

DOT1L-mediated H3K79me2 modification critically regulates gene expression during cardiomyocyte differentiation

P Cattaneo^{1,2,3}, P Kunderfranco¹, C Greco¹, A Guffanti⁴, GG Stirparo^{2,3}, F Rusconi¹, R Rizzi^{5,6}, E Di Pasquale^{1,2}, SL Locatelli^{2,3}, MVG Latronico¹, C Bearzi^{5,6}, R Papait^{*,1,2} and G Condorelli^{*,1,2,3}

Epigenetic changes on DNA and chromatin are implicated in cell differentiation and organogenesis. For the heart, distinct histone methylation profiles were recently linked to stage-specific gene expression programs during cardiac differentiation *in vitro*. However, the enzymes catalyzing these modifications and the genes regulated by them remain poorly defined. We therefore decided to identify the epigenetic enzymes that are potentially involved in cardiomyogenesis by analyzing the expression profile of the 85 genes encoding the epigenetic-related proteins in mouse cardiomyocytes (CMs), and then study how they affect gene expression during differentiation and maturation of this cell type. We show here with gene expression screening of epigenetic enzymes that the highly expressed H3 methyltransferase disruptor of telomeric silencing 1-like (DOT1L) drives a transitional pattern of di-methylation on H3 lysine 79 (H3K79) in CMs at different stages of differentiation *in vitro* and *in vivo*. Through a genome-wide chromatin-immunoprecipitation DNA-sequencing approach, we found H3K79me2 enriched at genes expressed during cardiac differentiation. Moreover, knockdown of *Dot1L* affected the expression of H3K79me2-enriched genes. Our results demonstrate that histone methylation, and in particular DOT1L-mediated H3K79me2 modification, drives cardiomyogenesis through the definition of a specific transcriptional landscape.

Cell Death and Differentiation (2016) 23, 555–564; doi:10.1038/cdd.2014.199; published online 19 December 2014

Cardiac morphogenesis is a complex process involving rapid specification and differentiation of the cell types that form the heart.^{1–3} During embryonic development, the activation of specific transcriptional programs defines the subsets of lineage-restricted genes that become expressed while the rest of the genome is left in a repressed state.^{4–7} Epigenetic regulation, such as via histone methylation, establishes the global chromatin environment that stabilizes these transcriptional programs, and is thus a crucial phenomenon for orchestrating these biological processes.^{8,9}

Recent reports have shown that histone modification has a fundamental role in defining the transcription program of CM differentiation *in vitro*. In fact, profiling of different methylation marks on various lysines of H3, such as H3K4me1, H3K4me3, H3K27me3, and H3K36me3, revealed that distinct chromatin patterns are coordinated with stage-specific expression of functionally relevant genes: programmed temporal alterations of chromatin structure distinguish the key regulator genes of cardiovascular development from other genes.^{10,11} However, the involvement of most epigenetic marks, and of the enzymes that catalyze these modifications, remains poorly understood during CM differentiation.

Here, we demonstrate that DOT1L H3K79 methyltransferase is required for CM gene transcription: *Dot1L* was highly expressed in mouse CMs at embryonic and neonatal stages; H3K79me2 marks mapped to lineage-specific CM genes in correlation with the transcriptional profile at distinct stages of differentiation; and a reduction in DOT1L interfered with CM differentiation *in vitro*.

Results

***Dot1L* is differentially expressed during differentiation and maturation of CMs.** To identify epigenetic enzymes potentially involved in cardiomyogenesis, we analyzed the expression profile of 85 genes encoding epigenetic-related proteins (15 histone acetyltransferases (HATs), 12 histone deacetylases (HDACs), 31 histone methyltransferases (HMTs), 14 histone demethylases (HDMs), 2 DNA methyltransferases (DNMTs), and 11 chromatin-remodeling factors) in two models: mouse embryonic stem cells (mES) at distinct stages of differentiation (i.e., pluripotent cells, mesodermal precursors, cardiac progenitors, and differentiated CMs) and

¹Laboratory of Cardiovascular Research, Humanitas Clinical and Research Center, Rozzano, Milan 20089, Italy; ²Operational Unit of Milan, Institute of Genetics and Biomedical Research, National Research Council of Italy (CNR), Rozzano, Milan 20089, Italy; ³Department of Biomedical Technologies and Translational Medicine, University of Milan, Milan 20133, Italy; ⁴GenOmnia, via Nerviano 31/b, Lainate, Milan, 20020, Italy; ⁵Casa di Cura Multimedicale Istituto di Ricovero Cura Carattere Scientifico, Milan 20138, Italy and ⁶Institute of Cell Biology and Neurobiology, National Research Council of Italy (CNR), Rome 00100, Italy

*Corresponding author: G Condorelli or R Papait, Cardiovascular Research, Humanitas Research Hospital, Via Manzoni 56, 20089 Rozzano, Milan, Italy. Tel: +39 02 8224521; Fax: +39 02 82245404; E-mail: gianluigi.condorelli@humanitasresearch.it or roberto.papait@humanitasresearch.it

Abbreviations: ChIP-seq, chromatin immunoprecipitation-sequencing; CM, cardiomyocyte; CMa, adult cardiomyocyte; CMp, pup cardiomyocyte; DNMT, DNA methyltransferase; DOT1L, telomeric silencing-1-like H3K79 methyltransferase; GEO, Gene Expression Omnibus; H, histone; HAT, histone acetyltransferase; HDAC, histone deacetylase; HDM, histone demethylase; HMT, histone methyltransferase; me, methylation; mES, mouse embryonic stem cell; TSS, transcription start site

Received 23.6.14; revised 02.10.14; accepted 06.10.14; Edited by JC Marine; published online 19.12.14

mouse cardiac cells isolated at different stages of maturation (immature cells from E8.5, E10.5, E12.5, and E14.5 embryos; neonatal CMs from 1-day-old pups; and mature CMs purified from 2-month-old adult mice). The stages of cardiac differentiation and maturation were confirmed by measuring the expression of markers of pluripotency (*Pou5F1* and *Nanog*), mesodermal populations (*Eomes*, *Mesp1*, and *Brachyury*), cardiac progenitors or immature CMs (*Nkx2-5*, *Isl1*, *Gata4*, and *Tbx5*), and mature CMs (*Myh7* and *Myh6*; Supplementary Figure 1).

