransferase SMYD1/BOPI is a critical regulator of *Hand2* expression²². To date, it is not known whether any particular epigenetic enzyme plays a similar role in left ventricle-specific regulation of *Hand1* expression.

Regulated CM proliferation is a major component of normal cardiogenesis. The rate of CM mitosis gradually decreases from mid-gestational stages until birth, reaching a state of complete mitotic withdrawal in the first week after birth, when the majority of CMs become binucleated²³. Therefore, a major goal in cardiac regeneration is the identification of strategies to promote proliferation of adults CMs.

To test whether DOT1L plays a role in cardiogenesis, we ablated a floxed *Dot1l* allele using a Cre allele that is active in CMs from early developmental timepoints and observed a perinatal lethal phenotype with abnormal cardiac morphology. This phenotype resulted from perturbations in highly specific gene expression programs that orchestrate normal cardiogenesis at two distinct timepoints: embryonic heart development and neonatal CM cell cycle withdrawal. In embryogenesis, DOT1L emerged as a major regulator of the expression of several transcription factors involved in chamber-specific gene expression profiles, with left ventricle-specific genes being particularly sensitive to the absence of this enzyme. In the neonatal period, DOT1L promoted expression of genes involved in CM maturation and cell cycle exit, with *Dot1l* knockouts exhibiting sustained CM proliferation, a feature of interest for cardiac regeneration. Integration of CM ChIP-seq and Hi-C data revealed that DOT1L regulated target genes via two mechanisms: methylation of H3K79 in gene bodies and methylation of H3K79 in regulatory elements (K79-REs). Our analyses identified two types of K79-REs: activating elements that

synergized with gene body H3K79me2 to promote expression of targets activated by DOT1L, and K79-REs that contributed to silencing of genes normally not expressed in CMs. In contrast to H3K79me2 being a general activator of gene transcription as previously thought, our results reveal H3K79me2 as an epigenetic mark that regulates highly specific gene programs by both activating and repressing target genes.

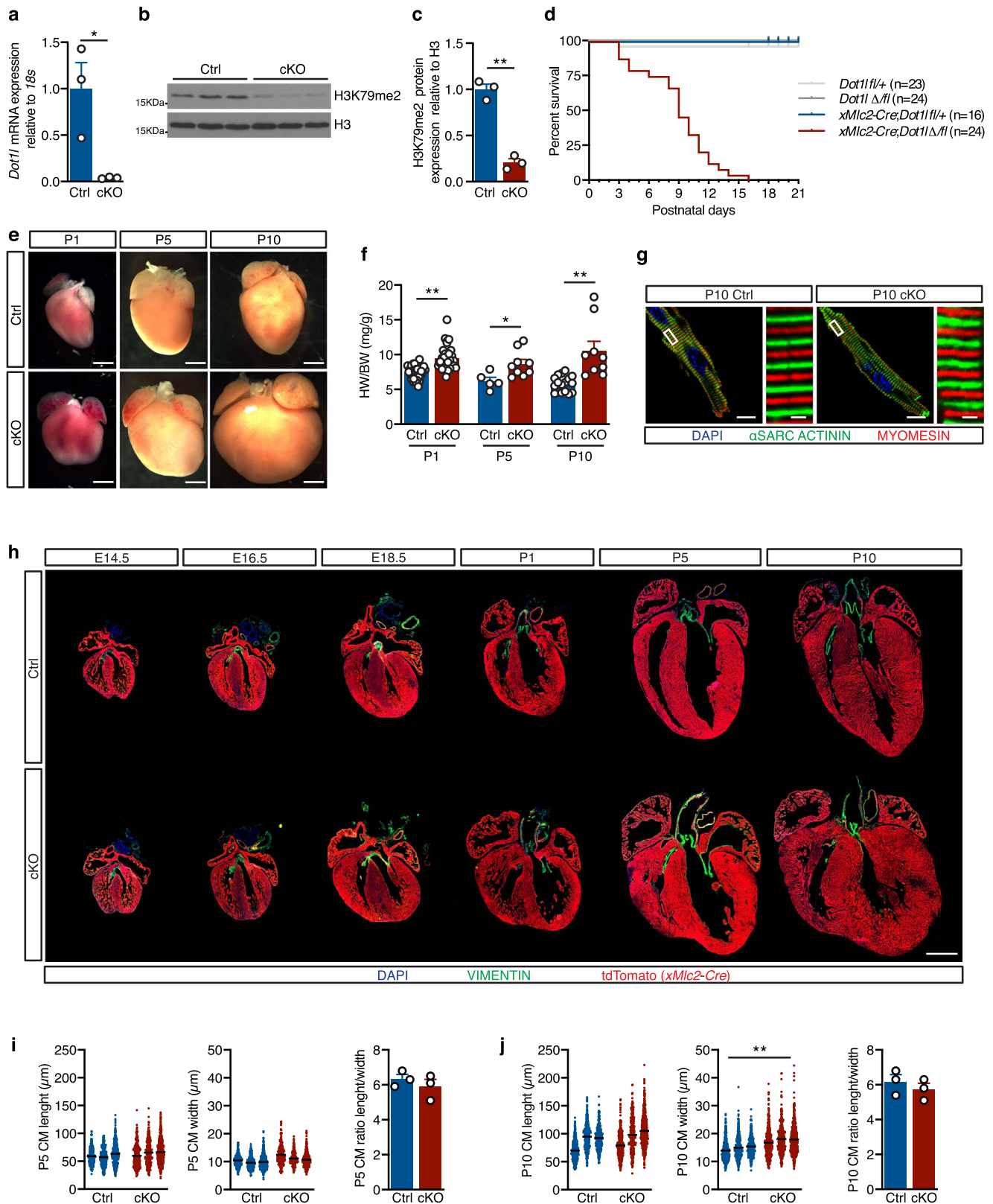
Results

Abnormal cardiogenesis in DOT1L cKOs

Following our previous observation that DOT1L is required for the differentiation of embryonic stem cells into CMs⁷, we decided to assess if this enzyme is required for normal cardiogenesis in vivo. Fluorescent RNA in situ hybridization studies showed that *Dot1l* is robustly expressed in CMs at least from E10.5 onwards (Supplementary Fig. 1a). This expression was observed throughout the heart, without being specific to any chamber. *Dot1l* expression was also observed in non-cardiomyocyte cells of the heart, as well as most tissues outside the heart, showing this enzyme is not cardiac-specific. However, at E16.5, *Dot1l* expression levels were higher in CMs than in the developing valves (Supplementary Fig. 1b). *Dot1l* has been previously ablated in the heart using an *αMHC-Cre* line, leading to an adult lethal phenotype with reduced dystrophin expression⁹. We have shown that *αMHC-Cre* is not optimal for myocardial-restricted knockouts due to its expression in non-myocyte lineages inside and outside the heart, and because it does not flox-out in all CMs at an early time point^{10,11}. To avoid these limitations and study CM-specific roles of DOT1L from early embryonic timepoints, we ablated a floxed *Dot1l* allele (loxP sites flanking exon 2) using the *xMlc2-Cre* allele²⁴ that drives highly specific and efficient recombination in CMs (Supplementary Fig. 1c) from the cardiac crescent stage²⁴. In our analyses, we compared *xMlc2-Cre+; Dot1l-WT/flox* mice (herein designated as *Dot1l* Ctrl or controls) with *xMlc2-Cre+; Dot1l-Δ/flox* littermate mice (herein designated as *Dot1l* cKO or mutants). Inclusion of a copy of the *Rosa26-tdTomato* reporter allele²⁵ in all crosses allowed highly specific labeling of CMs by the red fluorescent protein tdTomato for downstream flow cytometry and confocal microscopy applications (Supplementary Fig. 1d). Real time qPCR analyses revealed highly efficient ablation of the floxed allele in FACS-sorted E12.5 CMs (Fig. 1a).

DOT1L is the only enzyme catalyzing H3K79 methylation³. Consequently, CM ablation of *Dot1l* resulted in a strong reduction in H3K79me2 levels in E14.5 hearts (Fig. 1b, c). CM-specific ablation of *Dot1l* did not result in an embryonic lethal phenotype, as animals with the mutant genotype were observed at expected Mendelian ratios in all stages analyzed (Supplementary Fig. 1e). Despite being born at expected numbers, *Dot1l* cKO mice started dying shortly after birth, with 50% mortality by postnatal day 9 (P9), and the longest-lived mutant surviving until P16 (Fig. 1d). Macroscopic examination of organs at neonatal stages revealed that mutants had enlarged hearts with a rounded shape relative to controls (Fig. 1e). This phenotype was exacerbated as animals aged and peaked at P10 in surviving animals (Fig. 1e). This evident cardiac enlargement translated into increased heart weight (Supplementary Fig. 1f) and increased heart weight/body weight ratios (Fig. 1f) without changes in body weight (Supplementary Fig. 1g). Co-immunostaining for α Sarcomeric Actinin and Myomesin revealed absence of myofiber orientation defects in CMs from *Dot1l* cKO hearts (Fig. 1g).

To determine the timing of onset of the cardiac phenotype and assess for additional morphogenic malformations, we conducted a histological time course analysis spanning from mid-gestation to postnatal timepoints (Fig. 1h). cKO hearts were completely indistinguishable from control littermates until E14.5 and started exhibiting the enlarged, rounded phenotype between E16.5 and E18.5. Postnatally, mutant hearts displayed ventricular walls with increased thickness and areas of moderate persistent trabeculation (Fig. 1h). No additional major morphogenic abnormalities (septal or



outflow defects) were observed. Immunostaining for Vimentin revealed absence of valve malformations and absence of major foci of fibrosis in cKOs (Fig. 1h). Absence of fibrotic remodeling was also confirmed by qPCR for fibrotic markers *Col1a1* and *Col3a1* (Supplementary Fig. 2a) performed on whole heart tissue, immunostaining for Collagen1 (Supplementary Fig. 2b), and Masson trichrome staining (Supplementary Fig. 2c). To assess if the cardiac

enlargement observed in cKOs is a consequence of CM hypertrophy, we conducted length and width measurements of CMs isolated from *Dot1l* cKO and control hearts. These analyses revealed that, at P5, CMs from both genotypes had comparable sizes (Fig. 1i). At P10, CM length was similar in both groups, but CM width was moderately increased in cKOs, without causing a significant change in the CM length-to-width ratio (Fig. 1j). These results ruled

Fig. 1 | Cardiomyocyte-specific ablation of DOTIL from early developmental timepoints results in enlarged hearts and peri-natal lethality. **a** qPCR analysis using a primer within the floxed exon of *Dot1l* mRNA showing efficient ablation of this gene in E12.5 cKO FACS-sorted CMs ($N = 3$ biological replicates, unpaired t -test, two-sided $P = 0.0271$). **b** Western blot (**b**) and respective quantification (**c**) showing strongly reduced H3K79me2 levels in E14.5 hearts upon ablation of *Dot1l* ($N = 3$ biological replicates, unpaired t -test, two-sided $P = 0.0004$). **d** Kaplan–Meier survival curves showing postnatal lethality of *Dot1l* cKOs. **e** Whole mount images of postnatal day (P) 1, P5 and P10 hearts in Ctrl and *Dot1l* cKOs representative of the enlarged heart phenotype of *Dot1l* cKOs (scale bars = 1 mm). **f** Graph representing a significant increase in heart weight/body weight ratio (HW/BW (mg/g)) in *Dot1l* cKO vs Ctrl in all stages analyzed. (P1 Ctrl $N = 36$, P1 cKO $N = 30$, unpaired t -test, two-sided $P < 0.0001$; P5 Ctrl $N = 5$, P5 cKO $N = 9$, unpaired t -test, two-sided $P = 0.0223$; P10 Ctrl $N = 16$, P10 cKO $N = 9$ biological replicates, unpaired t -test, two-sided

$P = 0.0002$). **g** Confocal images showing no alterations in sarcomere organization and myofiber orientation in *Dot1l* cKOs. α Sarcomeric Actinin in green, Myomesin in red, DAPI (4',6-diamidin-2-phenylindol) in blue (scale bar = 10 μ m for lower magnification images on the left and 1 μ m for higher magnification images on the right). **h** Immunofluorescence time course depicting the dynamics of phenotypic manifestations in *Dot1l* cKOs in embryonic (E) and postnatal (P) stages. DAPI in blue, Vimentin in green and lineage traced *XMlc2-Cre*;tdTomato CMs in red (scale bar = 1 mm). Assessment of CM length (left), width (middle) and ratio of CM length/width (right) at P5 (**i**) and P10 (**j**) indicated no major changes in CM size. Measurements were performed on isolated CMs (mean of 413 CMs counted per heart from $N = 3$ biological replicates, unpaired t -test, two-sided $P = 0.0096$ for P10 CM width). In all graphs Ctrl indicates control mice (*XMlc2-Cre;Dot1l fl/+*), cKO indicates mutant mice (*XMlc2-Cre;Dot1l Δ/fl*). Data is presented as mean \pm SEM; * represents $P \leq 0.05$, ** $P \leq 0.01$. Source data are provided as a Source Data file.

out CM hypertrophy as a cause of the observed cardiac enlargement.

Echocardiographic assessment of heart function revealed that, compared with littermate controls, *Dot1l* cKO mice had reduced fractional shortening both at P5 and P10, increased left ventricular posterior wall thickness and increased left ventricular chamber diameter (Fig. 2a, b). Electrocardiographic analyses revealed that *Dot1l* cKO mutants had reduced heart rate both at P5 and P10, which translated into prolonged and irregular R-R intervals, with occasional beat drops (Fig. 2c–e). At P5, *Dot1l* cKOs also had prolonged QRS and P-R intervals, whereas at P10 the P-R intervals of mutants were comparable to those of controls, suggesting that the mutants with more severe conduction defects might die between P5 and P10. Despite the strong cardiac phenotype, *Dot1l* cKOs presented no obvious signal of distress prior to death. Therefore, it is possible that defects in cardiac conduction might account for the sudden lethal phenotype.

Altogether, these findings identify a critical role for DOTIL in normal cardiogenesis and suggest that the strong macroscopic and functional phenotypes observed in *Dot1l* cKOs are not a consequence of fibrotic remodeling or major alterations in CM structure or size.

