EBioMedicine 4 (2016) 74–85

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Mapping Human Pluripotent-to-Cardiomyocyte Differentiation: Methylomes, Transcriptomes, and Exon DNA Methylation "Memories"

Joshua D. Tompkins [a,e,*], Marc Jung [a,e], Chang-yi Chen [b,c,e], Ziguang Lin [b,c,e], Jingjing Ye [b,c,e], Swetha Godatha [a,e], Elizabeth Lizhar [a,e], Xiwei Wu [d,e], David Hsu [b,c,e], Larry A. Couture [b,c,e], Arthur D. Riggs [a,e]

[a] Department of Diabetes Complications and Metabolism, Duarte, CA 91010, USA
[b] Center for Biomedicine and Genetics, Duarte, CA 91010, USA
[c] Sylvia R. and Isador A. Deutch Center for Applied Technology Development, Duarte, CA 91010, USA
[d] Biomedical Informatics Core, Duarte, CA 91010, USA
[e] Beckman Research Institute at City of Hope National Medical Center, Duarte, CA 91010, USA

ARTICLE INFO

Article history:
Received 11 August 2015
Received in revised form 5 January 2016
Accepted 15 January 2016
Available online 19 January 2016

Keywords:
Human embryonic stem cells
Pluripotent
Cardiomyocytes
Differentiation
Cardiomyogenesis
DNA methylation
Long non-coding RNA
lncRNA
Epigenetic
Methylome
Transcriptome
Mesoderm
Good manufacturing practice, GMP, epigenetic memory, WNT, hedgehog, transforming growth factor, ROR2, PDGFRα, demethylation, TET, TDG, HOX, TBOX

ABSTRACT

The directed differentiation of human cardiomyocytes (CMs) from pluripotent cells provides an invaluable model for understanding mechanisms of cell fate determination and offers considerable promise in cardiac regenerative medicine. Here, we utilize a human embryonic stem cell suspension bank, produced according to a good manufacturing practice, to generate CMs using a fully defined and small molecule-based differentiation strategy. Primitive and cardiac mesoderm purification was used to remove non-committing and multi-lineage populations and this significantly aided the identification of key transcription factors, lncRNAs, and essential signaling pathways that define cardiomyogenesis. Global methylation profiles reflect CM development and we report on CM exon DNA methylation "memories" persisting beyond transcription repression and marking the expression history of numerous developmentally regulated genes, especially transcription factors.

1. Introduction

As a model system of human cardiomyogenesis and for cardiac cell replacement-based regenerative medicine, pluripotent-to-cardiomocyte differentiation strategies have recently undergone rapid advancement. In parallel, strategies for isolating mesoderm cells, a prerequisite lineage for cardiomyocytes (CMs), have also been under development. Such controlled cell fate specification, coupled with differentiation stage-specific cell isolation, provides not only an opportunity to describe molecular programs of cardiomyogenesis, but also an essential tool for investigating fundamental epigenetic mechanisms.

Pathological cardiac remodeling, often following myocardial infarction (MI), is marked by extensive CM hypertrophy and death. As a potentially unlimited cell replacement source, pluripotent-derived-CMs have been shown to electrically couple with host CMs, promote paracrine mediated cell survival, and aid widespread cardiac remuscularization post-MI in non-human primate models (Chong et al., 2014; Carpenter et al., 2012). hESC-derived-CMs were first isolated from regions of spontaneous contraction in embryoid bodies (Kehat et al., 2001), and CM differentiation has since been gradually improved via manipulation of key developmental pathways including: TGF-β, WNT, hedgehog (HH), FGF, and Notch signaling, among others (Freire et al., 2014; Lian et al., 2013). For clinical adoption, fully defined, serum and growth-factor-free differentiation strategies have evolved to include small molecule mimics of WNT signal modulators with yields of up to 98% pure CMs (Lian et al., 2013). Primitive mesoderm (PMESO) has been isolated using the cell