We found that CM commitment and maturation are accompanied by a change in expression of several epigenetic enzyme genes already reported to be important for cardiac development (e.g., *Jmjd3*,¹² *Ezh2*^{13,14}), as well as of HATs, HDACs, HMTs, HDMS, and DNMTs not yet studied in the heart (Figures 1a and c). In particular, *Dot1L*, which encodes the only known H3K79 HMT, underwent a major differential change in expression during CM differentiation and maturation. This finding was confirmed with conventional real-time qPCR and immunoblotting (Figures 1b and d; Supplementary Figure 1).

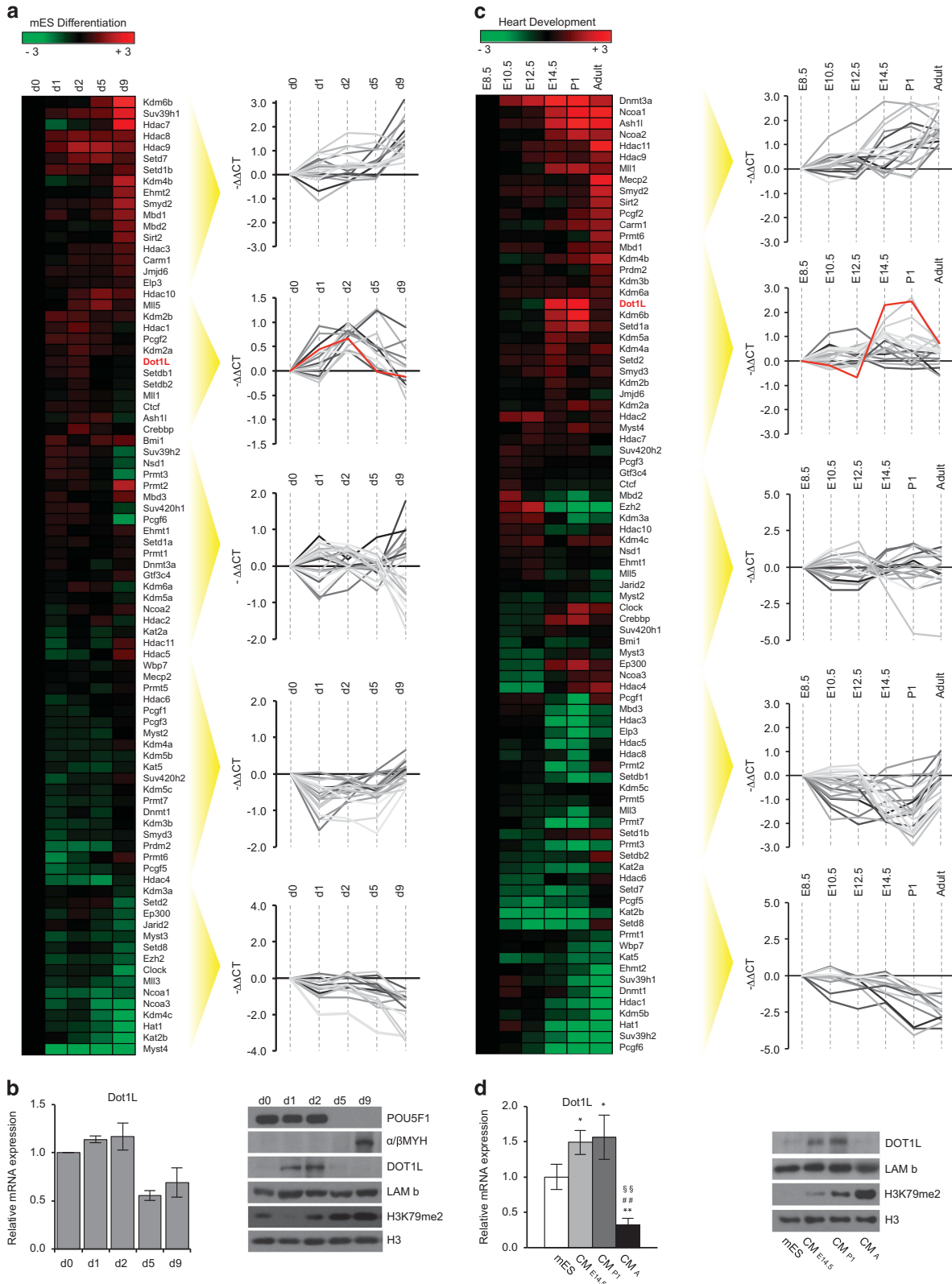
In addition to the dynamic changes in *Dot1L* expression, we observed increased deposition of H3K79me2 subsequent to the *Dot1L* expression peak (Figures 1b and d right panels, and Supplementary Figure 1). Interestingly, this epigenetic mark seems to be very stable, as it is abundantly found in adult CMs (C_{Ma}) even if these cells do not express high levels of DOT1L. From these findings, we inferred that DOT1L is a key enzyme for CM differentiation and maturation.

H3K79me2 regulates gene expression during CM differentiation and maturation. To gain insight into the role of DOT1L in cardiomyogenesis, we analyzed the genome-wide distribution of H3K79me2 in CMs during differentiation. To this end, we performed chromatin immunoprecipitation (ChIP) followed by massive parallel DNA sequencing (seq) for H3K79me2 and three other dynamic epigenetic marks of differentiation (i.e., H3K4me3, H3K27me3, and H3K9me3) on purified CMs isolated at two different stages of maturation: 1-day-old pup CMs (C_{Mp}), which are mononucleated cells that still retain proliferative capacity, and 2-month-old C_{Ma},¹⁵ which are mainly binucleated cells and have lost their ability to divide.¹⁶ In addition, ChIP-seq data sets of mES from the Gene Expression Omnibus (GEO) database were used as reference points for cells in a pluripotent, undifferentiated state (GSE12241 and GSE11724).^{17,18}

Comparison of the methylation profiles across the genome in mES, C_{Mp}, and C_{Ma} indicated that H3K4me3, H3K27me3, and H3K79me2 became redistributed within various genomic regions during differentiation, whereas there was no significant change in the distribution of H3K9me3 (Figure 2a). In particular, H3K79me2 became increasingly more present inside the gene body and less associated with the transcription start site (TSS), whereas H3K4me3 shifted from the overlap start region in mES to a region more proximal to the TSS, and H3K27me3 decreased around the TSS, becoming localized mainly up and downstream of the gene body in C_{Mp} and C_{Ma}. These findings indicate that the histone methylation profile is dynamic during CM differentiation.