DOTIL is essential for chamber-specific gene expression

To decipher gene expression networks misregulated in the absence of DOTIL, we used RNA-seq to assess the transcriptomes of FACS-sorted Ctrl and cKO CMs just prior to the onset of clear phenotypic changes (E16.5). An average of 100,000 CMs were sorted per E16.5 heart (including CMs from all four cardiac chambers), yielding enough RNA to prepare biological replicates from individual hearts (3 replicates per genotype group). Bioinformatics analyses revealed that, from a total of 12,110 expressed genes (RPKM ≥ 1 in either Ctrl or cKO), 1478 were misregulated in cKO versus Ctrl CMs ($\log_2FC \leq -0.5$ or ≥ 0.5 , false discovery rate ≤ 0.05). From these, 439 genes were downregulated and 1039 upregulated in cKOs (Fig. 3a and Supplementary Data 1a). Transcripts upregulated in cKOs corresponded to genes that are normally expressed at low levels in control CMs (62% of genes in the lower expression quartile), whereas downregulated genes corresponded to genes normally expressed at a medium/high level (Fig. 3b). Notably, the gene displaying the highest downregulation ($-5.4 \log_2FC$) in mutant CMs was the transcription factor *Hand1*, a central element in the establishment and maintenance of left ventricular identity (Fig. 3c and Supplementary Data 1a). In addition, amongst genes significantly downregulated in mutants there were several other genes asymmetrically expressed across the cardiac chambers: the transcription factors *Irx4* (ventricular specific¹⁷), *Tbx5* (left ventricle and atrial-specific²⁶), and the left ventricular-specific genes *Gja5* (encoding Connexin40) and *Cited1* (Fig. 3c and Supplementary Data 1a). These results revealed that transcriptional networks operating in left ventricular CMs (those derived from the first heart field) are particularly sensitive to the absence of DOTIL. In addition to chamber-specific genes, DOTIL also regulated critical cardiogenic factors that are

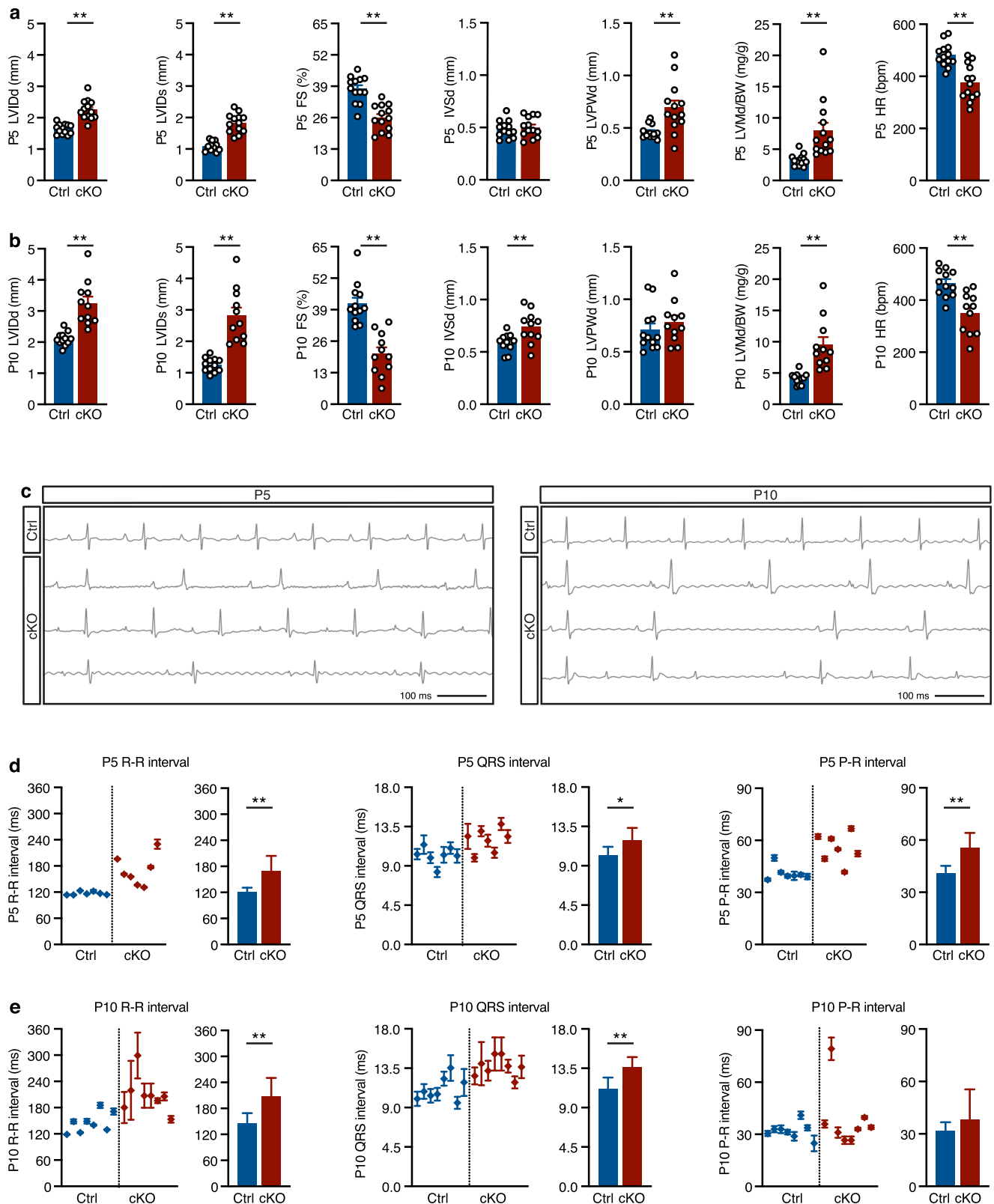
expressed throughout the heart. Examples of these are the transcription factors *Gata4* and *Mef2c*, as well as the epigenetic enzyme *Smyd1*. The transcription factor *Nkx2-5* also exhibited a trend toward downregulation (Supplementary Data 1a). These transcriptomic differences revealed that DOTIL activity in embryonic CMs is essential for establishment of gene expression networks that coordinate normal cardiogenesis. This requirement, however, was not generalized to all cardiac patterning genes, as transcript levels for *Hand2*, *Tbx20* and *Nppa* (ANF) were not altered in mutants (Supplementary Data 1a).

To validate RNA-seq results, we performed fluorescent RNA in situ hybridization (RNA-scope) in histological sections of control and cKO hearts. Strikingly, at E16.5, *Hand1* transcripts were completely undetectable in the left ventricle of mutants (Fig. 3d, bottom panel, quantification in Fig. 3e, right). Downregulation of *Hand1* was already evident at E10.5 when levels of *Hand1* expression in control hearts were higher than at E16.5 (Fig. 3d, top and middle panels, quantification in Fig. 3e, left), indicating that transcriptional consequences of DOTIL ablation preceded the first phenotypic manifestations by several days. Transcript levels for the ventricular-specific transcription factor *Irx4* and for the CM-specific epigenetic modifier *Smyd1* were significantly reduced in cKO hearts both at E10.5 and E16.5 (Fig. 3g and Supplementary Fig. 3, quantification in Fig. 3h, i), confirming RNA-seq results. RNA-scope analyses also showed that *Dot1l* cKO CMs expressed normal levels of *Hand2*, both at E10.5 and E16.5, further validating that DOTIL regulates specific transcriptional programs, rather than being a generic activator of transcription (Supplementary Fig. 4, quantification in Fig. 3f).

Our data revealed a role for DOTIL in regulating the expression of genes necessary for the establishment and maintenance of chamber identity, in particular those involved in left ventricular identity. While a number of genes were downregulated in cKOs, *Hand1* was particularly sensitive to absence of DOTIL, with *Hand1* transcripts being completely absent from midgestational DOTIL-deficient CMs (both in RNA-seq and RNA-scope). Consistent with these transcriptomic observations, cardiac-specific ablation of *Hand1* phenocopies *Dot1l* cKO, including an enlarged and round-shaped heart and peri-natal lethality²¹.

Transcriptional control via gene body H3K79me2

Several core cardiac transcription factors are differentially expressed in *Dot1l* cKO CMs. Thus, genes modulated in our transcriptomic analyses likely include a combination of targets directly regulated by DOTIL as well as those indirectly regulated owing to secondary effects. To identify genes directly regulated by DOTIL, we performed ChIP-seq assays for H3K79me2 in E16.5 control and cKO FACS-sorted CMs. These analyses revealed a total of 31,895 H3K79me2 peaks significantly enriched in control versus *Dot1l* cKO CMs (Supplementary Data 2a). The fact that the vast majority of H3K79me2 peaks in Ctrl were lost in cKOs is consistent with DOTIL being the sole histone K79 methyltransferase and validated the purity of our sorted CMs (Fig. 4a and



Supplementary Data 2a). Genome-wide differential peak distribution analysis revealed that H3K79me2 is mostly an intragenic modification (introns, exons, UTR and promoter/TSS) with only 1% of differential peaks being located in intergenic regions (Fig. 4b).

Genes downregulated in cKOs exhibited, in control CMs, maximal levels of H3K79me2 immediately after the transcription start site (TSS) that progressively decreased towards the transcription termination

site (TTS) (Fig. 4c). Gene body H3K79me2 (coverage ≥ 50 reads and fraction of gene body ≥ 0.20 in control CMs) was present in 341 downregulated genes (77% of all downregulated genes) and 98 upregulated genes (9% of all upregulated genes) (Fig. 4d, e and Supplementary Data 1a). This observation suggests that a significant part of gene downregulation observed in cKOs can be directly attributed to H3K79me2 in the gene body, whereas gene upregulation does not

Fig. 2 | Echocardiographic and electrocardiographic defects in Dot1L cKO hearts. Echocardiographic analyses conducted at P5 (a) and P10 (b) revealed significant defects in Dot1L cKO hearts at both timepoints, including increased left ventricular inner diameter, both in diastole (LVIDd) and systole (LVIDs), reduced fractional shortening (FS), and increased diastolic left ventricle mass to body weight ratios (LVmD/BW). Dot1L cKOs also exhibited increased diastolic left ventricular posterior wall thickness (LVPWd) at P5 and increased diastolic intraventricular septum thickness (IVSd) at P10. Additionally, Dot1L cKOs displayed reduced heart rate (HR) at both timepoints. (P5 Ctrl $N=13$, P5 cKO $N=13$, P10 Ctrl $N=12$, P10 cKO $N=11$ biological replicates; unpaired t -test, two-sided P5 LVIDd $P<0.0001$, P5 LVIDs $P<0.0001$, P5 FS $P<0.0001$, P5 LVPWd $P=0.0029$, P5 LVmD/BW $P=0.0017$, P5 HR $P<0.0001$; P10 LVIDd $P<0.0001$, P10 LVIDs $P<0.0001$, P10

FS $P<0.0001$, P10 IVSd $P=0.0083$, P10 LVmD/BW $P=0.0002$, P10 HR $P<0.0005$). c Representative ECG tracks from Ctrl and Dot1L cKO. ECG measurements revealed multiple defects in Dot1L cKOs both at P5 (d) and P10 (e), including increased and irregular R-R and QRS intervals. At P5 Dot1L cKOs also displayed significantly increased P-R intervals, however, this difference could not be detected in P10 mutants. (P5 $N=7$, P10 $N=8$ biological replicates, unpaired t -test, two-sided P5 R-R $P=0.004$, P5 QRS $P=0.0164$, P5 P-R $P=0.0017$, P10 R-R $P=0.0024$, P10 QRS $P=0.001$, P10 P-R $P=0.2279$). The left part of the graph represents the mean \pm SD of multiple beats measured for each mouse. The right graph represents the mean \pm SD of multiple biological replicates. Data is presented as mean \pm SD; * represents $P\leq 0.05$, ** $P\leq 0.01$. Source data are provided as a Source Data file.

seem to be directly associated with gene body H3K79me2. This is consistent with the view of gene body H3K79me2 as an activation mark⁵, however, it should be noted that this modification is not an absolute requirement for transcription, as 42% of all genes expressed in control CMs did not have significant levels of gene body H3K79me2 (Fig. 4f). Within genes showing gene body H3K79me2, there were two clear categories: genes dependent on DOT1L (modulated in cKOs) and genes not significantly affected by the loss of DOT1L, suggesting a scenario in which other epigenetic modifications can compensate for the loss of gene body H3K79me2 in a subset of genes. Notably, the majority of genes whose transcription depended on gene body H3K79me2 belonged to the highest quartiles of expression, suggesting this modification might play a role in supporting high transcriptional levels (Fig. 4f).

Transcriptional control via H3K79me2 in regulatory elements

Recent studies in leukemia showed that there is a distinct subset of enhancers that depend on H3K79me2⁶. We conducted analyses to investigate whether CMs have H3K79me2-dependent regulatory elements (REs) that, together with gene body H3K79me2, contribute to gene expression differences observed in cKOs. To identify regulatory regions, we generated CM-specific H3K27ac ChIP-seq datasets. At E16.5, control CMs had 52,495 H3K27ac peaks, mainly occupying introns, promoters/TSS or intergenic regions (Supplementary Fig. 5a). From these, 37% overlapped with H3K79me2 peaks (Fig. 4g and Supplementary Fig. 5b). To predict interactions between candidate regulatory elements and target genes, we applied an adapted implementation of the Activity-by-Contact (ABC) model^{27,28}. By integrating information on chromatin state (H3K27ac ChIP-seq) and genomic conformation (Hi-C), this method is more accurate at predicting actual interactions between REs and their target genes than methods simply assigning H3K27ac peaks to the closest neighboring gene²⁷. Because the promoter of a gene can serve as a RE for a distinct cis gene^{29,30}, our analyses also considered interactions between promoters and distal genes. Using the ABC algorithm, our CM-specific H3K27ac ChIP-seq dataset was combined with a publicly available CM Hi-C dataset³¹ (GSM2544836). To ensure we focused on meaningful candidate RE-target gene interactions, we only considered the most relevant interactions identified by the ABC method (adapted ABC-score ≥ 0.02): 373,610 interactions, corresponding to 47,547 H3K27ac peaks (Supplementary Data 3a and Supplementary Fig. 5c). From these, 36% (16,963 peaks) overlapped with a differential H3K79me2 peak (Supplementary Data 3a and Supplementary Fig. 5c). This large number of CM REs with H3K79me2 signal (from here on designated as K79-REs) could provide an alternative mechanism for gene expression regulation by DOT1L. Of note, about 34% of all K79-REs observed at E16.5 overlapped a promoter (Supplementary Data 3a and Supplementary Fig. 5c).