* Corresponding author at: Department of Diabetes Complications and Metabolism, Duarte, CA 91010, USA.
E-mail address: jtompkins@coh.org (J.D. Tompkins).

http://dx.doi.org/10.1016/j.ebiom.2016.01.021

surface marker ROR2 (Drukker et al., 2012) and cardiac mesoderm (CMESO) is known to express PDGFRα (Mummery et al., 2012). However, most candidate markers are, to variable extents, expressed on alternative cell types, and distinct lineages can require multiple surface markers for isolation (Evseenko et al., 2010). To date, no studies have globally assessed pluripotent-to-CM commitment using purified mesoderm cells.

It is widely understood that localized and genome-spanning networks of hierarchical epigenetic features dictate nucleosome positioning, chromatin architecture, and ultimately the interpretation of the genetic code. Epigenetic mechanisms, including DNA methylation and ncRNAs as related to cardiac development and disease, have been recently reviewed by Tompkins and Riggs (2015). DNA methylation, or more specifically 5-methylcytosine (5mc) typically in CpG sequences, is essential for normal development, and plays major roles in X-chromosome inactivation (XCI), imprinting, transposable element suppression, and tissue-specific gene expression (Smith and Meissner, 2013). Three enzymes are known to form 5mC, of which DNMT1 is generally considered the maintenance methyltransferase responsible for maintaining DNA methylation patterns through replication. DNMT3A and 3B, on the other hand, exhibit strong *de novo* activity (Goll and Bestor, 2005). Three enzymes, TET1, 2, and 3, are now known to actively demethylate DNA via oxidation of 5mC to 5-hydroxymethylC, which can be removed by base excision repair enzymes, such as TDG glycosylase (Kohli and Zhang, 2013). It has become clear that global DNA methylation patterns are, for the most part, stably maintained. However, methylation patterns can be dynamically altered during differentiation and presumably involve *de novo* methylation, active demethylation, and restriction of maintenance methylation activities.

The inverse relationship between promoter methylation and gene expression has been long known, extensively documented (reviewed by Smith and Meissner, 2013), and recently noted in hESC-derived-CMs at cardiac structural genes (Gu et al., 2014). In contrast, positive correlations exist between gene body methylation and expression, with models favoring transcription coupled DNA methylation and involving DNMT3B (Jjingo et al., 2012; Baubec et al., 2015). Considerable work remains, however, to comprehensibly understand functional roles for DNA methylation, especially outside of promoter contexts. Here, we use a state-of-the art, fully defined and scalable, good manufacturing practice (GMP)-compliant CM differentiation strategy and purify PMESO and CMESO cells to improve our understanding of global DNA methylation patterns over multiple stages of CM differentiation. We address lncRNA, gene expression, and transcription factors (TFs) that define human cardiomyogenesis, providing important genome-wide data sets. Lastly, we observed transcription and differentiation associated exon methylation to be enriched specifically at developmental TFs and to persist beyond gene silencing as a transcriptional "memory" of cellular differentiation.

2. Materials and Methods

2.1. hESC Cultures Through CM Differentiation

From an H7 hESC GMP suspension bank, CMs were derived under fully defined, GMP compliant conditions (City of Hope, Center for Biomedicine and Genetics) (Chen et al., 2012) (Fig. 1a) with adaptations from Lian et al. (2013).

2.2. Flow Cytometry, Cell Sorting

~2–2.5E5 accutase dissociated cells were washed with PBS 0.5% BSA, resuspended in 10 μl of antibody solution (30 min at 4 °C), washed again in PBS with 0.5% BSA (3×) and subjected to flow cytometry. Cell counts were conducted on a BD Accuri C6 Flow Cytometer (Cflow Plus). PMESO and CMESO sorts were performed on a BD FACS ARIA II Cell Sorter (FACS Diva V6.1.3).

See supplemental experimental procedures for additional details on CM differentiation, antibodies used, immunofluorescence, methyl binding domain (MBD) reactions, MBD-seq and RNA-seq library preparation, data processing, and validation.