To identify the genomic regions where these histone methylation marks became redistributed during cardiac differentiation, we analyzed the genomic distribution of each mark with two algorithms (MACS (Model-based Analysis of ChIP-seq) and SICER (Spatial Clustering Approach for the Identification of ChIP-Enriched Regions)) in order to minimize peak calling bias; peak data sets were then compared with DiffBind.¹⁹ Hierarchical clustering analysis using affinity data revealed that H3K79me2, like the other histone marks analyzed, was redistributed within the genome not only during CM specification (mES versus C_{Mp}) but also during CM maturation (C_{Mp} versus C_{Ma}; Figure 2b). Interestingly, H3K79me2 in mES and C_{Mp} clustered with the clade containing H3K4me3, forming a branch with reduced similarity to the profile of H3K79me2 in C_{Ma}. This suggests that these two histone modifications mark similar regions in mES and C_{Mp} but not in C_{Ma}, and indicates that CM maturation is associated with redistribution of H3K79me2. This finding was corroborated by Principal Component Analysis, which indicated a higher number of differential binding sites in mES versus C_{Ma} than in mES versus C_{Mp} (Supplementary Figure 2). For each histone modification analyzed, we report all significant differential sites in mES versus C_{Mp}, mES versus C_{Ma}, and C_{Mp} versus C_{Ma} in Supplementary Table 1. Gene Ontology analysis of the genes associated with these sites revealed that these histone methylation marks became redistributed at different functional classes of genes during CM differentiation: in mES, H3K79me2, and H3K4me3 were modulated at putative regulatory regions involved in RNA metabolism; in CMs, these modifications were enriched at genes encoding CM structural proteins and were involved in cardiac phenotype specification (myocardial fiber morphology and contractility), whereas H3K27me3 was enriched at genes involved in bone and nervous system formation (Supplementary Figure 3).

Figure 1 Genome-wide expression profiling of epigenetic enzyme genes reveals a potential role for disruptor of telomeric silencing 1-like (DOT1L) in cardiomyocyte (CM) differentiation and maturation. (a) Custom qRT-PCR array card assessing the relative expression of 85 epigenetic enzyme genes at different stages of cardiac differentiation of mouse embryonic stem cells (mES; days 1, 2, 5, and 9). Data are presented as $-\Delta\Delta Ct$ relative to expression at day 0. On the left, heatmap of differential expression levels; on the right, graphs highlighting five groups of differentially expressed genes. (b) On the left, relative qRT-PCR of *Dot1L* mRNA expression at day 0, 1, 2, 5, and 9 of mES differentiation; mRNA levels are normalized to 18 s and expressed as mean \pm S.D. On the right, representative Western blot of mES cells at 0, 1, 2, 5, and 9 days of differentiation. The expression levels for POU5F1, α/β MYH, and DOT1L are normalized to that of lamin B (LAM b), whereas the expression of H3K79me2 is normalized to that of unmodified H3. (c) Custom qRT-PCR array card assessing the relative expression of 85 epigenetic enzyme genes in mouse cardiac cells at different stages of development (embryos at E8.5, E10.5, E12.5, E14.5; P1 pups, and adult). Data are presented as $-\Delta\Delta Ct$ relative to E8.5 expression. The mean of three biological replicates is shown. On the left, heatmap of differential expression levels; on the right, graphs highlighting five groups of differentially expressed genes. (d) On the left, qRT-PCR of relative mRNA expression of *Dot1L* in mES, cardiomyocytes from embryonic day E14.5 (C_{Me14.5}), and cardiomyocytes from post-natal day 1 (C_{Mp1}) and 2-month-old adult (C_{Ma}) mice. Graphs of mean \pm S.D. of three biological replicates. * $P < 0.05$ versus mES; ** $P < 0.01$ versus mES; *** $P < 0.01$ versus C_{Me}; \$\$\$ $P < 0.01$ versus C_{Mp}. On the right, Western blot analysis for the expression of DOT1L (normalized to LAM b) and H3K79me2 (normalized to unmodified H3)



We then studied in detail the epigenetic profile at genes important for cardiogenic differentiation and maturation: *Pou5F1* and *Nanog* (genes encoding pluripotency transcription factors), *Nkx2.5* and *Gata4* (which encode

transcription factors involved in early CM development), *Nppa*, *Myl7*, and *Myl2* (which are involved in chamber specification), *Myh7*, *Myh6*, and *Tnni3* (genes regulated during cardiac maturation), and *Kcnb1*, *Kcnd2*, and *Tuba8* (genes important