To assess the relative contribution of H3K79me2 in gene body versus H3K79me2 in REs to the regulation of target genes, we quantified the effect of these variables in the cumulative distribution of log₂FC values. These analyses revealed that genes with higher fraction

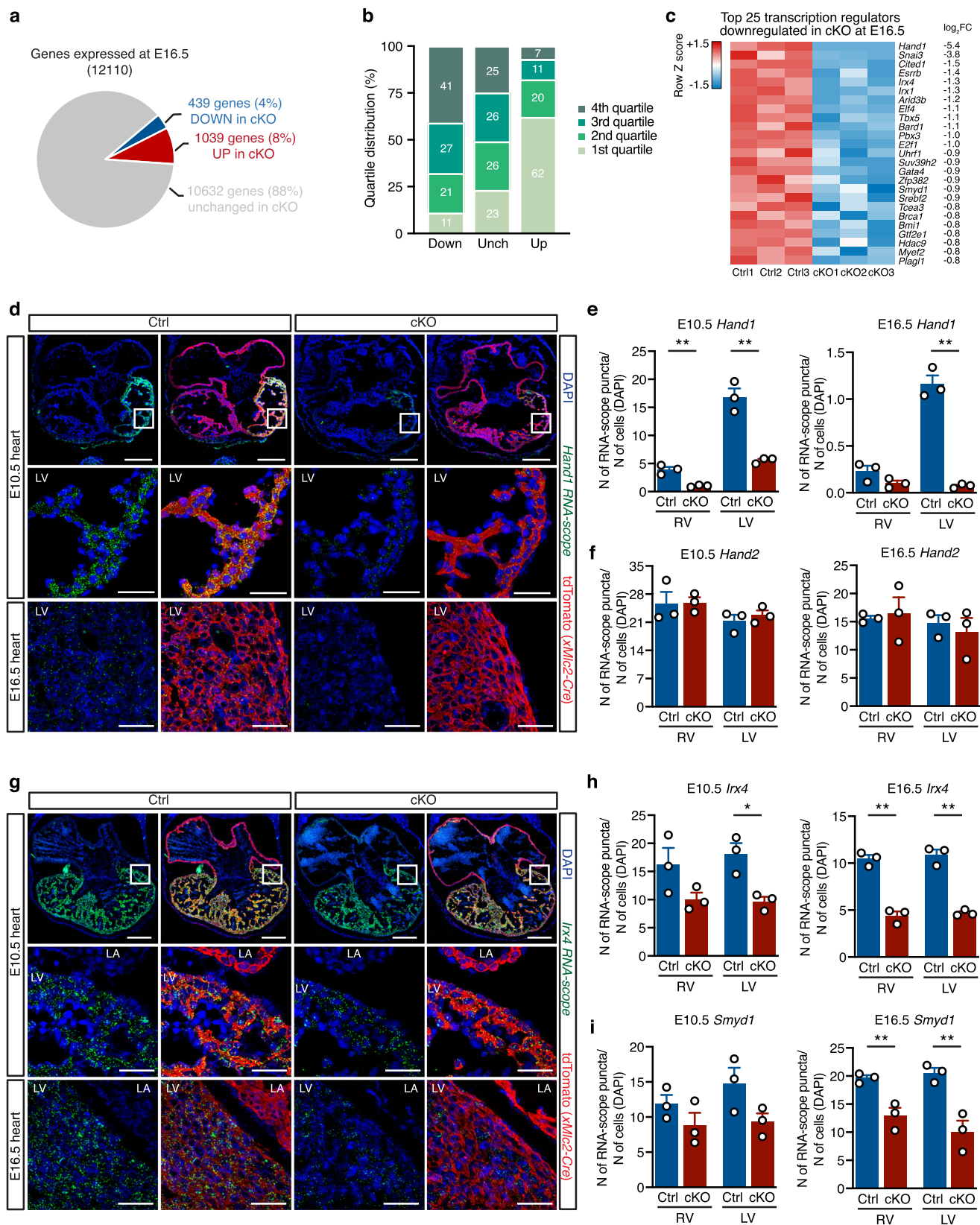
of H3K79me2 in the gene body were more downregulated in cKOs than genes with lower (less than 33%) fraction of H3K79me2 in the gene body (Supplementary Fig. 5d and Supplementary Data 4a). Genes with 6 or more K79-REs displayed stronger modulation than those associated with lower numbers of REs (Supplementary Fig. 5e and Supplementary Data 4a). Focusing exclusively on loci with gene body H3K79me2, those that interacted with K79-REs were more downregulated in mutant CMs than those without these elements, indicating that gene body and regulatory element H3K79me2 synergized to potentiate expression of target genes (Supplementary Fig. 5f and Supplementary Data 4a).

The vast majority (74.5%) of the genes downregulated in DOT1L-deficient CMs were associated with both gene body H3K79me2 and at least one K79-RE (Fig. 4h). Interestingly, 17.6% of all downregulated genes were associated with K79-REs in the absence of gene body H3K79me2. Overall, only 5.5% of downregulated genes were not associated with H3K79me2 in gene body or K79-REs, indicating most gene downregulation was a direct consequence of Dot1L cKO. On the other side of the scale, upregulated genes had reduced association with gene body H3K79me2 (as mentioned above), but, surprisingly, 65.4% of upregulated genes were associated with K79-REs in the absence of gene body H3K79me2. This observation suggests a role for DOT1L in gene silencing via H3K79me2-dependent regulatory elements. Functional annotation revealed that genes downregulated in cKOs with H3K79me2 in gene bodies and REs were involved in differentiation of striated muscle and cardiac morphogenesis (Fig. 4i and Supplementary Data 5a). On the other hand, genes upregulated in cKOs and associated with inhibitory K79-REs were related to non-muscle functions: extracellular matrix organization and skeletal system development (Fig. 4j and Supplementary Data 5b).

To search for cues as to how DOT1L achieves gene silencing via methylation of REs, we screened for enrichment in binding sites for known transcriptional regulators in the K79-REs of genes upregulated without gene body H3K79me2 but with K79-REs versus genes downregulated with gene body H3K79me2 and K79-REs. This analysis revealed that K79-REs associated with genes upregulated in cKOs (silencing K79-REs) were enriched in binding sites for MBD2, a member of the NuRD complex known to play an important role in transcriptional silencing³² (Fig. 4k). Enrichment in NuRD complex binding sites in the silencing K79-REs could provide an explanation for how DOT1L promotes silencing of target genes. However, it is also possible that upregulated genes reflect secondary effects (for example downregulation of a transcription regulator with repressor properties) rather than being a direct consequence of DOT1L inactivation.

H3K27ac relationship to H3K79me2

To assess whether H3K27ac in K79-REs is altered upon DOT1L ablation, we also performed H3K27ac ChIP-seq in E16.5 CMs sorted from cKO hearts. Comparison of ChIP-seq profiles from control and mutant CMs identified 2,354 differential H3K27ac peaks (Fig. 5a, b and Supplementary Data 6a). From these, 2,017 were upregulated and 337 downregulated in cKOs (Fig. 5a and Supplementary Data 6a). Most



(69%) of the H3K27ac peaks downregulated in cKOs overlapped with a differential H3K79me2 peak, whereas the majority (99%) of upregulated H3K27ac peaks did not (Fig. 5b, c). From the 16,963 K79-REs identified by the ABC method, the vast majority (98%) did not show any difference in H3K27ac enrichment in cKOs, indicating that, at a genome-wide scale, H3K79me2 is not necessary for maintenance of

H3K27ac in REs positive for both modifications (Supplementary Data 3a). Despite corresponding to a small fraction (2%) of all K79-REs, those with differential H3K27ac made a relevant contribution to differential gene expression in DotLL cKOs. From the 313 genes downregulated with H3K79me2 in gene body and K79-REs, 60 (19%) also had differential H3K27ac in at least one RE (Fig. 5d and Supplementary

Fig. 3 | DOT1L is required in cardiomyocytes for chamber-specific gene expression. **a** Pie chart representing the number of genes down- ($\log_2FC \leq -0.5$; $FDR \leq 0.05$) and up-regulated ($\log_2FC \geq 0.5$; $FDR \leq 0.05$) in Dot1L cKO CMs at E16.5. **b** Quartile distribution of gene expression in CMs at E16.5. Genes downregulated (Down) in Dot1L cKO were expressed at a high level in control CMs, whereas the majority of upregulated genes (Up) belonged to the bottom quartile of expression. Genes not significantly modulated (Unch) were evenly distributed across quartiles of expression. Data are shown as stacked percentage bar graph. **c** Heatmap showing the expression of the top 25 transcription regulators downregulated in Dot1L cKO CMs at E16.5, highlighting that multiple chamber-specific transcription regulators were significantly downregulated. RNA-scope analyses (**d**) and respective quantification (**e**) validating blunted expression of *Hand1* in Dot1L cKOs both at E10.5 and E16.5. ($N = 3$ biological replicates; unpaired t -test, two-sided, E10.5 RV $P = 0.0042$,

E10.5 LV $P = 0.0020$, E16.5 LV $P = 0.0003$). **f** Quantification of RNA-scope analysis validating no changes in expression of *Hand2* in Dot1L cKOs both at E10.5 and E16.5 (RNA-scope images presented in Supplementary Fig. 4) ($N = 3$ biological replicates). RNA-scope analyses (**g**) and respective quantification (**h**) validating reduced levels of *Irx4* in Dot1L cKOs both at E10.5 and E16.5. ($N = 3$ biological replicates; unpaired t -test, two-sided, E10.5 LV $P = 0.0153$, E16.5 RV $P = 0.0006$, E16.5 LV $P = 0.0005$). **i** Quantification of RNA-scope analysis validating reduced levels of *Smyd1* in Dot1L cKOs both at E10.5 and E16.5 (RNA-scope images presented in Supplementary Fig. 3). ($N = 3$ biological replicates; unpaired t -test, two-sided, E16.5 RV $P = 0.0057$, E16.5 LV $P = 0.008$). For panels **d** and **g** scale bars = 250 μ m for low magnification panels and 50 μ m for high magnification images. In panels **e**, **f**, **h** and **i** data is presented as mean \pm SEM; * represents $P \leq 0.05$, ** $P \leq 0.01$. ($N = 3$ biological replicates). Source data are provided as a Source Data file.

Data 4a). From the 653 upregulated genes without gene body H3K79me2 but with K79-REs, 39% also had differential H3K27ac in at least one RE (Fig. 5e and Supplementary Data 4a). As expected, for K79-REs associated with upregulated genes, differential H3K27ac peaks were mainly upregulated, whereas for K79-REs associated with downregulated genes, differential H3K27ac peaks were mainly downregulated (Supplementary Data 4a). Altogether, these results revealed that, in mutants, differential H3K27ac within a subset of K79-REs may contribute to altered gene regulation by those REs, however, differential H3K27ac is not a prerequisite for gene regulation by K79-REs.

In summary, our mechanistic data revealed that, during embryogenesis, DOT1L directly regulates critical cardiogenic transcription factors: *Hand1*, *Irx4*, *Gata4*, *Tbx5*, and *Mef2c*, all of which had gene body H3K79me2 combined with K79-REs (Fig. 5g and Supplementary Data 4a). Both *Hand1* and *Irx4* were associated with at least one K79-RE with downregulated H3K27ac in Dot1L cKOs (Fig. 5f, Supplementary Fig. 5g and Supplementary Data 4a). Interestingly, *Smyd1*, encoding the H3K4 methyltransferase necessary for right ventricular expression of *Hand2*²², was also a direct target of DOT1L, unveiling complex epigenetic mechanisms for transcriptional regulation in cardiogenesis. Our results also revealed that DOT1L has particular importance in maintenance of left ventricle-specific transcriptional networks, as it directly regulated not only transcription factors specifying the identity of this chamber (*Hand1* and *Tbx5*), but also their direct downstream targets *Cited1* and *Gja5* (Cx40)²¹ (Fig. 5g). On the other hand, *Nppa* (ANF), another classic target of both HAND1²¹ and TBX5³³ was not altered in Dot1L cKOs, further highlighting the specificity and complexity of this epigenetic control (Fig. 5g). These findings reveal previously unrecognized roles for DOT1L and improve our understanding of the complex transcriptional and epigenetic landscape governing mammalian cardiogenesis.

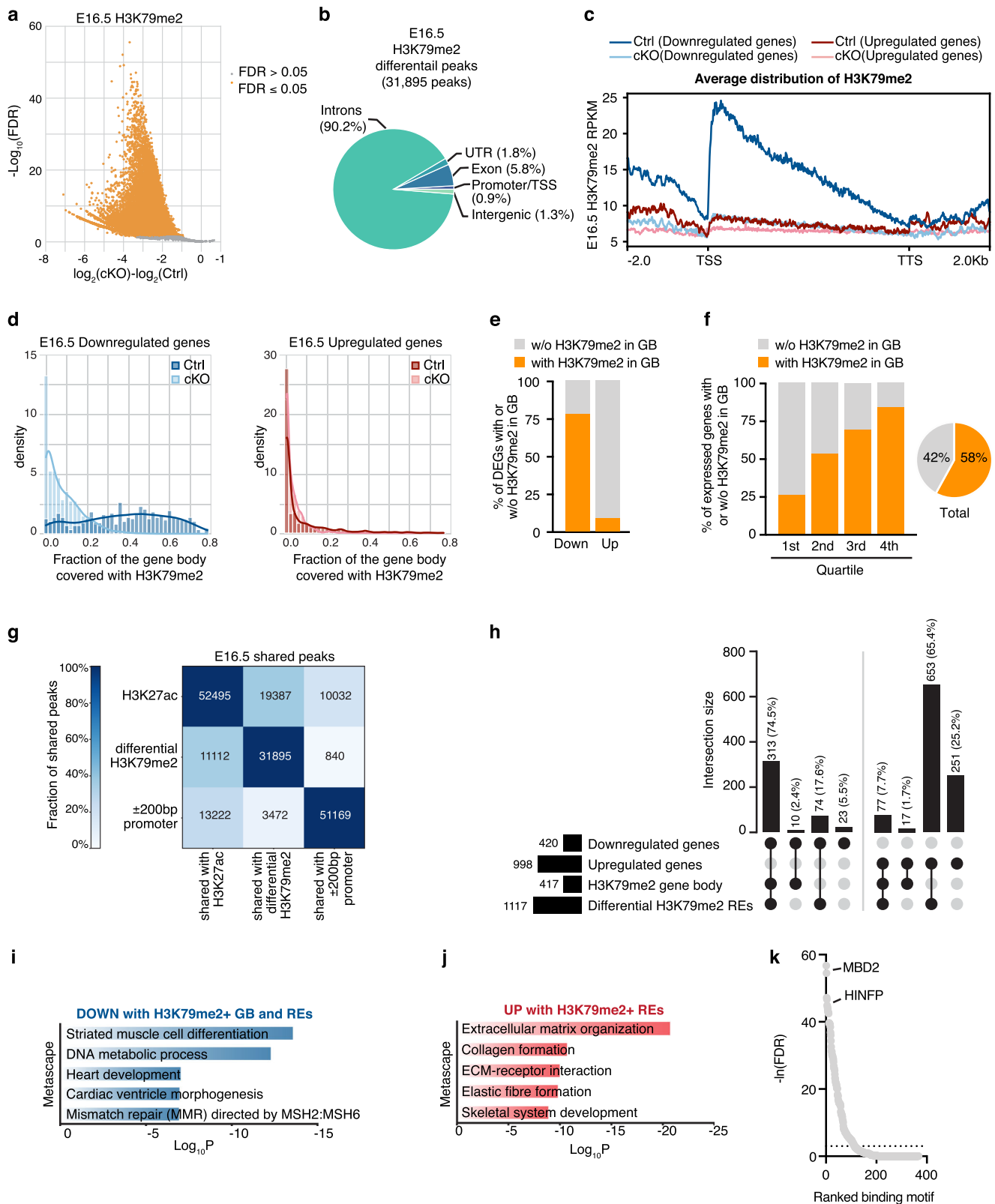
Impaired cardiomyocyte cell cycle withdrawal in Dot1L mutants

Defective expression of core cardiogenic factors (including chamber-specific genes) can account for the abnormal morphology of Dot1L cKO hearts, however, it does not explain the increased wall thickness observed in postnatal mutant hearts (Fig. 1h). Dot1L cKO and Ctrl CMs had comparable sizes (Fig. 1i, j), therefore increased wall thickness in mutants was not caused by CM hypertrophy. As increased myocardial wall thickening coincided with the developmental period in which CMs withdraw from cell cycle²³, we hypothesized this process might be affected in Dot1L cKOs. Flow cytometry analyses of cells isolated from P1 hearts revealed that, at this stage, about 50% of all cardiac cells were CMs (labeled by expression of the reporter gene tdTomato, Fig. 6a). Assessment of EdU incorporation rates (24 h EdU pulse) revealed that, in control hearts, 19% of all CMs were proliferative, whereas this number increased to 26% in mutant hearts (Fig. 6a), representing a statistically significant increase in CM proliferation (Fig. 6b). In mice, by P10 most CMs are withdrawn from cell cycle and 80% of CMs are binucleated²³. Consistently, fluorescent microscopy analyses of CMs isolated from P10 hearts revealed that less than 3% of control CMs were

EdU+(24 h EdU pulse, Fig. 6c, d), compared to 10% in mutant hearts (Fig. 6c, d). Nucleation analysis revealed that Dot1L cKO hearts had almost twice as many mononucleated CMs (34% in cKOs versus 19% in controls) at the expense of binucleated CMs (62% in cKOs versus 80% in Ctrl) (Fig. 6e). Quantification of proliferation ratios across different nucleation categories (Fig. 6f) revealed that Dot1L cKOs had a moderate increase in the percentage of EdU+ mononucleated CMs (8.1% in cKOs versus 5.2% in Ctrl) and a significant 4.9-fold increase in the percentage of EdU+ binucleated CMs (9.1% in cKOs versus 1.86% in Ctrl).