3. Results

3.1. Differentiation of hESCs into PMESO, CMESO, and CMs

To globally describe in vitro human cardiomyogenesis, we prepared highly pure populations of hESCs, primitive mesoderm (PMESO), cardiac mesoderm (CMESO), and CMs. Importantly, a scalable GMP suspension bank of hESCs was coupled with fully defined, small molecule, GMP compliant differentiation (Fig. 1a). By flow cytometry, >98% of banked hESCs were positive for pluripotency markers SSEA-4, Tra-1-60, and POU5F1, whereas, <2% presented SSEA-1, ROR2, and PDGFRα differentiation markers (Fig. S1a and b). A pilot 31 day RNA-seq time course demonstrated progressive downregulation of pluripotency and rapid early upregulation of mesoderm genes from day 1 (D1) to D4, and subsequent delayed expression of cardiac progenitor (CP) and finally CM markers (Fig. 1b). Candidate mesoderm markers ROR2 and PDGFRα peaked from D2 to D4 and D4 to D6, respectively. Based on these expression patterns, we selected D3 and D4 for isolation of PMESO and CMESO, respectively (Fig. S1e). Duplicate hESC, PMESO, CMESO, and Percoll density purified CM samples (D31) were subsequently prepared as well as D3 ROR2(−) and D4 ROR2(+)/PDGFRα(−) cells. CM cell fate was confirmed by synchronized beating activity, >94% cTnT content by FACS (Fig. S1c), immunofluorescence for cardiac troponin I (cTnI) and myosin heavy chain 7 (MYH7) (Fig. S1d), and hierarchical clustering of RNA-seq data with published adult cardiac tissue data (Fig. S1g; (Lindskog et al., 2015). A double-positive ROR2(+)/PDGFRα(+) sort was used to purify maturing mesoderm (D4, CMESO) and validated with PMESO (D3) for target population enrichment (Figs. 1c and S1f).

RNA-seq results demonstrated PMESO cells to be clearly enriched for early mesoderm and mesendoderm markers and CMESO cells for CMESO and CP markers (Figs. 1c and S1f). By hierarchical clustering of all RNA-seq samples, both PMESO and CMESO sub-clustered together (Fig. S5a), and displayed significantly lower residual pluripotent gene expression, minimal trophoblast and ectoderm gene expression, and increased epithelial-to-mesenchymal transition (EMT) expression (ie: loss of *CHD1*, increased *TWIST*, *SNAI2*, *FN1*) (Kovacic et al., 2012). By CMESO stage, some CM genes were already upregulated (Fig. S1f). Relative to D3 ROR2(−) cells, PMESO exhibited 1163 differentially expressed genes (3× fold change, p-value <0.05). By gene ontology (GO) analysis, upregulated transcripts were clearly enriched for heart and skeletal system development (mesoderm derivatives; Fig. 1c). PMESO terms also included: *pattern specification process* ($p = 7.53E-15$), *gastrulation* ($p = 2.0E-09$), and *mesoderm development* ($p = 1.3E-05$) (Table S1). Upregulated CMESO genes, relative to D4 ROR2(+)/PDGFRα(−), were predominately enriched for cardiac development terms (1248 differentially expressed genes), showing clear transition towards CM cell fate. D3 ROR2(−) cells, in distinct contrast, were enriched for *neuron development* ($p = 1.3E-7$; Fig. 1c), and at D4 ROR2(+)/PDGFRα(−) cells displayed enrichment for multiple developmental pathways including: *endocrine system development* ($p = 4.6E-05$), *neuron differentiation* ($p = 8.5E-05$), and *epithelial cell differentiation* ($p = 1.4E-04$), among others. Taken together, FACS-based purification of PMESO and CMESO populations removes otherwise contaminant transcriptional, and by extension, epigenetic signatures from heterogeneous multi-lineage and non-committing cells.