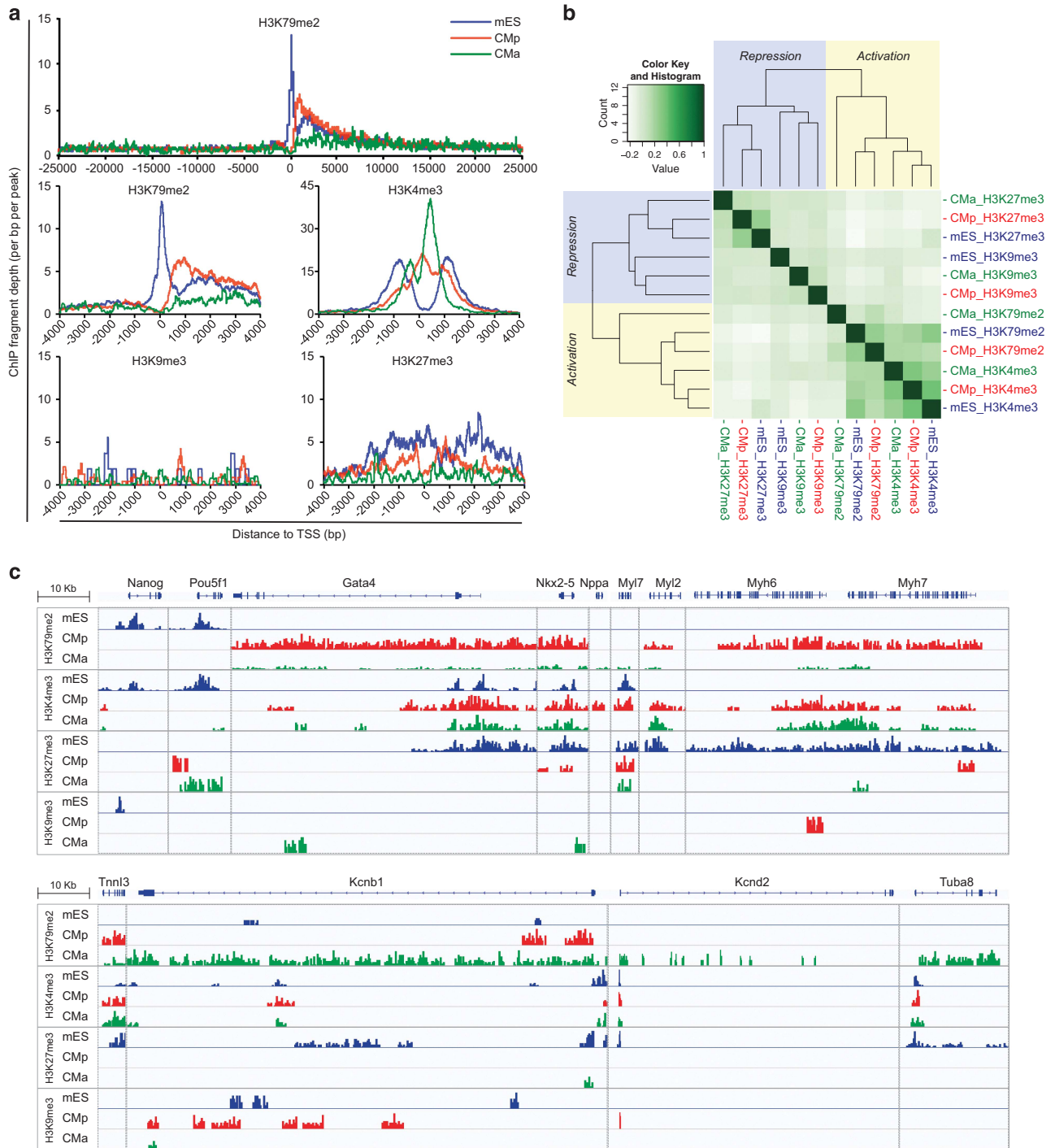


Figure 2 Genome-wide analysis of H3K79me2 and other histone methylation marks in cardiomyocytes (CM). (a–c) chromatin immunoprecipitation sequencing (ChIP-seq) analysis of H3K79me2, H3K4me3, H3K27me3, and H3K9me3 in mouse embryonic stem cells (mES) and CM from post-natal day 1 (CMp) and 2-month-old adult (CMa) mice. (a) Enrichment plots for the histone modifications studied: for H3K79me2, enrichment is shown within the region ± 25 kb around the annotated transcription start site (TSS); for H3K79me2, H3K4me3, H3K27me3, and H3K9me3, enrichments are shown within a region ± 4 kb around the annotated TSSs. Each graph gives the distribution in mES (blue line), CMp (red line), and CMa (green line). Data were generated with HOMER v4.5 software (<http://homer.salk.edu/homer/>; La Jolla, CA, USA).³² (b) correlation heatmap using affinity (read count) data generated with DiffBind R package.¹⁹ (c) Genome browser representations of enrichment profiles of H3K79me2, H3K4me3, H3K27me3, and H3K9me3 in mES (blue), CMp (red), and CMa (green) are shown at key genes of pluripotency (*Nanog*, *Pou5f1*), cardiac transcription regulation (*Gata4*, *Nkx2-5*), atrial/ventricular characterization (*Nppa*, *Myl7*, *Myl2*), cardiomyocyte specification (*Myh6*, *Myh7*, *Tnni3*), and maturation or function (*Kcnb1*, *Kcnd2*, *Tuba8*). The peak heights are set to correspond to the maximum enriched peak in the region for each mark

for CM structure and function). We found that the profile of H3K79me2, as well as of the other histone methylation marks analyzed, changed during cardiac differentiation at many of these loci: H3K79me2 and H3K4me3 were present at *Nanog* and *Pou5F1* in mES but not in CMp or CMa; H3K27me3 was acquired at *Pou5F1* in CMp and CMa; and the cardiac-specific genes *Gata4*, *Nkx2.5*, *Myh7*, *Myh6*, *TnnI3*, and *Myl2* acquired H3K79me2 and lost H3K27me3 in CMs, with H3K79me2 enrichment being more significant in CMp. There was a similar loss of H3K27me3 at *Kcnb1*, *Kcnd2*, and *Tuba8* in CMs compared with mES, but these loci acquired H3K79me2 in the transition from CMp to CMa. In contrast, we did not observe any significant change in the distribution of H3K79me2 at *Nppa* and *Myl7*; however, there was a redistribution of the bivalent domain H3K4me3/H3K27me3 (Figure 2c). These findings were validated by ChIP-qPCR assays (Supplementary Figure 4).

To investigate whether changes in the genomic distribution of H3K79me2, H3K4me3, H3K27me3, and H3K9me3 explain the transcriptional changes occurring during CM differentiation, we cross-compared the ChIP-seq libraries from mES, CMp, and CMa with the transcriptional profile of these three cell populations. To this end, we first assessed differential gene expression in mES, CMp, and CMa with an Illumina microarray. We identified 4607 significantly modulated genes in mES versus CMp, 5258 genes in mES versus CMa, and 2163 genes differentially expressed between CMp and CMa (P -value < 0.001; Supplementary Figure 5 and Supplementary Table 2). Pearson correlation analysis revealed that the transcriptional profile correlated directly with H3K4me3 and H3K79me2, and inversely with H3K27me3, whereas no significant correlation was found with H3K9me3 (Figures 3a and b).

To unveil biological pathways regulated by H3K79me2, we analyzed the network of genes displaying changes in H3K79me2 associated with a significant differential transcriptional profile. This analysis revealed that depending on the stage of differentiation, H3K79me2 regulated different transcriptional networks: in the transition from mES to CMs, the most modulated genes belonged to networks involved in cellular growth and assembly that lead to developmental disorders, whereas in CMp versus CMa they were implicated in cell cycle and DNA replication (e.g., cyclin B and AURKB; Figure 3c). Analysis of the differentially expressed genes that acquired or lost methylation on the bivalent domain K4/K27 identified similar networks but with different genes being affected (Supplementary Figure 6). This supported the notion that a specific pattern of histone methylation is used to mark functionally distinct genes having a similar pattern of expression, and that H3K79me2 plays a central role in the regulation of the transcriptional program involved in developing cardiac cells.

Dynamic changes of DOT1L-mediated H3K79 methylation. To examine whether the activity of DOT1L is required for the early events of cardiomyogenic differentiation *in vitro*, we used a lentiviral vector-mediated approach to knock down *Dot1L* in our model of *in vitro* CM differentiation. mESs were transduced with lentiviral vectors carrying either short hairpin (sh)RNA against *Dot1L* (shDOT1L) or scrambled sequences

(sh-Ctrl), and then assessed for expression of pluripotency and CM differentiation markers. We found that *Dot1L* knockdown affected mesodermal and cardiac commitment, with shDOT1L cells exhibiting reduced and delayed expression of *Gsc*, *Eomes*, *Mesp1*, and *Brachyury* (at days 3, 4 and 5), and *Myl2*, *Myh7*, *Myh6*, and *TnnI3* (at days 7 and 9) when compared with sh-Ctrl-transduced mES (Figure 4a). In contrast, there was no effect on markers of pluripotency (*Nanog* and *Pou5F1*), endoderm (*Sox17* and *Gata6*) or ectoderm (*Nestin*), or on cardiac-specific transcription factors (*Gata4*, *Nkx2.5*, *Isl1*, *Mef2c*) or atrial-specific genes (*Myl4* and *Myl7*; Supplementary Figure 7). Furthermore, *Dot1L* knockdown caused a decrement in the efficiency of cardiac differentiation, as assessed by flow cytometry quantification of TNNI3-positive cells (Figure 4b).