To validate these analyses, we assessed proliferation of control and Dot1L cKO CMs using the *Rosa26-Fucci2a* cell cycle reporter³⁴. Expression from this allele is Cre-dependent (thus, restricted to CMs in our analyses) and labels G1 CMs in red, and actively proliferating cells in yellow (G1/S) or green (S/G2/M)³⁴. Analysis of histological sections of P10 hearts (Fig. 6g, h and Supplementary Fig. 6a, d) revealed that in controls the majority of CMs (67%) were in G1. G1/S CMs could also be detected (32%), but S/G2/M cells were rare (1%). On the other hand, Dot1L mutant CMs showed significantly different cell cycle distribution across all cycling phases: 34% in G1, 61% in G1/S and 5% in S/G2/M (Fig. 6g, h and Supplementary Fig. 6a, d). In addition, increased proliferation of mutant CMs was also confirmed by phospho-Histone 3 (pH3) antibody staining in histological sections (Fig. 6g, i and Supplementary Fig. 6b, c), as well as EdU analysis (2 h EdU pulse, Supplementary Fig. 2d–g) of Dot1L mutants and controls in the *Rosa26-Fucci2a* background. As left ventricle-specific transcripts were particularly affected in E16.5 Dot1L cKO hearts, we wondered whether a similar chamber tropism was operative for the neonatal proliferative phenotype. Quantification of chamber-specific rates of proliferation (as assessed by Fucci2a reporter, pH3 or EdU incorporation, Supplementary Fig. 6a, b, f) revealed similar results in all compartments analyzed (RV, septum, LV, RA and LA), suggesting that the cell cycle phenotype is mechanistically distinct from the decreased expression of patterning genes observed during embryogenesis.

To gain insight into gene expression networks underlying the sustained proliferation of Dot1L cKO CMs, we used neonatal (P1) FACS-sorted CMs to perform mechanistic analyses similar to those described for E16.5: RNA-seq, H3K79me2 and H3K27ac ChIP-seq, followed by bioinformatic analysis leveraging a publicly available CM Hi-C dataset³¹ and employing the ABC method to predict interactions between regulatory elements and their target genes (Fig. 7, Supplementary Fig. 7 and Supplementary Data 1b–4b and 6b). These analyses revealed that, similar to E16.5, the majority of genes downregulated in cKO CMs (72.8%) were directly regulated by DOT1L via H3K79me2 both in gene body and REs (Fig. 7b). On the other hand, the vast majority of genes upregulated in cKOs (92.4%) did not have gene body H3K79me2, but 79.9% had interactions with K79-REs (Fig. 7c). Similar to what we had observed in E16.5 CMs, from all K79-REs, the vast majority (96%) did not show altered H3K27ac signal in Dot1L cKO CMs (Supplementary Data 3b). Differential H3K27ac was not a requirement for gene expression regulation by K79-REs as, within differentially expressed



genes associated with K79-REs, 65% of downregulated genes and 60% of upregulated genes did not show differences in H3K27ac between genotype groups (Fig. 7b, c and Supplementary Data 4b). Functional annotation revealed that multiple genes upregulated in cKOs and associated with K79-REs had important roles in the neuronal system and in synaptic structure or activity. qPCR analyses (Supplementary Fig. 8) showed that genes in these categories were already upregulated

in Dot1L cKO hearts in early cardiogenesis (E10.5), suggesting that their modulation is a direct consequence of DOT1L absence, rather than an indirect effect. On the other hand, genes downregulated in cKOs with H3K79me2 in gene body and REs were involved in CM differentiation (Fig. 7d and Supplementary Data 5b). Within these was the gene encoding the cell cycle regulator p27 (*Cdkn1b*) (Fig. 7e). Interestingly, p27 knockouts have a phenotype of increased heart size due to

Fig. 4 | DOT1L controls transcription of target genes via a combination of H3K79me2 in gene bodies and regulatory elements. **a** Volcano plot displaying H3K79me2 ChIP-seq peaks significantly enriched in E16.5 Dot1L cKO vs Ctrl CMs. **b** Pie chart indicating the genomic distribution of differential H3K79me2 ChIP-seq peaks in E16.5 CMs. **c** Metagene profiles showing the average distribution of H3K79me2 input-normalized density relative to Transcription Start Site (TSS) and Transcription Termination Site (TTS) with ± 2 Kb flanking regions. **d** Fraction of gene body covered with H3K79me2 in downregulated genes (left graph) and in upregulated genes (right graph). **e** Graph representing the percentage of down- and up-regulated genes in E16.5 cKO CMs with (Coverage ≥ 50 reads and Fraction of gene body ≥ 0.2) or without (Coverage < 50 reads or Fraction of gene body < 0.2) gene body H3K79me2 in E16.5 Ctrl CMs. **f** Percentage of genes with H3K79me2 in the gene body (GB) or without H3K79me2 in the gene body across the distinct quartiles of expression. This modification was abundant amongst highly expressed genes (4th quartile of RNA expression) and progressively decreased towards the

lower quartiles of expression. Globally more than half (58%) of all genes expressed in E16.5 CMs had gene body H3K79me2. **g** Heatmap indicating the number of total and shared regions between H3K27ac ChIP-seq peaks in Ctrl CMs, differential H3K79me2 ChIP-seq peaks and promoters (± 200 bp around 5'TSS) in E16.5 CMs. Different intensities of colors indicate the fraction (%) of shared peaks. **h** UpSet plot indicating the number and percentage of genes up and downregulated in cKOs versus Ctrl with or without H3K79me2 in gene body (GB) and/or regulatory elements (REs) in E16.5 CMs. Metascape pathway analysis of genes downregulated with H3K79me2 in GB and K79-REs (i) and upregulated without gene body H3K79me2 but with K79-REs (j) in Dot1L cKO versus Ctrl E16.5 CMs. Top 5 enriched categories are shown, sorted by $\text{Log}_{10} P$ value. **k** Motif enrichment analysis ranking transcription factors (TFs) enriched in K79-REs associated with upregulated genes without H3K79me2 in GB versus K79-REs associated with downregulated genes with H3K79me2 in GB.

defective CM cell cycle arrest that resembles the one of Dot1L cKOs³⁵. This observation strongly suggested that disrupted p27 expression accounts, at least in part, for the sustained CM proliferation observed in Dot1L cKOs (Fig. 7h). In addition to p27, it is likely that this phenotype is also mediated by other upstream factors. To identify other factors potentially involved in the sustained proliferation observed in Dot1L cKOs, we determined transcriptional regulators downregulated in Dot1L cKOs (Fig. 7f) and assessed transcription factor binding sites enriched in the K79-REs associated with genes downregulated in Dot1L cKOs with H3K79me2 in their gene body and in K79-REs (Fig. 7g). Together, these analyses suggested that TFs such as MEIS1 and MEIS2 might also play a role in the observed proliferative phenotype, which is consistent with known functions for these TFs in CM development post birth^{36,37}. Altogether these observations suggested that in the neonatal period DOT1L directly promotes expression of genes involved in CM maturation and cell cycle withdrawal whilst repressing expression of genes associated with non-CM functions.

Discussion

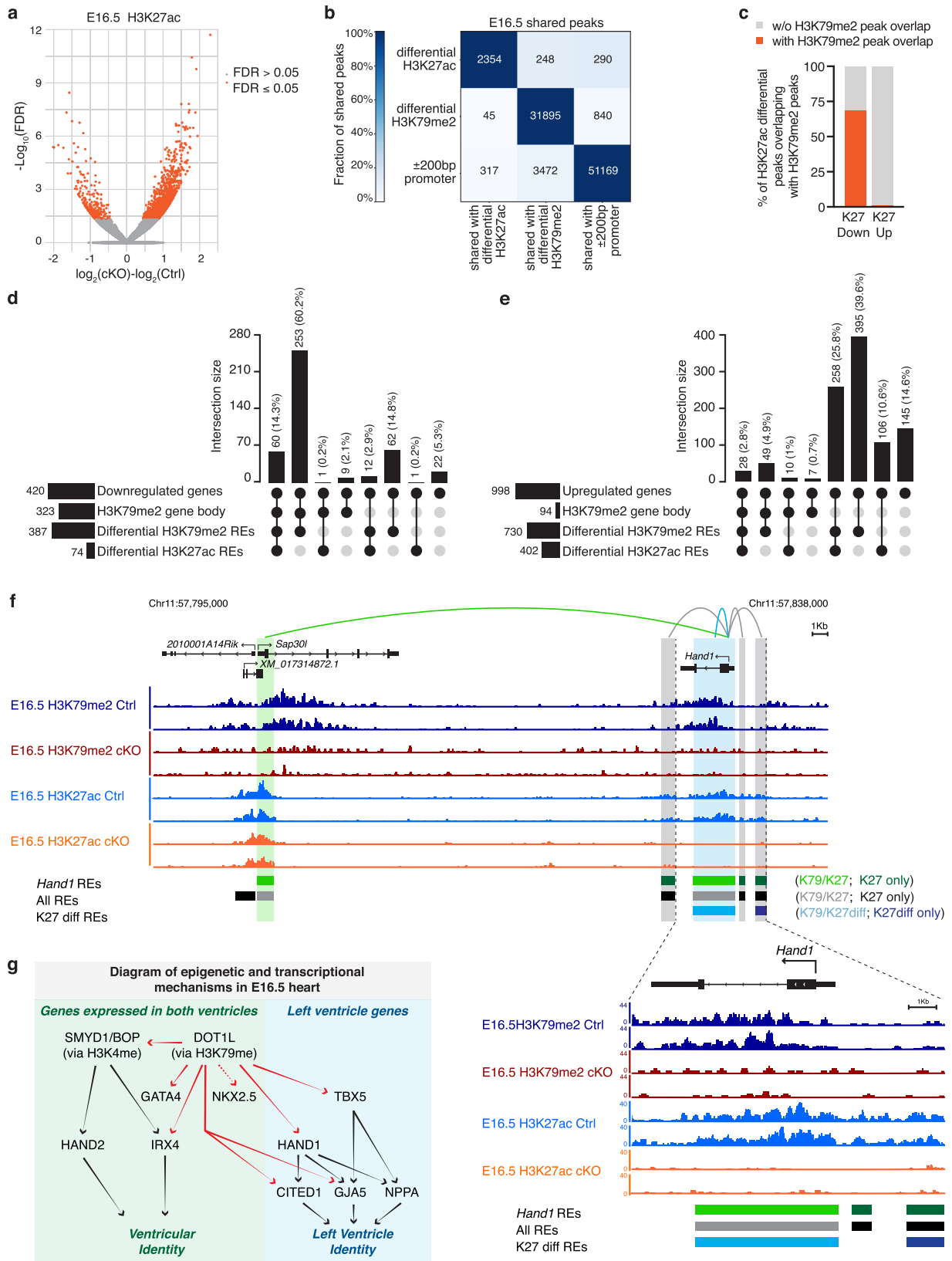
Devising a blueprint for transcriptional and epigenetic mechanisms regulating normal heart development is of critical importance for our understanding of congenital heart disease and to pave the way for regenerative therapies of the adult heart post injury. Our previous studies using embryonic stem cells led to the hypothesis that DOT1L might play a relevant role in cardiogenesis, which we confirmed with the in vivo experiments reported here. In addition, we deciphered previously unrecognized mechanisms of action of this epigenetic modifier. *Dot1l* is ubiquitously expressed in the heart and its CM-specific conditional ablation resulted in a fully penetrant phenotype of an enlarged and rounded heart that culminated in neonatal lethality (Figs. 1 and 2). Genome-wide transcriptomic and ChIP-seq assays provided detailed mechanistic insight into gene expression programs underlying this phenotype, highlighting DOT1L as a direct regulator of two processes that take place at distinct developmental stages: expression of critical cardiogenic factors in embryogenesis (Figs. 3, 4 and 5) and neonatal CM maturation and cell cycle withdrawal (Figs. 6 and 7). Detailed bioinformatic analyses revealed that DOT1L regulated expression of its targets via methylation of H3K79 in gene body, regulatory regions, or a combination of both. Changes in H3K27ac profiles were observed in a significant number of K79-REs associated with genes differently expressed in Dot1L cKOs, but were not a requirement for gene expression regulation by K79-REs.