3.2. Dynamic Gene-coding Transcriptional Programs During Stage-specific CM Differentiation

To investigate temporal patterns of gene expression over our 4 point time-series (hESCs, PMESO, CMESO, CM), we used Grid Analysis of

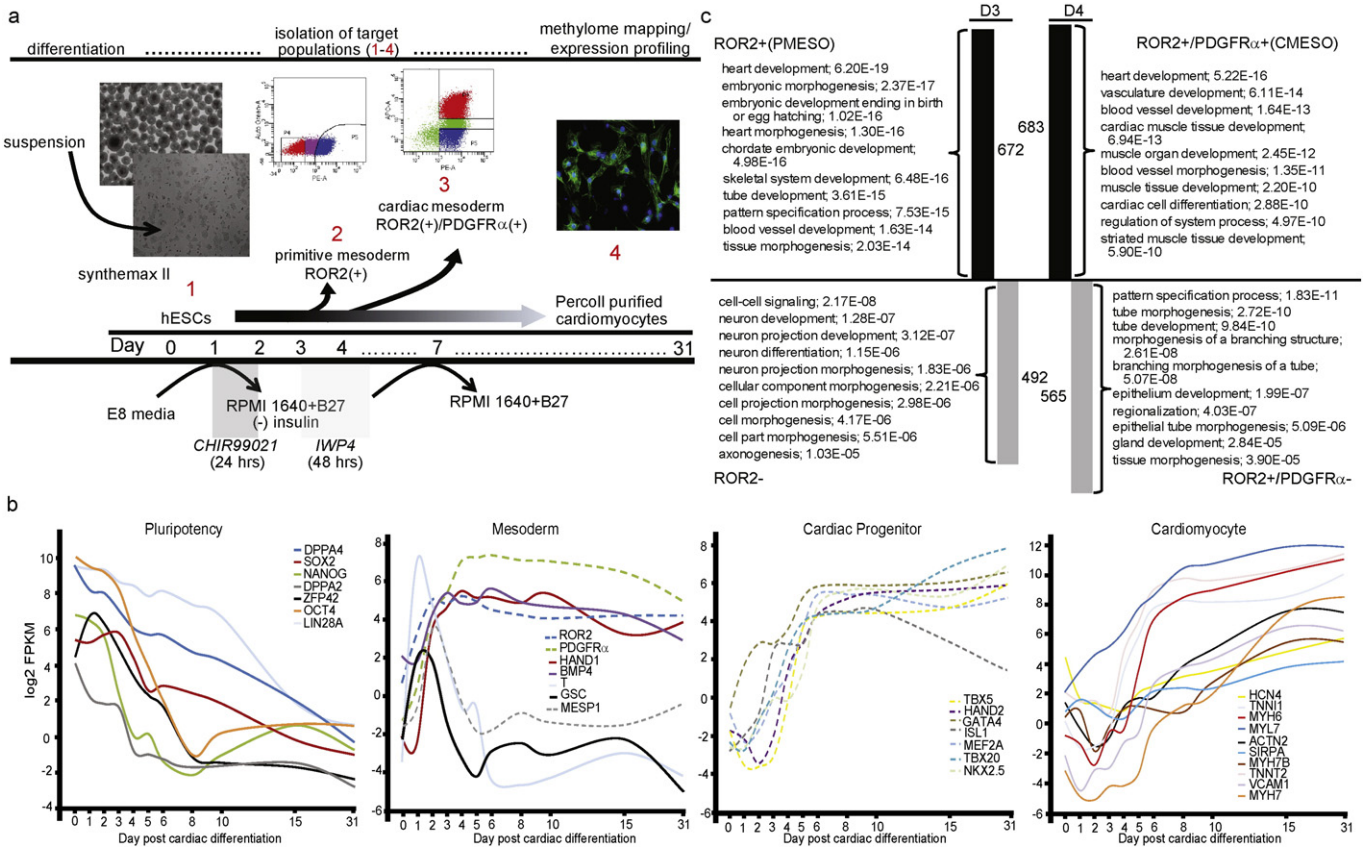


Fig. 1. CM differentiation, sample preparation, and validation. a) Experimental strategy with 4 target populations marked red. Briefly, suspension hESCs were adhered to Synthemax II plates and CM differentiation induced by CHIR99021 and IWP4. CMs were Percoll density purified and PMESO and CMESO by FACS. b) Pilot time course of known developmental markers. Left to right: pluripotent, mesoderm, cardiac progenitor, and CM marker expression. c) Differential expression of PMESO and CMESO compared to D3 ROR2(−) and D4 ROR2(+)/PDGFRα(−), respectively. Black bars = upregulated; gray = downregulated. Top 10 ontology terms are displayed (BP term, EASE p-value).

Time-Series Expression (GATE) (MacArthur et al., 2010). GATE dynamically colors time-series data according to transcript level, grouping genes with common expression patterns across a 2-dimensional array. A "snap shot" of PMESO and CM stages of all differentially expressed genes (2917) is displayed in Fig. 2a. By selecting stage-specific up or downregulated transcripts, it becomes evident that CM differentiation is marked by progressive upregulation of heart development genes and concomitant downregulation of mitotic and organelle fission genes. Pathway specific regulation of the cell cycle is highlighted in Fig. S2m. Genes enriched for glycolysis (Table S2; BP: *glycolysis*) were suppressed during mesoderm stages, but were generally increased within CMs which demonstrated massive upregulation of mitochondrial oxidative phosphorylation genes (Fig. 2a; Table S2; Fig. S2n–p). Collectively, these observations are well in line with differentiation associated CM cell cycle exit (Sdek et al., 2011), mitochondrial fission gene downregulation, and switches to predominately mitochondrial oxidative metabolism from anaerobic glycolysis (Chung et al., 2007). Besides obvious CM stage pathway upregulation (e.g., *cardiac muscle contraction*, p = 2.52E − 5), cellular component GO demonstrates cardiac structural gene upregulation. *Contractile fiber*, *myofibril*, *I