Western blot analysis of relative protein levels validated that the knockdown of *Dot1L* mRNA at day 1 and day 2 of differentiation (Figure 4a) directly translated into a reduced DOT1L protein level at the same time-points (Figure 4c); as expected, levels of H3K79me2 were also decreased. This difference was attenuated at later days of differentiation, when lentiviral vector-mediated knockdown became less efficient (Figures 4a and c). Consistent with the gene expression analysis and the assessment of CM differentiation efficiency, the initial reduction in DOT1L and H3K79me2 resulted in a lower expression of α/β -myosin heavy chains later on (day 9).

We then analyzed how the reduction in DOT1L expression affected the profile of H3K79me2 at genes regulated during CM differentiation. H3K79me2 levels at the loci of *Nanog* and *Pou5F1* were high at day 0 and decreased afterwards, with shRNA-mediated downregulation of *Dot1L* having no effect on this profile (Figure 4d). Markers of mesodermal precursors (*Brachyury* and *Eomes*) had very high levels of H3K79me2 at their loci from day 5 onwards, whereas at *Myh7*, *Myh6*, and *TnnI3* the level of H3K79me2 peaked at day 9. Of note, H3K79me2 was significantly reduced at these genes on day 5 in mES transduced with sh-DOT1L when compared with sh-Ctrl mES. The decrease was not observed on day 9 of differentiation, a time at which we found a major effect on expression of *Myh7*, *Myh6*, and *TnnI3* mRNAs. This result suggests that the presence of H3K79me2 at these genes during an early phase of differentiation (days 2–5) is required for their expression. Thus, H3K79me2 could be a pre-activation mark for key differentiation genes.

Altogether, these findings demonstrate that DOT1L is required for CM differentiation by regulating the expression of specific genes.

Discussion

We demonstrate here that DOT1L-mediated H3K79me2 modification is necessary for the definition of the epigenetic signature of CM differentiation and maturation. Previous reports have proposed a role for DOT1L in cardiac development: constitutive disruption of *Dot1L* in mice resulted in growth impairment, angiogenesis defects of the yolk sac, cardiac dilatation, and lethality at E10.5,²⁰ whereas late-stage cardiac-specific deletion of *Dot1L* in mice produced a severe cardiac phenotype resembling that of dilated cardiomyopathy.²¹ However, the role of DOT1L-mediated

H3K79me2 modification in CM differentiation and maturation remains poorly studied.

Genome-wide analysis of the distribution of H3K79me2 during CM differentiation revealed that this mark regulates distinct transcriptional networks at different stages of differentiation, i.e., networks of genes involved in cellular growth and assembly during CM differentiation (i.e., mES → Cmp), and networks of cell cycle and DNA replication genes during CM maturation (i.e., Cmp → Cma). This strongly suggests that di-methylation at H3K79 is a requirement for the establishment of the gene expression program of CMs during cardiogenesis, tailoring the program for stage-specific biological processes.

We also found that H3K79me2 was enriched in neonatal CMs at genes that became activated only in later stages of development, such as *Myh6* and *Tnni*. This suggests that H3K79me2 is involved in a pre-activating mechanism in which the histone modification is used during early lineage commitment to mark the genes for subsequent activation during maturational phases. This hypothesis is supported by our *in vitro* knockdown assessment of the role of DOT1L in CM differentiation. These experiments demonstrated how the activity of this enzyme is necessary for defining the epigenetic signature driving cells toward cardiac commitment, as a decrease in H3K79me2 at *Myh6*, *Myh7*, and *Tnni3* during the initial days of differentiation impaired expression of these genes at subsequent stages of differentiation.

To date, DOT1L is the only known methyltransferase catalyzing mono-, di-, and tri-methylation of H3K79.^{22–25} We and others have demonstrated that its expression is elevated during cardiac development but reduced in Cma. However, H3K79me2 is abundantly present in Cma, with the transition from Cmp to Cma characterized by its redistribution across the genome.

A critical mechanism for differentiation is maintenance of a proper balance in the methylation levels of H3K4 and H3K27 at the promoter regions of poised developmental genes: this phenomenon governs the highly dynamic regulation of gene expression that defines a differentiating cell.^{26,27} Two recent reports have shown that the bivalent K4/K27 domain controls gene expression during differentiation of embryonic stem cells into cardiovascular cells.^{10,11} Our results confirm this *in vivo* and reveal that the activation of key cardiac genes (e.g., *Myh6* and *Myh7*) is driven not only by the loss of the repressive H3K27me3 mark, but also by the acquisition of the activating H3K79me2 modification. Therefore, it is possible that H3K79me2 cooperates with the bivalent domain

K4/K27 in the regulation of the expression of key CM differentiation genes.

In addition, we found that the epigenetic profile of neonatal-stage CMs was closer to that of mES than to that of Cma, suggesting that the recently reported¹⁶ regenerative potential of neonatal mouse heart could be linked to this less mature epigenetic landscape. In fact, it is conceivable that in Cmp a chromatin structure similar to that of mES could be permissive to the binding of transcription factors driving an embryonic-like gene expression profile that would ultimately lead to a greater regenerative potential.

Besides demonstrating the importance of DOT1L in the epigenetic regulation of cell-fate acquisition processes during cardiac development, and highlighting the role of H3K79 methylation in stabilizing the signature of the cardiac gene expression program, the high-resolution methylation maps of the CM genome reported here could lead to the identification of novel genes and transcriptional regulatory networks involved in cardiac differentiation.