In embryogenic stages, DOT1L directly regulated several genes involved in defining distinct cardiac chambers. Amongst these, there was a clear enrichment in genes required for left ventricular identity (*Hand1*, *Tbx5*, *Cited1*, *Gja5*). From these, *Hand1* emerged as a gene particularly sensitive to absence of DOT1L activity. Likely reflecting this tight regulation, the dilated and rounded heart phenotype with neonatal lethality displayed by Dot1L mutants strongly resembles the

phenotype of *Hand1* conditional mutants²¹. How disruption of a left ventricle-specific transcription factor results in a rounded heart morphology is not currently clear and addressing this question will require a full understanding of gene expression networks downstream of HAND1. *Hand1* cardiac cKOs also display ventricular septal and outflow defects, whereas Dot1L mutants do not. This is likely a consequence of residual *Hand1* transcripts at earlier embryonic stages – *Hand1* transcripts were completely absent from E16.5 Dot1L cKO CMs but could be detected at E10.5 (although at much lower levels than in stage-matched controls, Fig. 3). Interestingly, postnatally *Hand1* is restricted to the cardiac conduction system and disrupting this specific expression causes conduction system problems¹⁸. Therefore, it is conceivable that electrocardiographic anomalies observed in Dot1L cKOs are, at least in part, caused by blunted *Hand1* expression. The dependence of *Hand1* expression on DOT1L is of particular relevance for our understanding of epigenetic regulation of cardiogenesis. Previous studies have shown that SMYD1 (that catalyzes H3K4 methylation) is critical for activation of *Hand2*²², but an epigenetic mechanism specifically regulating *Hand1* has remained elusive. Together with these previous observations, our results suggest a model in which DOT1L and SMYD1, two enzymes expressed throughout the heart, are essential for maintenance of normal levels of asymmetrically expressed cardiogenic factors (Fig. 5g).

DOT1L depletion resulted in absence of *Hand1* expression without affecting expression of *Hand2*, whereas absence of SMYD1 results in blunted activation of *Hand2* without affecting expression of *Hand1*²². DOT1L also regulated expression of *Smyd1* itself, revealing a complex epigenetic mechanism controlling normal cardiogenesis. It is unclear if SMYD1 is involved in regulating *Dot1l* expression. Interestingly, despite the fact that *Smyd1* transcripts were significantly downregulated in Dot1L cKOs, its critical target *Hand2* was not altered in these mutants, suggesting that the levels of SMYD1 protein present in Dot1L cKOs are sufficient to sustain mechanisms of transcriptional regulation dependent on this enzyme. Despite regulating distinct *Hand* genes, DOT1L and SMYD1 also share some common targets. For example, transcription of *Irx4* is dependent on the action of both enzymes (our own data and ref. 22). IRX4 is required for ventricular myocyte identity, and it is conceivable that DOT1L and SMYD1 act in a coordinated way to ensure a rheostatic control of *Irx4* transcription. How DOT1L and SMYD1, two enzymes that are expressed throughout the heart, achieve regulation of chamber-specific genes (ventricular versus atrial, left versus right) is not currently known and will be the subject of future studies. Given that several of DOT1L targets are also targeted by critical regulators of cardiogenesis, such as NKX2.5 and TBX5, it is possible that the specificity of these epigenetic enzymes is derived from their interactions with distinct transcription factors.

Initial studies suggested that H3K79 methylation is a generic marker of active transcription⁵, but our datasets clearly revealed that, in CMs, H3K79me2 methylation is not a basic requirement for gene



transcription, as in control animals, multiple genes are active without having gene body H3K79 methylation. From the group of genes that bear this modification in controls and lose it in mutant hearts, two major categories emerge: those that are modulated upon DOT1L loss and those that are unaffected by the absence of this enzyme. The latter may reflect mechanisms of epigenetic redundancy. Our bioinformatic

analyses integrating differential H3K79me2 ChIP-seq peaks with a mark of regulatory regions³⁸ (CM H3K27ac ChIP-seq) and 3-dimensional genomic conformation (CM Hi-C) suggested a model in which, in addition to gene body H3K79me2, DOT1L mediates expression of target genes via H3K79me2 in cis-regulatory elements. In loci with gene body H3K79me2, those associated with K79-REs were more

Fig. 5 | H3K27ac Relationship to H3K79me2. **a** Volcano plot displaying H3K27ac ChIP-seq peaks significantly enriched in E16.5 Dot1L cKO vs Ctrl CMs. **b** Heatmap indicating the number of total and shared regions between differential H3K27ac ChIP-seq peaks, differential H3K79me2 ChIP-seq peaks and promoters (± 200 bp around 5'TSS) in E16.5 CMs. Different intensities of colors indicate the fraction (%) of shared peaks. **c** Graph representing the percentage of differential H3K27ac ChIP-seq peaks in E16.5 cKO vs Ctrl CMs overlapping (orange) or not overlapping (gray) with H3K79me2 ChIP-seq peaks. H3K27ac downregulated peaks are more often differential for H3K79me2 than H3K27ac upregulated peaks (oddsratio = 0.004; p -value = 5.93×10^{-212}). UpSet plots indicating the number and percentage of genes down- (**d**) and upregulated (**e**) in E16.5 Dot1L cKO versus Ctrl CMs with or without H3K79me2 in gene body (GB) and/or regulatory elements (REs) and with or without differential H3K27ac REs. **f** Browser tracks displaying H3K79me2 and H3K27ac

ChIP-seq profiles of Ctrl (dark and light blue respectively) and Dot1L cKO (red and orange respectively) E16.5 CMs in the genomic region containing the *Hand1* locus. Loops display all regulatory interactions between REs and the *Hand1* gene, as identified by the ABC analysis. Gray loops identify interactions with REs without H3K79me2, green loops identify interactions with K79-REs, whereas blue loops represent interactions with K79-REs that additionally have H3K27ac REs differentially enriched between Dot1L Ctrl and cKOs. For reference, all REs are displayed in the middle lane in black or gray, regardless of their regulatory association with the *Hand1* gene. Differential H3K27ac REs are indicated in the bottom lane in light blue when they overlap a K79-RE or dark blue when they do not. **g** Diagrammatic representation summarizing the involvement of DOT1L in mammalian cardiogenesis. Genes directly regulated by DOT1L (via H3K79me2) are highlighted in red.

downregulated in Dot1L cKOs. The magnitude of downregulation correlated with two variables: extent of H3K79me2 gene body coverage and the number of K79-REs interacting with the gene, suggesting these two forms of H3K79me2 synergize to promote expression of target genes. Interestingly, our analyses suggested that K79-REs could also be of an inhibitory nature, blocking transcription of target genes that lack gene body H3K79me2. To our knowledge this is the first report of an involvement of DOT1L/H3K79me2 in this form of gene expression regulation, however, it is also conceivable that gene upregulation observed in Dot1L cKOs reflects secondary effects (for example, downregulation of a transcription regulator with repressor properties).

In mice, after the first week of life, most CMs are binucleated and completely withdrawn from the cell cycle²³. Resistance of mature CM to proliferation is a major hurdle for cardiac regeneration^{39,40}. Thus, devising strategies to promote reacquisition of proliferative potential in CMs is a major priority in cardiac regenerative medicine^{39,40}. FACS and histological studies using biochemical and genetic strategies to identify CM-specific proliferation revealed that Dot1L cKO CMs retain proliferative potential at P10, a stage at which control counterparts are already withdrawn from the cell cycle (Fig. 6). Notably, Dot1L cKO hearts also had a higher percentage of mononucleated CMs than control hearts. Transcriptome analyses suggested that DOT1L depleted CMs are less mature than control CMs and fail to activate expression of p27, a known repressor of cell cycle. Permanent ablation of DOT1L results in severe consequences and lethality, but our results revealing sustained proliferation of DOT1L-depleted CMs suggest that temporary inhibition of DOT1L in postnatal hearts might be a strategy to promote re-acquisition of mitotic potential in CMs. However, this sort of approach would need to be finely titrated in follow up studies, as there are potential risks associated with the activation of genes normally repressed in CMs or downregulation of DOT1L-dependent genes necessary for normal CM function.

Methods

Mouse strains and experiments

Animal experiments were conducted according to protocols approved by the Institutional Animal Care and Use Committee at University of California, San Diego and the German local ethic committee (Regierungspräsidium Darmstadt, Hessen). All transgenic lines used were kept on an outbred background (Mus Musculus, Black-Swiss, Charles River laboratories). Mice were maintained in plastic cages with filtered air intake ports (Techniplast) on a 12 h light cycle and have free access to water and food (Teklad LM-485 irradiated diet, Harlan Laboratories, catalog number 7912). All mouse housing rooms are maintained at 72 ± 2 degrees Fahrenheit and 30–70% relative humidity. Adult (2–12 months old) males and females were used for breeding. For analyses conducted in embryonic stages, embryos were staged according to the embryonic day (E) on which dissection took place,

with noon of the vaginal plug day being considered as E0.5 and birth typically occurring at E19. Experimental mice (males and females) were analyzed from embryonic day E10.5 to postnatal day 10. *Dot1L*^{fl^{ox}} mice were obtained from the KOMP Repository (CSD29070) <https://www.komp.org/geneinfo.php?geneid=54455>. Floxed-out *Dot1L*^{fl^{ox}} mice were generated by crossing the *Dot1L*^{fl^{ox}} allele with the epiblastic *Meox2-Cre* allele obtained from JAX laboratories (Stock No: 026858). *xMlc2-Cre* mice²⁴ were gently provided by Timothy Mohun. The *Rosa26-tdTomato (Ai9) (tdTom)* indicator allele²⁵ was purchased from JAX (Stock No: 007905). The *Rosa26-Fucci2A* cell cycle reporter allele was gently provided by Ian James Jackson³⁴. All experiments were performed using littermate cKOs and Ctrl. Images shown are representative examples of experiments with $n \geq 3$ biological replicates. In experiments assessing proliferation, mice received an injection of EdU (at P1 25 μ L of a 3 g/L stock and at P10 50 μ L of a 3 g/L stock) 2 h or 24 h prior to euthanasia.

For genotyping, genomic DNA was extracted by adding 250 μ L of 50 mM NaOH to a tail tip biopsy and heating at 95 °C for 30 min. The solution was then neutralized by adding 50 μ L of 1 M Tris-HCl (pH 8.0). Genotyping PCR reactions (36 amplification cycles) were performed using Taq DNA Polymerase with ThermoPol Buffer (New England Biolabs, M0267L), dNTPs from Promega (U1511) and 1 μ L of DNA solution. The following genotyping primers were used:

XMlc2Cre-Fw 5'-TAGGATGCTGAGAATCAAATGT-3';
 XMlc2Cre-Rev 5'-TCCCTGAACATGCCATCAGGTTTC-3';
 Dot1L-Fw 5'-CCATATTAGTGTTCAAGGGCTACT-3';
 Dot1L-fl/wt-Rev 5'-AGCATAAGGATGCCAACTACTAAC;
 Dot1L-null-Rev 5'-AAGGAGTCTCTCATCATAGCTACTT-3';
 Rosa26-tdTomato-Fw 5'-CTGTTCCCTGTACGGCATGG-3';
 Rosa26-tdTomato-Rev 5'-GGCATTAAAGCAGCGTATCC-3';
 Rosa26-wt-Fw 5'-AAGGGAGCTGCAGTGGAGTA-3';
 Rosa26-wt-Rev 5'-CCGAAAATCTGTGGGAAGTC-3';
 Rosa26-wt-Fw 5'-CAAAGTCGCTCTGAGTTGTTATCAG-3';
 Rosa26-wt-Rev 5'-GGAGCGGGAGAAATGGATATGAAG-3';
 Rosa26-Fucci2a-Rev 5'-TCACCCAGGAGTCATTTGAT-3'.

Echocardiography and electrocardiography

Echocardiography was performed at P5 and P10 using a Vevo3100 imaging system (VisualSonics) with MX550 (for P10) or MX700 (for P5) probes. Two-dimensional-guided M-mode images of the short axis at the papillary muscle level were recorded. Data were analyzed by VevoLab 3.2.5 software using Auto-LV technology to minimize the variability between the measurement and confirmed by an experienced researcher. For electrocardiography (ECG) measurements the PR interval was measured from the beginning of the P-wave to the beginning of the QRS complex; QRS duration was measured from the first deflection of the Q-wave to the nadir of the S-wave (defined as the point of minimum voltage in the terminal portion of the QRS complex); the R-R interval was obtained as the average R-R interval over the sampling period.