band*, *Z disk*, *A band*, and *sarcomere* were among the most highly enriched terms at each differentiation stage (Table S2). Lastly, we identified 289 alternatively spliced genes between CMs and earlier time points. Significant overlap was observed with published results of alternative splicing in human cardiac precursors (Fig. S3i; (Salomonis et al., 2009)) and transcripts were enriched for functions in muscle tissue development and contraction. Gene lists and GO results are provided in Table S2 and alternative splicing at TPM1 is shown in Fig. S3h.

3.3. Transcription Factor and miRNA Regulation of Coding Gene Expression

Wnt, TGF-β, and hedgehog (HH) pathways were significantly enriched among all differentially expressed transcripts over the 4 point time series and in PMESO and CMESO samples relative to alternative lineage cells (e.g., D3 ROR2(−) and D4 ROR2(+)/PDGFRα(−); Tables S1 and S2). These critical developmental pathways are intimately connected (Fig. S2a–c) (Freire et al., 2014) and several distinct subclasses were uniquely enriched at each differentiation stage (Figs. 2b–d; S2L).

Differential gene expression was also investigated with GATEs integrated databases of transcription factor (TF) interactions by predicted protein binding (TFs_predicted_binding_sites.gmt), by prior chromatin immunoprecipitation (ChIP) studies (TFs_chip_interactions.gmt), and for miRNA regulation (predicted_microRNAs.gmt). The highest ranked predicted TF was myocyte-specific enhancer factor 2 A (MEF2a) (V$RSRFC4_01; $p < 0.0001$). *MEF2a* is preferentially expressed at PMESO and CMESO stages, highly upregulated in CMs (Figs. 2e, S1f, S2d), and positively correlated with predicted target gene expression. In contrast, polycomb complex (PRC) 1 member, polyhomeotic homolog I (*PHCI*), was preferentially silenced over CM commitment and inversely correlated with target gene expression (Figs. 2f and S2e). The ~13 fold reduction in PMESO, relative to hESCs and D3 ROR2(−) cells (Fig. S1f), is consistent with a developmental repressive PHC1 function. Both MEF2a and PHC1 targets are involved in CM development and contraction functions; the highest ranked GO term for MEF2a and PHC1 were *striated muscle tissue development* (p = 1.27E − 5) and *heart development* (p = 3.3E − 15), respectively