Materials and Methods

Standard techniques were used for purification and culture of mouse CMs, immunoblotting, RNA extraction, FACS analysis, and viral particle production and use (see Supplementary Methods). Mouse ES culture and cardiogenic differentiation were carried out as previously reported.²⁸

ChIP assay. Genome-wide localization of histone modifications (H3K4me3, H3K27me3, H3K79me2, and H3K9me3) for each stage was determined via ChIP followed by high-throughput seq with SoLiD 5500 (Life Technologies, Carlsbad, CA, USA) carried out as described elsewhere.¹⁵ Briefly, 5×10^6 cells were used for each immunoprecipitation. Cells were cross-linked for 10 min at room temperature using 1% formaldehyde. Cross-linking was quenched by adding glycine to a final concentration of 0.125 M. The cells were then collected, resuspended in lysis buffer (5 mM PIPES pH 8, 85 mM KCl, 0.5% NP40, and protease inhibitors), and incubated on ice for 15 min before proceeding with sonication to generate 200–400 bp fragments. The efficiency of sonication was assessed with agarose gel electrophoresis. Chromatin samples were pre-cleared for 1 h with protein-G beads and then immunoprecipitated overnight at 4 °C with the following specific antibodies: anti-H3 dimethyl Lys79 (abcam ab3594, Cambridge, UK), anti-H3 trimethyl Lys4 (active motif 39159), anti-H3 trimethyl Lys 27 (Millipore-Upstate 07–449, Billerica, MA, USA), anti-H3 trimethyl Lys9 (Millipore-Upstate 07–442), anti-H3 (abcam ab1791), and rabbit IgG (Millipore-Upstate 12–370). After incubation, the immunocomplexes were bound to protein-G beads for 2 h and subsequently washed with low-salt wash buffer (0.1% SDS, 2 mM EDTA, 20 mM Tris HCl pH 8, 1% Triton X-100, 150 mM NaCl, and protease inhibitors), high-salt wash buffer (0.1% SDS, 2 mM EDTA, 20 mM Tris HCl pH 8, 1% Triton X-100, 500 mM NaCl, and protease inhibitors), and TE buffer. Immunocomplexes were eluted in elution buffer (1% SDS, 100 mM NaHCO₃, and protease inhibitors), and cross-linking reverted at 65 °C overnight. Samples were treated with proteinase K, extracted with phenol/chloroform, and precipitated with ethanol. Purified DNA was sequenced with SoLiD

Figure 3 The epigenetic code associated with histone methylation correlates with the transcriptional program of cardiomyocytes (CM). (a) Expression data for genes differentially marked by H3K79me2, H3K4me3, H3K27me3, and H3K9me3 in mouse embryonic stem cells (mES) and CM from post-natal day 1 (Cmp) and from 2-month-old adult (Cma) mice. Genes are ordered by magnitude of occupancy of differential marks; relative gene expression is shown. Samples with higher than average expression are represented in red, and samples with lower than average expression are represented in green. Results from three different comparative analyses are shown: mES versus Cmp (left), mES versus Cma (middle), and Cmp versus Cma (right). (b) Pearson correlation analysis between fold change of chromatin immunoprecipitation sequencing (ChIP-seq) values (ChIP-seq FC) and logarithmic-scale fold change in gene expression (Illumina log FC). The correlation index R for each comparison is indicated inside each graph. The red lines indicate direct and inverse correlations between enrichment of the mark and the level of gene expression. Results from three different comparative analyses are shown: mES versus Cmp (left), mES versus Cma (middle), and Cmp versus Cma (right) for H3K79me2, H3K4me3, H3K27me3, and H3K9me3. (c) Network analysis of direct relationships among genes undergoing acquisition of H3K79me2 and transcriptional activation. The analysis was performed using Ingenuity Pathway Analysis software (v. 8.5, Ingenuity Systems, www.ingenuity.com). Networks are displayed graphically as nodes and edges; different shapes indicate different gene functions. Results from three different comparative analyses are shown: mES versus Cmp (top), mES versus Cma (middle), and Cmp versus Cma (bottom). Green indicates gene downregulation and red indicates gene upregulation in Cmp (mES versus Cmp), Cma (mES versus Cma), or Cma (Cmp versus Cma), with a more intense shade indicating a greater change