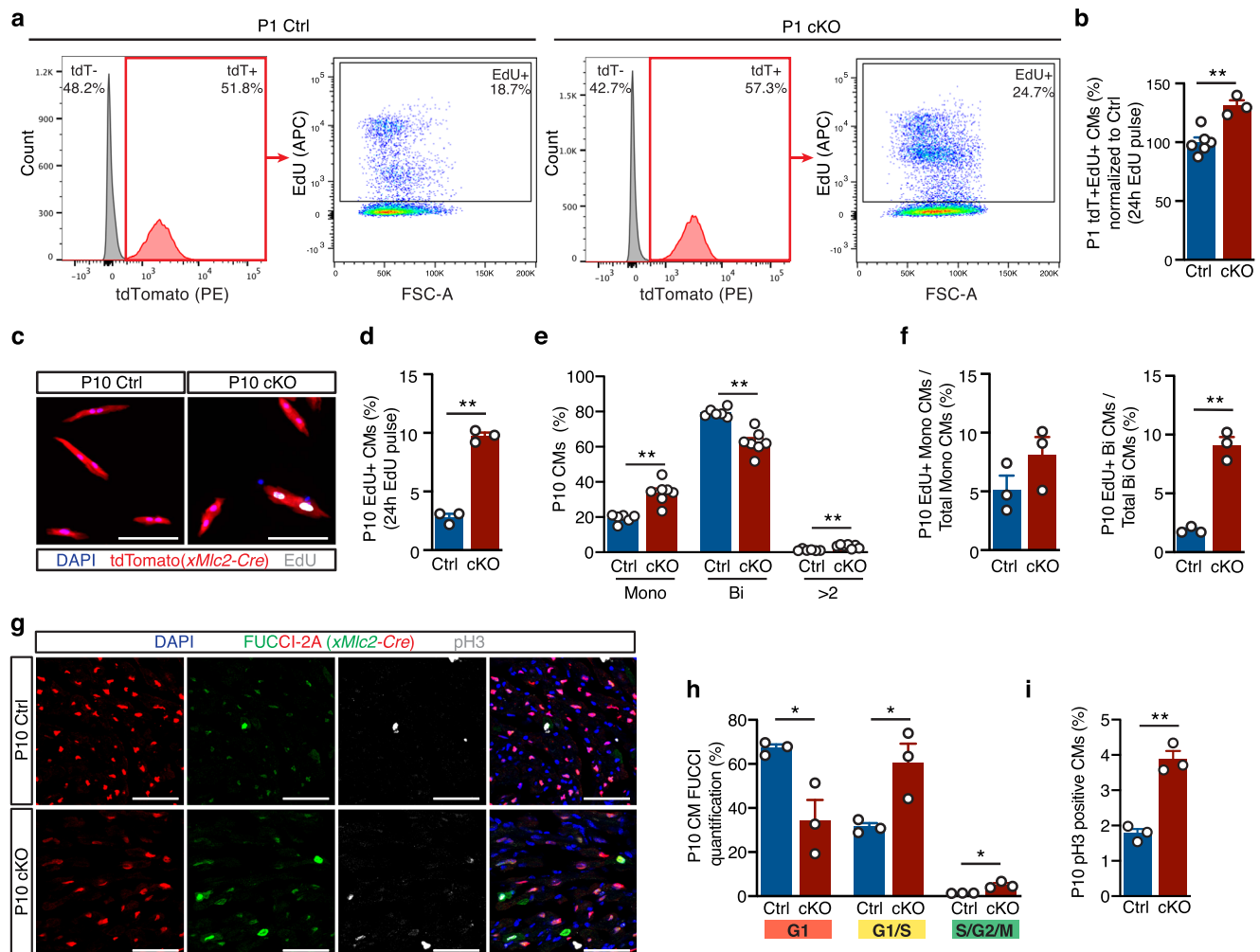


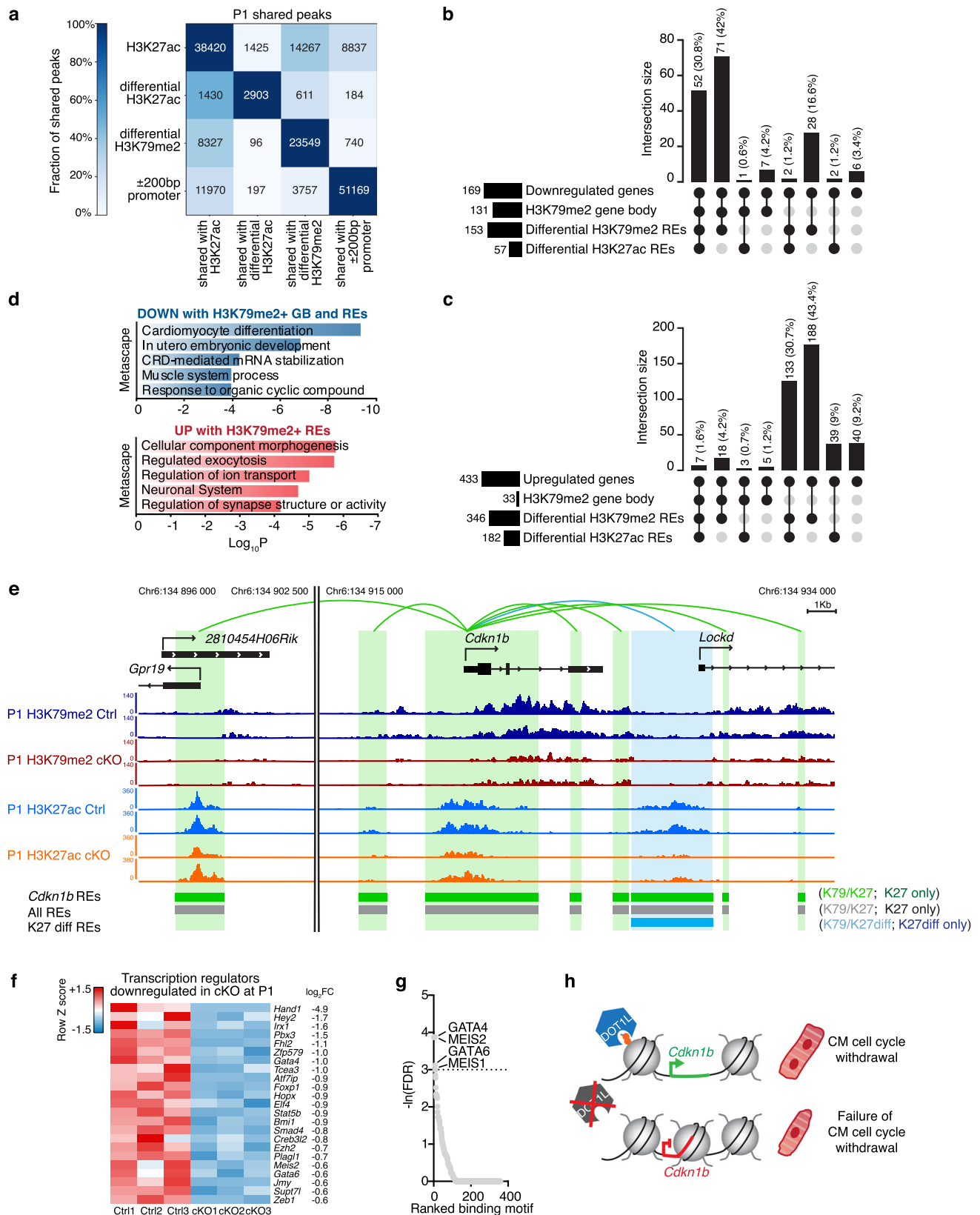
Fig. 6 | Neonatal Dot1L cKO cardiomyocytes fail to undergo cell cycle withdrawal. Representative FACS analysis (**a**) and respective quantification (**b**) showing significantly increased EdU incorporation within P1 CMs (tdTomato+ cells) of Dot1L cKO vs Ctrl hearts (Ctrl $N=6$, cKO $N=3$ biological replicates, unpaired t -test, two-sided, $P=0.0025$). Representative immunofluorescence images (**c**), and respective quantification (**d**) showing significantly increased rates of EdU incorporation in P10 CMs isolated from Dot1L cKO vs Ctrl hearts. DAPI in blue, endogenous tdTomato signal driven by *xMlc2-Cre*; *Rosa26*-tdTomato in red and EdU in white. (Scale bar = 100 μ m; Mean of 755 CMs counted per heart from $N=3$ biological replicates, unpaired t -test, two-sided, $P<0.0001$). **e** Quantification of relative percentage of mononucleated (Mono), binucleated (Bi) or multinucleated (>2) CMs in P10 Dot1L Ctrl and cKOs. At P10, Dot1L cKO hearts had more mononucleated (Mono) and less binucleated (Bi) CMs than littermate Ctrl (mean of 697 CMs counted per heart from $N=6$ Ctrl and $N=7$ cKO biological replicates, unpaired t -test, two-sided, Mono $P=0.0002$, Bi $P<0.0001$, >2 $P=0.0007$). **f** Quantification of percentage of EdU+

CMs within mononucleated (left graph) and binucleated (right graph) CMs of P10 Dot1L Ctrl and cKOs (Mean of 755 CMs counted per heart from $N=3$ biological replicates, unpaired t -test, two-sided $P=0.0006$). Immunofluorescence images (**g**) and respective quantification (**h**, **i**) of P10 Dot1L Ctrl and cKO hearts on a *Rosa26*-Fucci2A background. Red-only nuclei represent CMs in G1; red + green (yellow) nuclei represent CMs in G1/S; green-only nuclei correspond to CMs in S/G2/M, phosphor-Histone 3 (pH3) staining in white. Dot1L cKO hearts had a significantly higher percentage of CMs in G1/S and in S/G2/M compared to Ctrl (h, mean of 2070 CMs counted per heart from $N=3$ biological replicates, unpaired t -test, two-sided G1 $P=0.0256$, G1/S $P=0.0308$, S/G2/M $P=0.0126$) and of phospho-Histone3+ (pH3) CMs (i, mean of 1019 CMs counted per heart from $N=3$ biological replicates, unpaired t -test, two-sided $P=0.0014$). (Scale bar = 50 μ m; sections have been quantified from all the compartments of the heart). In all graphs from **b** to **i** data is presented as mean \pm SEM; * represents $P\leq 0.05$, ** $P\leq 0.01$. Source data are provided as a Source Data file.

Immunohistochemistry

Tissues were isolated in cold PBS and fixed in 4% paraformaldehyde at 4 $^{\circ}$ C overnight. Tissues were dehydrated in a sucrose gradient (5% for 1 h, 12% for 1 h, 20% overnight) and embedded in a 1:1 mix of 20% Sucrose and Tissue-Tek Optimal Cutting Temperature compound (OCT, Sakura). 10 μ m thick histological sections were cut using a cryostat (Leica CM3050) and kept at -20 $^{\circ}$ C (short-term) or -80 $^{\circ}$ C (long-term) until being processed for immunofluorescence. Sections or isolated CMs were blocked in 10% donkey serum before incubation with antibodies. Primary antibodies were incubated at 4 $^{\circ}$ C overnight. The following primary antibodies were used: anti- α Sarcomeric Actinin (1:400 Abcam #ab68167), anti-Myomesin (1:200 mMaC #B4 developed by Perriard, J.-C. was obtained from the

Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa), anti-Vimentin (1:100 Abcam #ab45939), anti-TNNT (1:100 Thermo Fisher #MA5-12960), anti-PDGFR- α (1:200 R&D Systems #AF1062), anti-Collagen 1 (1:200 Abcam #ab34710), anti-tdTomato (1:100 Sicgen #ab8181-200), anti-GFP (1:600 Abcam #ab13970), anti-phosphoH3 (1:100 Millipore #06-570). EdU incorporation was detected using a Click-iT[®] EdU kit (Thermo Fisher Scientific #C10340). Secondary antibodies were incubated at room temperature for 1 h and 30 min. The following secondary antibodies were used at a concentration of 1:400: donkey anti-rabbit Alexa fluor 647 (Thermofisher #A31573), donkey anti-goat Alexa fluor 488 (Thermofisher #A11055), donkey anti-goat Alexa fluor 555 (Thermofisher #A21432), donkey



anti-chicken Alexa fluor 488 (Jackson Immuno Research #703-545-155), donkey anti-rabbit Alexa fluor 555 (ThermoFisher #A31572), donkey anti-mouse Alexa fluor 488 (ThermoFisher #A21202), donkey anti-goat Alexa fluor 647 (ThermoFisher #A21447). Masson Trichrome staining of tissue sections was performed using the Masson trichrome stain kit (Sigma, HT15), following instructions

provided by the manufacturer. Immunofluorescence images were acquired using a Keyence BZX-700 fluorescent microscope, an Olympus FV1000 or a Leica SP8 scan confocal microscope. Bright field images were acquired using a Nikon Eclipse C1 microscope. Image processing was performed with Fiji and Volocity software.

Fig. 7 | Mechanisms underlying sustained proliferation of DOT1L cKO cardiomyocytes. **a** Heatmap indicating the number of total and shared regions between differential H3K27ac ChIP-seq peaks, differential H3K79me2 ChIP-seq peaks and promoters (± 200 bp around 5'TSS) in P1 CMs. Different intensities of colors indicate the fraction (%) of shared peaks. UpSet plots indicating the number and percentage of genes down- (**b**) and upregulated (**c**) in P1 Dot1L cKO versus Ctrl CMs with or without H3K79me2 in gene body (GB) and/or regulatory elements (REs) and with and/or without differential H3K27ac REs. **d** Metascape pathway analysis of downregulated genes with H3K79me2 in GB and K79-REs (top) and upregulated genes without gene body H3K79me2 but with K79-REs (bottom) in Dot1L cKO vs Ctrl P1 CMs. Top 5 enriched categories are shown, sorted by $\text{Log}_{10} P$ value. **e** Browser tracks displaying H3K79me2 and H3K27ac ChIP-seq profiles of Ctrl (dark and light blue respectively) and Dot1L cKOs (red and orange respectively) P1 CMs in the genomic region harboring the *Cdkn1b* locus (encoding p27). Loops display all

regulatory interactions between REs and the *Cdkn1b* gene, as identified by the ABC analysis. Gray loops identify interactions with REs without H3K79me2, green loops identify interactions with K79-REs, whereas blue loops represent interactions with K79-REs that additionally have H3K27ac REs differentially enriched between Dot1L Ctrl and cKOs. For reference, all REs are displayed in the middle lane in black or gray, regardless of their regulatory association with the *Cdkn1b* gene. Differential H3K27ac REs are indicated in the bottom lane in light blue when they overlap a K79-RE or dark blue when they do not. **f** Heatmap showing the expression of all transcription regulators downregulated in Dot1L cKO CMs at P1. **g** Motif enrichment analysis ranking transcription factors (TFs) enriched in REs associated with downregulated genes with H3K79me2 in GB and K79REs versus upregulated genes without H3K79me2 in GB and with K79REs. **h** Diagrammatic representation of mechanism of defective CM cell cycle withdrawal in the absence of DOT1L.

RNA scope

RNA scope fluorescent in situ hybridizations (ISH) were conducted using the RNA scope Multiplex Fluorescent Reagent Kit v.2 (Advanced Cell Diagnostics #323100) following standard protocols provided by the manufacturer, using the following RNA scope probes (ACDbio): Mm-Dot1L-C2 (#533431-C2); Mm-Hand1-C2 (#429651-C2); Mm-Irx4 (#504831); Mm-Hand2 (#499821); Mm-Smyd1-C3 (#1152831-C3).

Cardiomyocyte isolation

Embryonic and neonatal CMs were isolated using a modified version of the a previously described protocol⁴¹. Embryonic or postnatal day 1 hearts were isolated and transferred into ice cold HBSS. Embryonic single cell suspensions were obtained by performing eight rounds of enzymatic digestion (5 min each) with Collagenase type II (0.7 mg/ml Worthington) at 37 °C under agitation. Postnatal day 1 single cell suspensions were obtained by performing an overnight digestion with trypsin (0.5 mg/ml) at 4 °C followed by eight rounds of enzymatic digestion (5 min each) with Collagenase type II (1 mg/ml Worthington) at 37 °C. Cells were collected in cold medium containing 10% horse serum and 5% fetal bovine serum to stop the enzymatic reaction and centrifuged at 210 rcf to allow initial separation of CMs from other cardiac cells. Cell preps were resuspended in FACS buffer (HBSS, 5% FBS, 2.5 mM EDTA) and incubated for 30 min with APC-conjugated, FACS-validated antibodies against CD31 (endothelial cells, clone: MEC13.3; BioLegend #102510), CD45 (leukocytes, clone: 30-F11, BioLegend #103112), CD140a (fibroblasts, clone: APA5, eBioscience #17-1401-81) and TER119 (erythroid cells, clone: TER-119, BioLegend #116212) to avoid doublets between CMs and stromal cells. Live/Dead (Invitrogen #L34957) or DAPI staining was performed to exclude dead cells. Embryonic and postnatal day 1 live and single control and mutant CMs were sorted based on the red fluorescence emitted by the Cre reporter tdTomato using a FACS Aria II or Influx Cell sorter (BD Biosciences) and collected in TRIzol reagent for RNA extraction or cross-linked as described below for chromatin analysis.

Postnatal day 10 CMs were isolated using a Langendorff system using Collagenase type II (0.7 mg/ml Worthington). After perfusion cells were dissociated from the heart, collected in conical tubes and allowed to settle by gravity in order to obtain separation of viable, rod-shaped CMs from dead CMs and from interstitial cells of the heart. For histological analysis, isolated CMs were fixed in 4%PFA and processed for immunostaining.

Protein isolation and Western Blot analysis

Total protein extracts were prepared by lysing samples in RIPA buffer. Protein lysates in Laemmli buffer were separated by electrophoresis on 12% SDS-PAGE gels and transferred for 2 h at 4 °C on to a PVDF membrane (BioRad). After blocking for an hour in 5% dry milk, membranes were incubated overnight at 4 °C with the primary antibody in blocking buffer. The following primary antibodies were used: H3K79me2 (1:1000 abcam #ab3594) and anti-H3 (1:5000 abcam #ab1791). Blots

were washed and incubated with a horseradish peroxidase (HRP)-conjugated secondary antibody generated in Rabbit (1:2000; Cell Signaling Technology #7074) for 1.5 h at room temperature. Immunoreactive protein bands were visualized using an enhanced chemiluminescence reagent (Thermo Fisher Scientific). Protein quantification was achieved using ImageJ software.