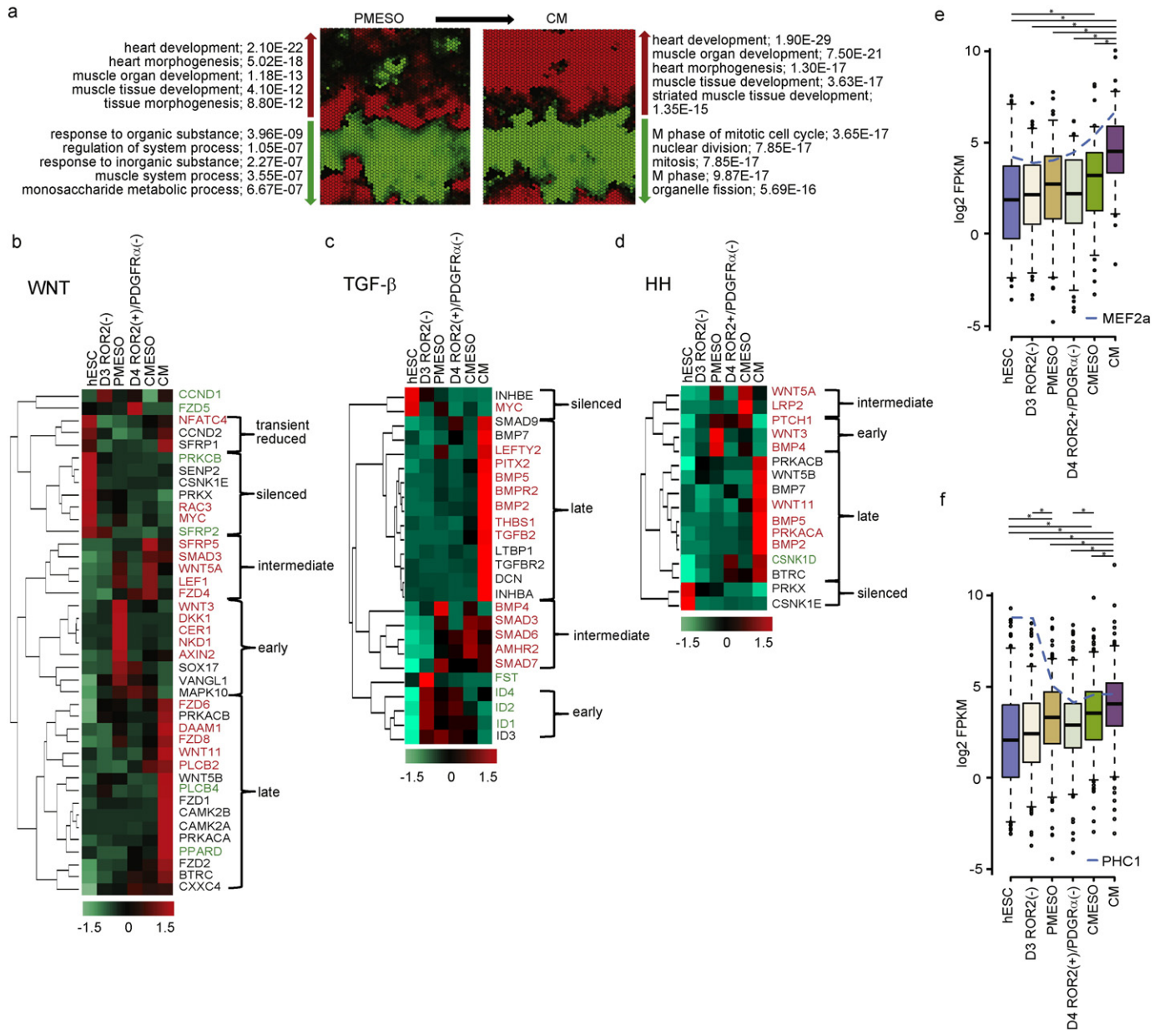


Fig. 2. Differential expression during hESC-to-CM commitment. a) "Snap shot" of PMESO and CM stages by GATE visualization and top 5 ontology terms for up- and downregulated genes (BP term, EASE p-value). Hexagons are individual genes clustered by temporal differentiation expression patterns. Red hexagons = upregulated; green = downregulated. b–d) Expression hierarchical clustering of Wnt (b), TGF-β (c), and HH (d) pathway genes. Most individual sub-clusters were clearly defined by stage-specific expression (i.e., early, late) and individual genes were colored based on enrichment in PMESO and/or CMESO stages relative to alternative lineage cells. At PMESO and/or CMESO, red genes = enrichment; green genes = suppression; no color = no or unclear enrichment. Scale = row z-score. e–f) Expression distribution of predicted MEF2a (e) and PHC1 (f) targets during CM differentiation. Center lines = medians; box limits = 25th and 75th percentiles with whiskers extended 1.5× the interquartile range from those percentiles (R software); outliers = open circles. (*$p < 0.05$; Mann–Whitney U-test). Blue dashed lines = *MEF2a* (E) and *PHC1* (F) expression.

(Fig. S2d and e; Table S3). For potential miRNA regulation, much like repressive PHC1, miR-302/367 silencing-marked differentiation, was enriched in alternative lineage cells, and was inversely correlated with strong target gene induction (Fig. S2h–k).

3.4. lncRNA Associations with Coding Gene Transcription