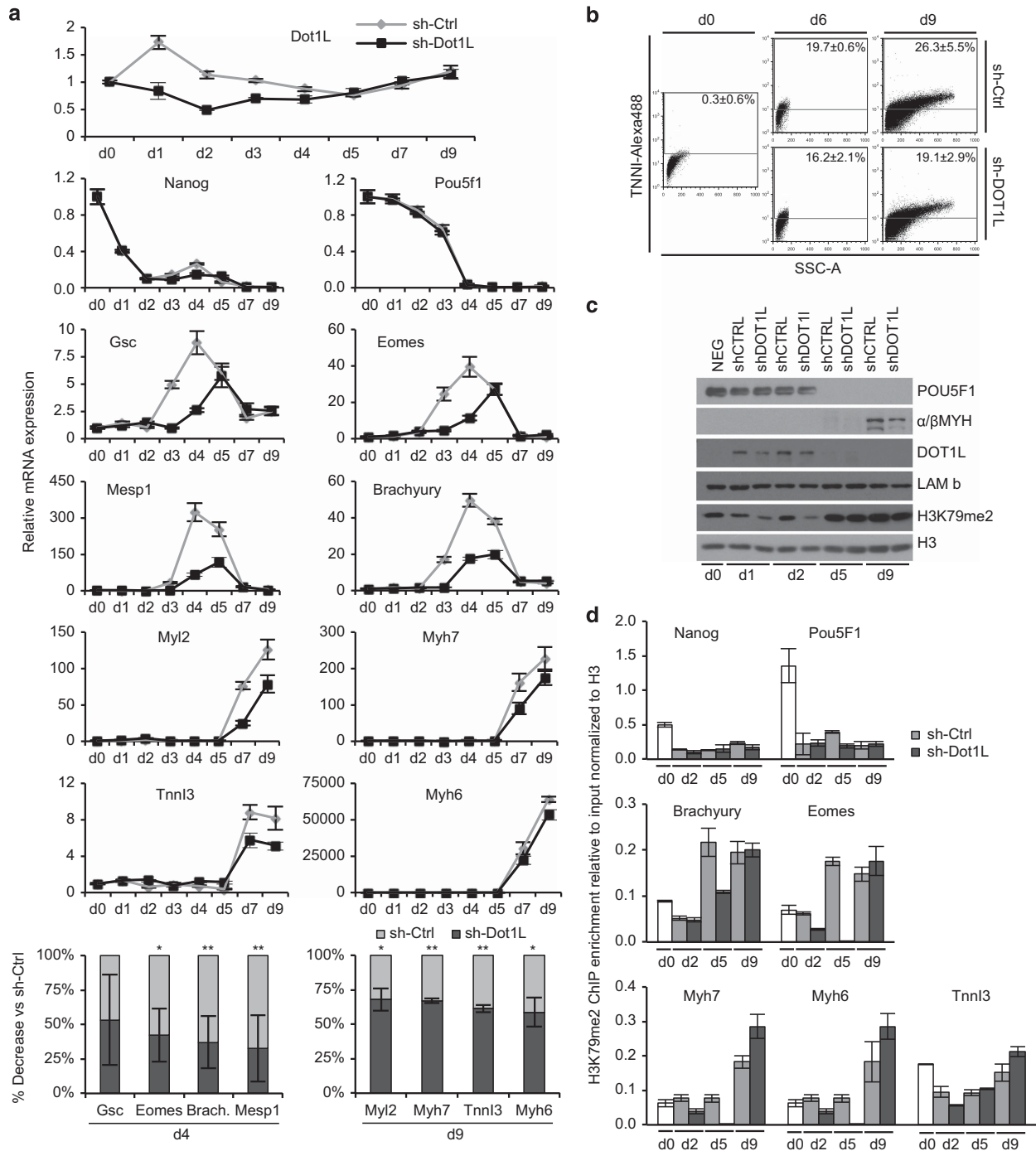


Figure 4 Disruptor of telomeric silencing 1-like (DOT1L)-mediated di-methylation (me) of H3K79 in cardiogenic differentiation of mouse embryonic stem cells (mES). **(a)** mRNA levels of *Dot1L*, pluripotency, mesodermal, and cardiomyocyte (CM)-specific genes in undifferentiated (day 0) mES and after days 1, 2, 3, 4, 5, 7, and 9 of differentiation in control (sh-Ctrl, gray line) and *Dot1L*-knockdown (sh-Dot1L, black line) cells. Levels are normalized to 18 s or *Gapdh*, expressed as the mean \pm S.D. of experimental replicates, and plotted as relative mRNA expression (top). Percentage decrease of *Gsc*, *Eomes*, *Mesp1*, and *Brachyury* expression at day 4 and *Myl2*, *Myh7*, *Tnn3*, and *Myh6* at day 9 in sh-Dot1L versus sh-Ctrl transduced cells. Mean \pm S.D. of three independent experiments is plotted (** $P < 0.01$; * $P < 0.05$ sh-Dot1L versus sh-Ctrl; bottom). **(b)** FACS analysis of the percentage of cells expressing TNNI at days 0, 6, and 9 of differentiation in control (sh-Ctrl) and *Dot1L*-knockdown (sh-DOT1L) cells. The expression levels for POU5F1, α/β MYH, and DOT1L are normalized to that of lamin B (LAM b), whereas the expression levels of H3K79me2 are normalized to that of unmodified H3. **(c)** Protein levels, assessed by Western blotting, at days 0, 1, 2, 5, and 9 of differentiation in control (sh-Ctrl) and *Dot1L*-knockdown (sh-DOT1L) mES. The expression levels for POU5F1, α/β MYH, and DOT1L are normalized to that of lamin B (LAM b), whereas the expression levels of H3K79me2 are normalized to that of unmodified H3. **(d)** Chromatin immunoprecipitation (ChIP) assay of H3K79me2 binding to the gene locus of *Nanog*, *Pou5F1*, *Brachyury*, *Eomes*, *Myh7*, *Myh6*, and *Tnn3* in undifferentiated (day 0) mES (white bars) and at days 2, 5, and 9 after differentiation in control (sh-Ctrl, gray bars) and *Dot1L*-knockdown (sh-Dot1L, black bars) cells. Levels were determined by qPCR and are expressed as fold change to the input and relative to H3.

5500 (Life Technologies) by Genomnia Srl (Lainate, Milan, Italy) or evaluated by qPCR on an ABI 7900HT with SYBR green PCR master mix (Applied Biosystem, Carlsbad, CA, USA) using specific primers designed close to the promoter region and the gene TSS: A = -1000 bp/-500 bp from TSS; B = +500 bp/+1000bp from TSS; C = +3500 bp/+4000 bp from TSS. Values obtained were normalized to the input and to the H3 content. The sequences of the primers used for ChIP-qPCR are available upon request.

ChIP-seq analysis pipeline. ChIP-seq data analysis was performed using distinct bioinformatics software. For H3K4me3, H3K27me3, and H3K9me3 modifications in mES, bed files were downloaded from the GEO (H3K4me3: GSM307618; H3K27me3: GSM307619; H3K9me3: GSM307621). Genome coordinates were converted from mm8 to mm9 mouse reference genome using Batch Coordinate Conversion (LiftOver) created by the UCSC Genome Bioinformatics Group. For the H3K79me2 modification in mES, raw seq reads were downloaded from GEO (GSM307150-GSM307151) and mapped to the mouse genome (version mm9) using BOWTIE (version 0.12.8).

To profile histone modifications in CMs, seq reads were mapped to the mouse genome (version mm9) using BOWTIE (version 0.12.8). Uniquely mapped reads with no more than two mismatches were used for binding peak detection. To identify peaks, two peak-calling software were used: MACS (Model-based Analysis of ChIP-seq)²⁹ and SICER (Spatial Clustering Approach for the Identification of ChIP-Enriched Regions).³⁰ Both the software detected binding peaks by comparing IP and input control. We used the following parameters: MACS: effective genome size = 1.87e+09, band width = 300, model fold = 5.30, *P*-value cutoff = 1.00e-05; SICER: windows size = 200, gap size = 600, redundancy threshold = 1, FDR = 0.05. Occupancy analysis and differential binding affinity analysis were assessed with the R Bioconductor package DiffBind (v. 1.2.0).¹⁹ The final set of binding peaks contained those that were called by both softwares. Genomic annotation of the peaks identified from the ChIP-seq data were performed using the R Bioconductor package ChIPpeakAnno (v. 2.4.0).³¹ A summary of the genes associated with H3K79me2-, H3K4me3-, H3K27me3-, and H3K9me3-occupied regions is provided in Supplementary Table S1. The complete ChIP-seq data sets are available from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>), under accession number GSE45174.