Quantification of proliferation by Flow Cytometry

FACS quantification of rates of EdU incorporation was carried out using littermate Dot1L cKOs and Ctrl. EdU detection was done in cell suspensions using the Click-iT® EdU Alexa 647 kit (Thermo Fisher Scientific; C10340), according to the manufacturer's instructions. TdTomato signal was used to discriminate CMs from other lineages of cardiac cells. Stained cells were analyzed using a FACS Canto II flow cytometer (BD Bioscience). DIVA (v8.0.1) and FlowJo (v10.8.1) software (BD Pharmingen) were used for data acquisition and analysis.

RNA extraction, qRT-PCR and RNA-seq

RNA was extracted using Tryzol (Invitrogen #15596026) and Direct-zol RNA Kits (Zymo Research #R2061) following instructions provided by the manufacturers. All transcriptomics analyses (qRT-PCR or RNA-seq) were performed on FACS-sorted CMs, except qPCR in Supplementary Figs. 3a and 8 that were performed on whole heart tissue. For E16.5, P1, P5 and P10 biological replicates were prepared from single hearts. For E12.5, preparation of each biological replicate required pooling of hearts with same genotypes. For each stage analyzed, a minimum of 3 biological replicates, prepared from littermate Dot1L Ctrl and Dot1L cKO hearts, were used. For qRT-PCR experiments, cDNA was produced using the SuperScript VILO cDNA Synthesis Kit (Invitrogen #11754050). qRT-PCR was performed using SYBR Select Master Mix for CFX (Applied biosystems #4472952) on a Bio-Rad CFX96 Real-Time PCR system or Perfecta SYBR green FastMix (VWR # 733-1389) on an Applied Biosystems ViiA7 Real-Time PCR. The following primers were used:

Dot1L-Ex2-Fw 5'-TGCTGCTCATGAGATTATTGAGA-3' (primer hybridizing within the loxP-flanked exon2 of the floxed *Dot1l* allele);

Dot1L-Ex4-Rev 5'-ATGGCCCGTTGTATTGTGC-3';

18s-Fw 5'-AAATCAGTTATGGTTCCTTTGGTC-3';

18s-Rev 5'-GCTCTAGAATTACCAGCTTATCCAA-3';

Col1a1-Fw 5'-CATGTTGAGCTTTGTGGACCT-3';

Col1a1-Rev 5'-GCAGCTGACTTCAGGGATGT-3';

Col3a1-Fw 5'-ACGTAGATGAATTGGGATGCAG-3';

Col3a1-Rev 5'-GGGTTGGGGCAGTCTAGTG-3'.

Gap43-Fw 5'-ATAACTCCCCGTCCTCCAAGG-3'

Gap43-Rev 5'-GTTTGGCTTCGTACAGCGT-3'

Gpc2-Fw 5'-CTGCCCGGCATGAAAGTTTA-3'

Gpc2-Rev 5'-GCGACCATAGGAATCGGAGAA-3'

Nefl-Fw 5'-TGATGTCTGCTCGCTCTTTC-3'

Nefl-Rev 5'-CTCATCCTTGGCAGCTTCTT-3'

Nefm-Fw 5'-ACAGCTCGGCTATGCTCAG-3'

Nefm-Rev 5'-CGGGACAGTTGTAGTCGCC-3'
 Rims4-Fw 5'-CTACTTCCCGTGCATGAATC-3'
 Rims4-Rev 5'-CCTCCATAGTTAAGGTTGCCCT-3'
 Shank1-Fw 5'-TGCATCAGACGAAATGCCTAC-3'
 Shank1-Rev 5'-AACAGTCCATAGTTCCAGCAGC-3'
 Sybu-Fw 5'-GCGATGAAGACTTTACCAGAA-3'
 Sybu-Rev 5'-CCTCGTTGCGTGAGAAAGA-3'
 Ntn2-Fw 5'-GTGATGCGCCTGAAGGATTAT-3'
 Ntn2-Rev 5'-TTCTCATGGGAACAGAACCTTTC-3'
 Gabra4-Fw 5'-ACAATGAGACTCACCATAAGTGC-3'
 Gabra4-Rev 5'-GGCCTTTGGTCCAGGTGTAG-3'
 Gabrr1-Fw 5'-CGAGGAGCACACGACGATG-3'
 Gabrr1-Rev 5'-GTGAAGTCCATGTCAACCTCTG-3'
 Snap25-Fw 5'-CAACTGGAACGCATTGAGGAA-3'
 Snap25-Rev 5'-GGCCACTACTCCATCCTGATTAT-3'
 Amph-Fw 5'-TCCGGGATATTAGCAGCAA-3'
 Amph-Rev 5'-TGGCTCGTAGACTTCATGTAGAG-3'
 Cadps2-Fw 5'-CTTGGTTGTCCGCTACGTTGA-3'
 Cadps2-Rev 5'-GTTGAGCCATTGTTGACAGGC-3'
 Erc2-Fw 5'-AAAGCAGCAGACCCAGAACA-3'
 Erc2-Rev 5'-TGGTGGTGGTGAATGGTG-3'
 Sh3gl2-Fw 5'-AACGATTGAATACCTCCAACCC-3'
 Sh3gl2-Rev 5'-TTCATTCCATGTCCAATGAGTC-3'

For RNA-seq experiments libraries were generated from 25 ng of RNA using the TruSeq Stranded mRNA library Prep kit (Illumina# 20020594) and sequenced on a HiSeq 4000 System (Illumina) using a single read 50 protocol.

ChIP-seq

ChIP-seq was essentially performed as described^{42,43}. Briefly, cells were fixed in 1% formaldehyde/PBS for 10 min at room temperature, the reactions quenched by adding 2.625 M glycine to 125 mM final concentration, 20% BSA to 0.5% final concentration and cells pelleted by centrifugation for 5 min at 1000 g, 4 °C. Cells were washed twice with ice-cold 0.5% BSA/PBS and cell pellets were snap-frozen in liquid nitrogen and stored at -80 °C. For H3K79me2 ChIP-seq fixed cells were thawed on ice, resuspended in 500 µl ice-cold buffer L2 (0.5% Empigen BB, 1% SDS, 50 mM Tris/HCl pH 7.5, 1 mM EDTA, 1× protease inhibitor cocktail) and chromatin was sheared to an average DNA size of 300–500 bp by administering 7 pulses of 10 s duration at 13 W power output with 30 s pause on wet ice using a Misonix 3000 sonicator. The lysate was diluted 2.5-fold with ice-cold L2 dilution buffer (20 mM Tris/HCl pH 7.4@20 °C, 100 mM NaCl, 0.5% Triton X-100, 2 mM EDTA, 1× protease inhibitor cocktail), and one percent of the lysate was kept as ChIP input. For each immunoprecipitation, aliquots of diluted lysate equivalent to 150,000 to 1 million cells, 20 µl of Dynabeads Protein A (for rabbit polyclonal antibodies) and 2 µg anti H3K79me2 antibody (Abcam #ab3594) were combined and rotated overnight at 8 RPM and 4 °C. For H3K27ac ChIP-seq fixed cells were thawed on ice, resuspended in 100 µl ice-cold RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% IGEPAL CA-630 0.25% Na-Deoxycholate, 0.1% SDS, 1 mM EDTA, 1× protease inhibitor cocktail). Chromatin was sheared for 90 min on an Active Motif PIXIL high-throughput sonicator using standard settings (Pulse [N]: 50, PRF:1kHz, Burst Rate: 20 Hz). Two microliter of each lysate was kept to generate ChIP input libraries, and the remainder used to immunoprecipitate H3K27ac-associated DNA by adding 2 µg anti-H3K27ac antibody (Active Motif 39133, lot 31521015) and 20 µl of Dynabeads Protein A (Thermo) and rotating overnight at 8 RPM and 4 °C. The following day, in the case of either H3K79me2 or H3K27ac ChIP-seq, beads were collected on a magnet and washed three times with 150 µl each of ice-cold wash buffer I (10 mM Tris/HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 2 mM EDTA), wash buffer III (10 mM Tris/HCl pH 7.5, 250 mM LiCl, 1% IGEPAL CA-630, 0.7% Deoxycholate, 1 mM EDTA) and twice with ice-cold TET (10 mM Tris/HCl pH7.5, 1 mM EDTA, 0.2% Tween-20). Libraries were prepared

directly on the antibody/chromatin-bound beads: after the last TET wash, beads were suspended in 25 µl TT (10 mM Tris/HCl pH7.5, 0.05% Tween-20), and libraries were prepared using NEBNext Ultra II reagents according to the manufacturer's protocol but with reagent volumes reduced by half, using 1 µl of 0.625 µM Bio Nextflex DNA adapters per ligation reaction. DNA was eluted, proteins digested and crosslinks reversed by adding 4 µl 10% SDS, 4.5 µl 5 M NaCl, 3 µl EDTA, 1 µl proteinase K (20 mg/ml), 20 µl water, incubating for 1 h at 55 °C, then overnight at 65 °C. DNA was cleaned up by adding 2 µl SpeedBeads 3 EDAC (Cytiva) in 61 µl of 20% PEG 8000/1.5 M NaCl, mixing and incubating for 10 minutes at room temperature. SpeedBeads were collected on a magnet, washed twice by adding 150 µl 80% EtOH for 30 s each, collecting beads and aspirating the supernatant. After air-drying the SpeedBeads, DNA was eluted in 25 µl TT and the DNA contained in the eluate was amplified for 12 cycles in 50 µl PCR reactions using NEBNext High-Fidelity 2X PCR Master Mix or NEBNext Ultra II PCR master mix and 0.5 µM each of primers Solexa 1GA and Solexa 1GB. Libraries were cleaned up as above by adding 36.5 µl 20% PEG 8000/2.5 M NaCl and 2 µl Speedbeads, two washes with 150 µl 80% EtOH for 30 s each, air-drying beads and eluting the DNA into 20 µl TT. ChIP library size distributions were estimated following 2% agarose/TBE gel electrophoresis of 2 µl library, and library DNA amounts measured using a Qubit HS dsDNA kit on a Qubit fluorometer. ChIP input material (1 percent of sheared DNA) was treated with RNase for 15 min at 37 °C in EB buffer (10 mM Tris pH 8, 0.5% SDS, 5 mM EDTA, 280 mM NaCl), then digested with Proteinase K for 1 h at 55 °C and crosslinks reversed at 65 °C for 30 min to overnight. DNA was cleaned up with 2 µl SpeedBeads 3 EDAC in 61 µl of 20% PEG 8000/1.5 M NaCl and washed with 80% ethanol as described above. DNA was eluted from the magnetic beads with 25 µl of TT and library prep and amplification were performed as described for ChIP samples. For H3K79me2 ChIP-seq, libraries were sequenced single-end for 75 cycles (SE75) to a depth of 20–25 million reads, for H3K27ac ChIP-seq libraries were sequenced single-end for 76 cycles (SE76) to a depth of 15–22 million reads on an Illumina NextSeq 500 instrument.

Bioinformatic analyses

RNA-seq. Sequencing reads were processed to remove Illumina barcodes and aligned to the UCSC *Mus musculus* reference genome (build mm10) using STAR v.2.5.1b with default parameters⁴⁴. ReadsPerGene.out.tab files were then processed with edgeR⁴⁵. RNA expression was calculated in reads per kilobase per million mapped reads (RPKM) considering the sum of exon length. Differentially expressed coding genes were selected based on the following parameters: FDR ≤ 0.05, RPKM ≥ 1 in at least one biological condition, and log₂ Fold Change ≤ -0.5 for genes downregulated and ≥ 0.5 for genes upregulated in Dot1L cKO vs Dot1L Ctrl. Pathway analysis was performed using METASCAPE⁴⁶.

ChIP-seq alignment and peak-calling. All analyses were performed using the mouse reference genome GRCm38 (mm10) and the gencode gene annotation version vM25. H3K79me2 and H3K27ac ChIP-seq data was processed in the same manner. Bowtie⁴⁷ was applied to align the fastq files to the mouse reference genome. First, the required index structure was built using: `bowtie2-build -f -seed 123 -threads 20 Mus-musculus.GRCm38.dna.primary_assembly.fa mouse_GRCm38_mm10`. To get the alignment files we ran: `bowtie2 -x mouse_GRCm38_mm10 -U<fastq-file> -S<output-file-name>.sam -q -t -p 30`. Conversion of the resulting files (sam) to bam format was done using samtools (version samtools 1.10)⁴⁸. To allow easy visualization in a genome browser, bam files were converted to bigWig (bw) files using deeptools bamCoverage function. Next, peak-calling was performed with MACS2 (version macs2 2.2.7.1)⁴⁹: `macs2 callpeak -t<treatment>.bam -c<input>.bam -n-prefix-name--outdir-output-dir>-f BAM -g 1.87e9 -B`, where <treatment> is either the aligned reads of the Dot1L Ctrl or Dot1L cKO and

<input> the corresponding input signal. For all following analyses we used the narrowPeak files. To compute differential ChIP-seq sites between Ctrl and Dot1L cKOs, we applied DiffBind (version 2.10.0 for H3K79me2 and version 3.4.11 for H3K27ac bioconductor-diffbind)⁵⁰.

Determining genomic distribution of H3K79me2 or H3K27ac peaks.

To determine the location of peaks in relation to distinct genomic regions, the gene annotation was split into exons, UTRs and gene bodies. Promoter regions of 400 bp length, centered around the most 5' TSS, were added. All those genomic regions were made exclusive, so that each base pair of a gene had one unique label. The remaining base pairs located within gene bodies but not overlapping any other feature were labeled as introns. Based on those annotations, we determined the location of each base pair of each ChIP-seq peak, using bedtools (v2.25.0)⁵¹, and visualized their distribution with pie charts. ChIP-seq tracks were visualized with Integrative Genomics Viewer (IGV) (v2.9.2).

Computing coverage and fraction of the gene body covered by H3K79me2. The profiles of the H3K79me2 signal at the gene body were computed using deepTools⁵². We used the bam-files resulting from the bowtie2 analysis and the up- and downregulated genes in bed-file format (with strand information).

First, we applied deepTools bamCompare to determine the mean signal of the two replicates of Ctrl and cKO for each time point: *bam-Compare -b1<treatment_rep1>.bam -b2<treatment_rep2>.bam -o<treatment>_mean_rep1_rep2.bw -of bigwig --scaleFactorsMethod None --operation mean --effectiveGenomeSize 2652783500 -p 25 --normalizeUsing RPKM --binSize 20 --skipNonCoveredRegions*, where <treatment> is either Ctrl or cKO.