lncRNAs, defined as non-coding transcripts >200 bp in length, have previously been reported to be expressed at lower levels than coding genes (Cabili et al., 2011; Chakraborty et al., 2014). If we consider >1 FPKM as an expressed transcript, it is clear that relative to the wide expression range of coding genes, most lncRNA expression falls within a narrow range (>1 and <10 FPKM) with subtle range widening by CM stage (Figs. 3a and S5h). In general, hESC-to-CM differentiation was marked by increasing densities of expressed coding genes and to a lesser extent lncRNAs, and decreasing proportions of silent coding genes (FPKM < 1). lncRNAs paired in *cis* with the nearest coding gene were hierarchically clustered, and a number of sub-clusters emerged with clear stage-specific enrichment or exclusion (Fig. 3b). For each sub-cluster, a positive correlation exists between lncRNAs and *cis* paired coding gene expression (Fig. 3c). For these lncRNA associated genes, the highest ranked GO term was *cell fate determination* ($p = 0.004$) and ~15% of lnc-associated genes were classified under *regulation of transcription, DNA dependent* (53/370; $p = 0.004$). In line with the high tissue specificity of lncRNAs (Cabili et al., 2011), numerous *cis*-associated developmental TFs were identified, including *MEIS1* and *2* (Fig. 3d), *TBX2* (Fig. 5G), *GATA3* (Fig. 6c), and *BMP4* (Table S4). Differentiation-induced lncRNA-coding gene pairs are also involved in cell adhesion