Density profiles of chromatin marks at promoters and gene bodies were computed with HOMER v.4.5 (<http://homer.salk.edu/homer/>; La Jolla, CA, USA) (Empirical Motif Optimizer).³² Functional annotation of differentially enriched modifications was performed using GREAT.³³ Genome browser representations of H3K79me2-, H3K4me3-, H3K27me3-, and H3K9me3-enrichment profiles were obtained with IGV 2.1 visualization software.³⁴

Gene expression microarray analysis pipeline. BeadChips were scanned with the Illumina iScan system. Raw data were background-subtracted and normalized using the quantile normalization method of the R Bioconductor package lumi.³⁵ Normalized data were filtered for genes with significant expression levels compared with negative control beads. Differentially expressed genes (adjusted *P*-value < 0.001) were identified with the Illumina *t*-test error model (limma R Bioconductor package).³⁶ The transcript with the highest median expression was selected to represent the expression of the gene if it was represented with several transcripts. Functional annotation of significant genes identified by microarray analysis was searched with the web-accessible program Database for Annotation, Visualization and Integrated Discovery (DAVID) version 2009, National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH; <http://david.abcc.ncifcrf.gov/>).³⁷ The complete Illumina gene expression data sets are available from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>), under accession number GSE44829.

Correlation of ChIP-seq and gene expression data. Genes were ordered by magnitude of differential mark occupancy, and heatmaps of relative gene expression generated. Genes with higher than average expression are marked in red, and those with lower than average expression are marked in green (scale in standard deviations). Scatter plots were produced to see how the expression level agreed with degree of histone modification. The Pearson correlation coefficients between histone modifications and the expression level of genes were calculated using R. Network analysis considering direct and indirect relationship among genes was performed using Ingenuity Pathway Analysis software (v. 8.5, Ingenuity Systems, www.ingenuity.com).

Statistical analysis. Data are presented as mean ± S.D. *P*-values were determined by two-tailed *t*-test. *P* < 0.05 was considered statistically significant.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements. This work was supported by an 'Advanced' European Research Council (ERC) grant (Cardioepigen), grants from Fondation LeDucq and from MIUR (PRIN #2010RNXM9C_004) to CG, and from the Progetto Bandiera Epigenomica (EPIGEN) and the Progetto Bandiera Invecchiamento (Ageing Project) to PR.

Author contributions

PC and PR designed the experiments, carried out research and wrote the manuscript; GC, RF, RR, PED, LSL and BC carried out research; KP, GA and SGG performed bioinformatics analyses; LMVG wrote the manuscript; CG supervised the whole research. All authors discussed the results and commented on the manuscript.

- Evans SM, Yelon D, Conlon FL, Kirby ML. Myocardial lineage development. *Circ Res* 2010; **107**: 1428–1444.
- Garry DJ, Olson EN. A common progenitor at the heart of development. *Cell* 2006; **127**: 1101–1104.
- Martin-Puig S, Wang Z, Chien KR. Lives of a heart cell: tracing the origins of cardiac progenitors. *Cell Stem Cell* 2008; **2**: 320–331.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J *et al*. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006; **125**: 315–326.
- Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 2007; **130**: 77–88.
- Meissner A. Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol* 2010; **28**: 1079–1088.
- Bruneau BG. The developmental genetics of congenital heart disease. *Nature* 2008; **451**: 943–948.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z *et al*. High-resolution profiling of histone methylations in the human genome. *Cell* 2007; **129**: 823–837.
- Kouzarides T. Chromatin modifications and their function. *Cell* 2007; **128**: 693–705.
- Wamstad JA, Alexander JM, Truty RM, Shrikumar A, Li F, Eilertson KE *et al*. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* 2012; **151**: 206–220.
- Paige SL, Thomas S, Stoick-Cooper CL, Wang H, Maves L, Sandstrom R *et al*. A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. *Cell* 2012; **151**: 221–232.
- Ohtani K, Zhao C, Dobrev G, Manavski Y, Kluge B, Braun T *et al*. Jmjd3 controls mesodermal and cardiovascular differentiation of embryonic stem cells. *Circ Res* 2013; **113**: 856–862.
- Delgado-Olguin P, Huang Y, Li X, Christodoulou D, Seidman CE, Seidman JG *et al*. Epigenetic repression of cardiac progenitor gene expression by Ezh2 is required for postnatal cardiac homeostasis. *Nat Genet* 2012; **44**: 343–347.
- He A, Ma Q, Cao J, von Gise A, Zhou P, Xie H *et al*. Polycomb repressive complex 2 regulates normal development of the mouse heart. *Circ Res* 2012; **110**: 406–415.
- Papait R, Cattaneo P, Kunderfranco P, Greco C, Carullo P, Guffanti A *et al*. Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. *Proc Natl Acad Sci USA* 2013; **110**: 20164–20169.
- Porrello ER, Mahmoud AI, Simpson E, Hill JA, Richardson JA, Olson EN *et al*. Transient regenerative potential of the neonatal mouse heart. *Science* 2011; **331**: 1078–1080.
- Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S *et al*. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 2008; **134**: 521–533.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G *et al*. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007; **448**: 553–560.
- Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ *et al*. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 2012; **481**: 389–393.
- Jones B, Su H, Bhat A, Lei H, Bajko J, Hevi S *et al*. The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet* 2008; **4**: e1000190.
- Nguyen AT, Xiao B, Neppel RL, Kallin EM, Li J, Chen T *et al*. DOT1L regulates dystrophin expression and is critical for cardiac function. *Genes Dev* 2011; **25**: 263–274.
- Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K *et al*. Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr Biol* 2002; **12**: 1052–1058.

23. van Leeuwen F, Gafken PR, Gottschling DE. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* 2002; **109**: 745–756.
24. Shanower GA, Muller M, Blanton JL, Honti V, Gyurkovics H, Schedl P. Characterization of the grappa gene, the *Drosophila* histone H3 lysine 79 methyltransferase. *Genetics* 2005; **169**: 173–184.
25. Nguyen AT, Zhang Y. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev* 2011; **25**: 1345–1358.
26. Song K, Nam YJ, Luo X, Qi X, Tan W, Huang GN *et al*. Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature* 2012; **485**: 599–604.
27. Epstein JA, Franklin H. Epstein lecture. Cardiac development and implications for heart disease. *New Engl J Med* 2010; **363**: 1638–1647.
28. Keller GM. *In vitro* differentiation of embryonic stem cells. *Curr Opin Cell Biol* 1995; **7**: 862–869.
29. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc* 2012; **7**: 1728–1740.
30. Zang C, Schonnes DE, Zeng C, Cui K, Zhao K, Peng W. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* 2009; **25**: 1952–1958.
31. Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS *et al*. ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 2010; **11**: 237.
32. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P *et al*. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010; **38**: 576–589.
33. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB *et al*. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010; **28**: 495–501.
34. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G *et al*. Integrative genomics viewer. *Nat Biotechnol* 2011; **29**: 24–26.
35. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008; **24**: 1547–1548.
36. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 2005; **21**: 2067–2075.
37. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**: 44–57.

Supplementary Information accompanies this paper on Cell Death and Differentiation website (<http://www.nature.com/cdd>)