Next, we visualized the data using deepTools computeMatrix and plotProfile functionalities:

```
computeMatrix scale-regions -S Ctrl_mean_rep1_rep2.bw cKO_mean_rep1_rep2.bw -R downregulated_genes_FDR_0.05.bed upregulated_genes_FDR_0.05.bed -o inputProfile.mat.gz --endLabel TTS --beforeRegionStartLength 2000 --afterRegionStartLength 2000 --regionBodyLength 5000 -p 20 --skipZeros
```

```
plotProfile -m inputProfile.mat.gz -out profile.pdf --perGroup --startLabel TSS --endLabel TTS --plotTitle "Average distribution of H3K79me2 relative to the distance from TSS and TTS" --samplesLabel "Ctrl (upregulated genes)" "cKO (upregulated)" --regionsLabel "downregulated genes" "upregulated genes" --plotFileFormat pdf
```

To compute the fraction of the gene body which is covered by H3K79me2, we applied bedtools (bedtools v2.25.0) coverage function. Given a set of genomic regions and a bam-file, the function computes the number of reads that overlap each genomic region and the fraction of bases that have a non-zero coverage based on the bam-file. We determined the number of overlapping reads and the fraction for all annotated mouse genes for Ctrl and cKO. Next, the mean fraction (of rep1 and rep2) of the up- and downregulated genes was computed. The density plots are based on all up- and downregulated genes.

Calling regulatory interactions with H3K27ac ChIP-seq and Hi-C.

For the identification of enhancer regions H3K27ac ChIP-seq was performed on FACS sorted CMs at E16.5 and P1. A Hi-C matrix of mouse CMs was downloaded from GEO (GSM2544836)³¹. The Hi-C matrix was normalized using Knight-Ruiz normalization with the Juicebox dump command (version 1.22.01)⁵³ with a resolution of 5000 bp for each chromosome.

```
java -jar juicer_tools_1.22.01.jar dump observed KR<hic-file>chr$ chr$ BP 5000<output-file-name>
```

The regulatory interactions between promoter and candidate enhancers were assessed with an adapted version of the ABC-score of Fulco et al.²⁷ from the STARE software²⁸. The command was as follows: *STARE_ABCpp -b<H3K27ac-peak-file>-n 7 -f<folder-with-HiC-files>-k 5000 -t 0.02 -a gencode.vM25.annotation.gtf -w 5000000 -o<output-*

*path>*For all genes on the autosomes (GRCm38p6) the most 5' TSS was taken and all candidate enhancers in a 5 Mb window centered on the TSS were scored according to the following equation:

$$\overline{ABCscore}_{e,g} = \frac{A_{e,g} \cdot C_{e,g}}{\sum_{i \in E_g} A_{i,g} \cdot C_{i,g}} \quad (1)$$

The interaction of an enhancer (*e*) with a gene (*g*) is described by the gene-specific activity of the enhancer ($A_{e,g}$) and the contact to the gene ($C_{e,g}$). The adapted ABC-score returns the relative contribution of an interaction in relation to all other candidate interactions of that gene (E_g). The gene-specific activity of an enhancer is defined as follows:

$$A_{e,g} = A_e \frac{C_{e,g}}{\sum_{j \in G_e} C_{e,j}} \quad (2)$$

where the activity of an enhancer (A_e), measured as signal of the H3K27ac ChIP-seq peak, is taken relative to the contacts that an enhancer has to its candidate target genes (G_e). The contact between an enhancer and a gene's TSS is the contact frequency of the respective bins in the normalized Hi-C matrix. All candidate enhancer-gene interactions with an adapted score ≥ 0.02 were used for further analyses. Chromosomes X and Y were excluded from the analysis. Statistical differences between the cumulative log₂(fold-change of gene expression) distributions were calculated pairwise with a two-tailed Kolmogorov-Smirnov test (SciPy 1.4.1, stats module, ks_2samp function).

Mapping differential ChIP-seq signal to genes. An RE was considered differential for H3K27ac/H3K79me2 if $\geq 10\%$ of the enhancer's length was covered by a differential ChIP-seq signal.

The REs differential for H3K27ac were mapped to genes via the adapted ABC-score. If a RE had a significant change in H3K27ac signal in cKO versus Ctrl, it was accounted for all genes that had an ABC interaction to that RE in Ctrl. Additionally, we considered H3K27ac signal changes in all REs present only in cKO that were mapped to a gene via ABC interactions in cKOs but not inCtrls.

Motif enrichment analysis. To identify transcription factors likely regulating the activity of K79-REs associated to genes differentially expressed in Dot1L cKOs, a motif enrichment analysis was performed. To this end, K79-REs associated to differentially expressed genes were divided into two groups: those associated with genes downregulated in Dot1L cKOs with H3K79me2 in their gene body (activating K79-REs) and those associated with genes upregulated in Dot1L cKOs without H3K79me2 in their gene body (silencing K79-REs). TF binding motifs (total of 515) were downloaded from the JASPAR database⁵⁴ and only those significantly expressed (RPKM ≥ 1) in E16.5 or P1 CMs were considered in subsequent analyses. Using TRAP⁵⁵, a TF-affinity value per TF for each RE sequence was computed. The value is defined as the sum over all binding site probabilities of the given TF for the current sequence. The TRAP analysis was performed separately for the RE sequences of the up- and downregulated genes. A one-tailed Mann-Whitney test (using R's Wilcox test function, confidence level of 0.975) to identify TFs enriched in the activating K79-REs versus silencing K79-REs and vice versa was performed. The resulting p-values were adjusted by applying the Benjamini-Hochberg procedure⁵⁶. All TFs with an adjusted p-value of ≤ 0.05 were considered significant.

Pathway enrichment analysis. Pathway enrichment analysis of genes downregulated with H3K79me2 in gene body and regulatory elements and genes upregulated with H3K79me2 in regulatory elements but not H3K79me2 in gene body was performed using Metascape web tool⁴⁶ using standard settings.

Quantification of experiments, statistical analysis and reproducibility. CM cell length and width and RNA-scope puncta were measured using Image J. Statistical significance of differences in the survival of DOT1L mice was assessed using Kaplan–Meier survival analysis with the log-rank method of statistics. In Fig. 2 graphs data are expressed as mean \pm SD, in all other graphs, data are expressed as mean \pm SEM, with a minimum of 3 biological replicates (the exact replicate number is described in the legend to each Fig.). Statistical significance of differences among groups was tested by 2-tailed Student's *t* test. A value of $P \leq 0.05$ was considered statistically significant. *represents $P \leq 0.05$, ** $P \leq 0.01$. Analyses were performed with GraphPad Prism software. Each experiment was repeated independently with similar results at least 2–3 times.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All RNA-seq and ChIP-seq data that support the finding of this study have been deposited in the Gene Expression Omnibus (GEO) under accession number code [GSE184192](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE184192) Previously published Hi-C data that were re-analyzed here are available under accession codes [GSM2544836](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM2544836) from GEO. Source data are provided with this paper.

References

- Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425–432 (2007).
- Singer, M. S. et al. Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* **150**, 613–632 (1998).
- Feng, Q. et al. Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr. Biol.* **12**, 1052–1058 (2002).
- McLean, C. M., Karemaker, I. D. & van Leeuwen, F. The emerging roles of DOT1L in leukemia and normal development. *Leukemia* **28**, 2131–2138 (2014).
- Steger, D. J. et al. DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol. Cell Biol.* **28**, 2825–2839 (2008).
- Godfrey, L. et al. DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation. *Nat. Commun.* **10**, 2803 (2019).
- Cattaneo, P. et al. DOT1L-mediated H3K79me2 modification critically regulates gene expression during cardiomyocyte differentiation. *Cell Death Differ.* **23**, 555–564 (2016).
- Jones, B. et al. The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet* **4**, e1000190 (2008).
- Nguyen, A. T. et al. DOT1L regulates dystrophin expression and is critical for cardiac function. *Genes Dev.* **25**, 263–274 (2011).
- Guimarães-Camboa, N. et al. HIF1 α represses cell stress pathways to allow proliferation of hypoxic fetal cardiomyocytes. *Dev. Cell* **33**, 507–521 (2015).
- Wu, T. et al. PRDM16 is a compact myocardium-enriched transcription factor required to maintain compact myocardial cardiomyocyte identity in left ventricle. *Circulation* **145**, 586–602 (2022).
- Evans, S. M., Yelon, D., Conlon, F. L. & Kirby, M. L. Myocardial lineage development. *Circ. Res.* **107**, 1428–1444 (2010).
- Paige, S. L., Plonowska, K., Xu, A. & Wu, S. M. Molecular regulation of cardiomyocyte differentiation. *Circ. Res.* **116**, 341–353 (2015).
- Olson, E. N. Gene regulatory networks in the evolution and development of the heart. *Science* **313**, 1922–1927 (2006).
- Harvey, R. P. Patterning the vertebrate heart. *Nat. Rev. Genet* **3**, 544–556 (2002).
- Srivastava, D. Making or breaking the heart: from lineage determination to morphogenesis. *Cell* **126**, 1037–1048 (2006).
- Bao, Z. Z., Bruneau, B. G., Seidman, J. G., Seidman, C. E. & Cepko, C. L. Regulation of chamber-specific gene expression in the developing heart by *Ir4*. *Science* **283**, 1161–1164 (1999).
- Vincentz, J. W. et al. Variation in a left ventricle-specific Hand1 enhancer impairs GATA transcription factor binding and disrupts conduction system development and function. *Circ. Res.* **125**, 575–589 (2019).
- Srivastava, D. et al. Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nat. Genet* **16**, 154–160 (1997).
- Thomas, T., Yamagishi, H., Overbeek, P. A., Olson, E. N. & Srivastava, D. The bHLH factors, dHAND and eHAND, specify pulmonary and systemic cardiac ventricles independent of left-right sidedness. *Dev. Biol.* **196**, 228–236 (1998).
- McFadden, D. G. et al. The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. *Development* **132**, 189–201 (2005).
- Gottlieb, P. D. et al. Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nat. Genet* **31**, 25–32 (2002).
- Soonpaa, M. H., Kim, K. K., Pajak, L., Franklin, M. & Field, L. J. Cardiomyocyte DNA synthesis and binucleation during murine development. *Am. J. Physiol.* **271**, H2183–H2189 (1996).
- Breckenridge, R., Kotecha, S., Towers, N., Bennett, M. & Mohun, T. Pan-myocardial expression of Cre recombinase throughout mouse development. *Genesis* **45**, 135–144 (2007).
- Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
- Bruneau, B. G. et al. Chamber-specific cardiac expression of *Tbx5* and heart defects in Holt-Oram syndrome. *Dev. Biol.* **211**, 100–108 (1999).
- Fulco, C. P. et al. Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet* **51**, 1664–1669 (2019).
- Hecker, D., Ardakani, F. B. & Schulz, M. H. The adapted Activity-By-Contact model for enhancer-gene assignment and its application to single-cell data. *bioRxiv* <https://doi.org/10.1101/2022.01.28.478202> (2022).
- Li, G. et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84–98 (2012).
- Dao, L. T. M. et al. Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet* **49**, 1073–1081 (2017).
- Rosa-Garrido, M. et al. High-resolution mapping of chromatin conformation in cardiac myocytes reveals structural remodeling of the epigenome in heart failure. *Circulation* **136**, 1613–1625 (2017).
- Lai, A. Y. & Wade, P. A. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat. Rev. Cancer* **11**, 588–596 (2011).
- Bruneau, B. G. et al. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor *Tbx5* in cardiogenesis and disease. *Cell* **106**, 709–721 (2001).
- Mort, R. L. et al. Fucci2a: a bicistronic cell cycle reporter that allows Cre mediated tissue specific expression in mice. *Cell Cycle* **13**, 2681–2696 (2014).
- Poolman, R. A., Li, J. M., Durand, B. & Brooks, G. Altered expression of cell cycle proteins and prolonged duration of cardiac myocyte hyperplasia in p27KIP1 knockout mice. *Circ. Res.* **85**, 117–127 (1999).
- Mahmoud, A. I. et al. Meis1 regulates postnatal cardiomyocyte cell cycle arrest. *Nature* **497**, 249–253 (2013).

37. Alam, P. et al. Inhibition of senescence-associated genes Rb1 and Meis2 in adult cardiomyocytes results in cell cycle reentry and cardiac repair post-myocardial infarction. *J. Am. Heart Assoc.* **8**, e012089 (2019).
 38. Creyghton, M. P. et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* **107**, 21931–21936 (2010).
 39. Guo, Y. & Pu, W. T. Cardiomyocyte maturation: new phase in development. *Circ. Res.* **126**, 1086–1106 (2020).
 40. Eschenhagen, T. et al. Cardiomyocyte regeneration: a consensus statement. *Circulation* **136**, 680–686 (2017).
 41. Louch, W. E., Sheehan, K. A. & Wolska, B. M. Methods in cardiomyocyte isolation, culture, and gene transfer. *J. Mol. Cell Cardiol.* **51**, 288–298 (2011).
 42. Heinz, S. et al. Transcription elongation can affect genome 3D structure. *Cell* **174**, 1522–1536.e1522 (2018).
 43. Texari, L. et al. An optimized protocol for rapid, sensitive and robust on-bead ChIP-seq from primary cells. *STAR Protoc.* **2**, 100358 (2021).
 44. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
 45. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
 46. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
 47. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
 48. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
 49. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
 50. Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
 51. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
 52. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
 53. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
 54. Fornes, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
 55. Sobreira, T. J., Durham, A. M. & Gruber, A. TRAP: automated classification, quantification and annotation of tandemly repeated sequences. *Bioinformatics* **22**, 361–362 (2006).
 56. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
- Research Foundation (Cardio-Pulmonary Institute; Exc2026-1). N.G.-C. was supported by a junior research group grant from the German Center for Cardiovascular Research (DZHK). J.C. and S.M.E. are funded by grants from the National Heart, Lung, and Blood Institute, and the Leducq Foundation. Confocal microscopy was performed at the University of California, San Diego School of Medicine Microscopy Core Facility, which is supported by the grant P30 NS047101. Tim Mohun from the Francis Crick Institute of London (UK) kindly provided the xMlc2-Cre mice. Ian James Jackson from the Institute of Genetic and Molecular Medicine, University of Edinburgh (UK) kindly provided the Rosa26-Fucci2a mice.

Author contributions

P.C. and N.G.-C. conceived the study and performed most of the experiments, analyzed data, acquired funding sources and prepared the manuscript. S.P., G.A., V.L., L.Z., R.C. and S.S. contributed to experiments. M.H., G.F. and S.H. performed ChIP-seq experiments. P.K. analyzed RNA-seq datasets. N.B. and D.H. under supervision by M.H.S. performed additional bioinformatics analyses on ChIP-seq. J.C., G.C. and S.D. contributed to project comments and supervision. S.M.E. conceived and supervised the study and acquired funding sources. All authors commented and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-35070-2>.

Correspondence and requests for materials should be addressed to Paola Cattaneo or Sylvia M. Evans.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

Acknowledgements

P.C. was supported by the Marie Curie International Outgoing Fellowship within the 7th European Community Framework Program under grant agreement No 623739 - The Cardiac Code and a grant from the German Center for Cardiovascular Research (DZHK) and the German