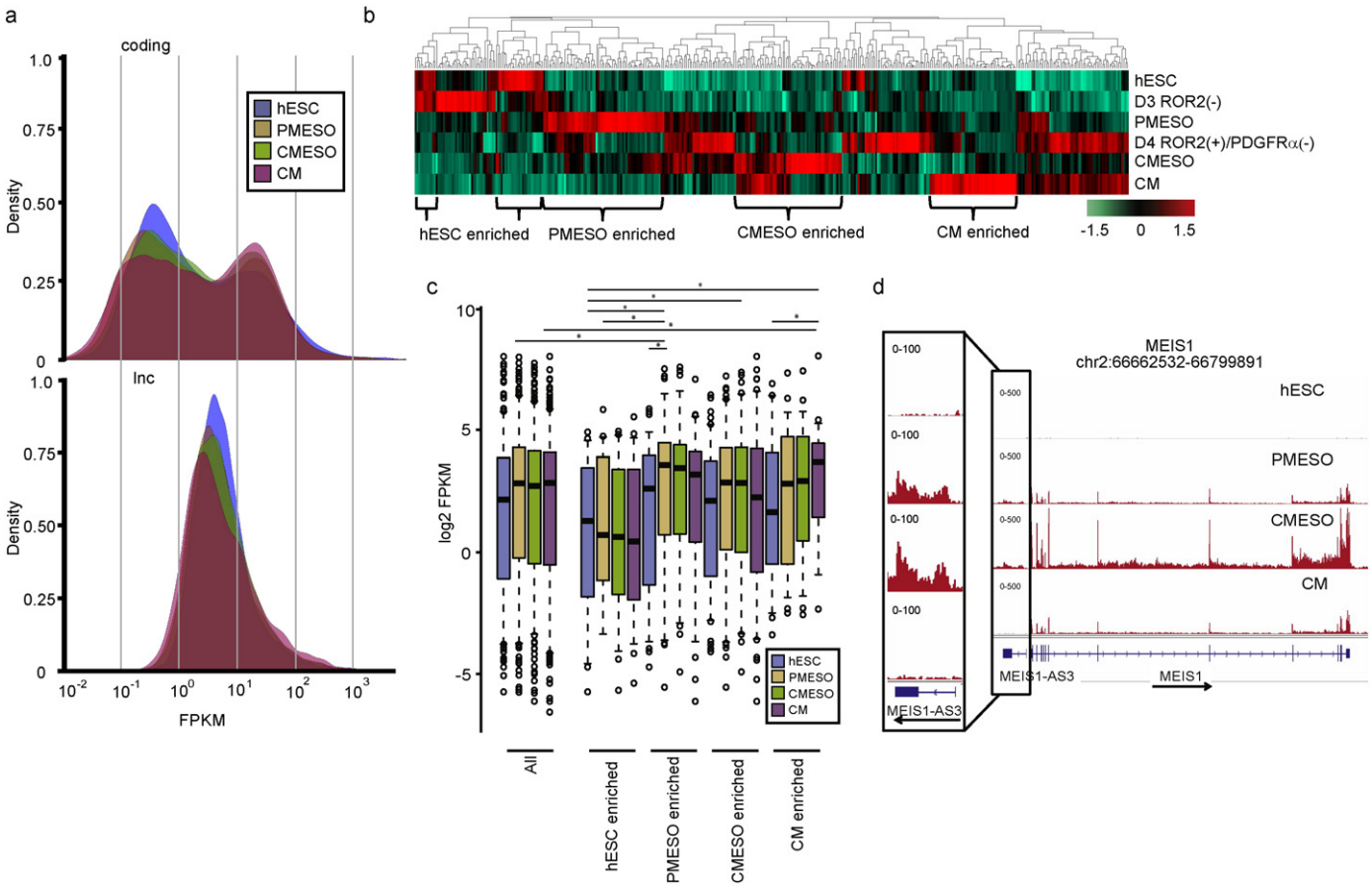


Fig. 3. lncRNA expression and CM differentiation. a) The frequency of FPKM values (density profile) for coding and lncRNA gene expression was plotted for each differentiation stage (expanded in Fig. S5h). b) Hierarchical clustering of lncRNAs by expression. Scale = row z-score. c) Stage-specific lncRNA sub-clusters were correlated with nearest *cis* matched coding gene expression. Box plots display paired coding gene expression (Fig. 2f for box plot description). d) Positive association of lncRNA (*MEIS1-AS3*) and higher level *MEIS1* expression.

(e.g., *CDH1*, *CDH11*, *COL4A6*) and cell cycle functions (e.g., *CDC16*, *CDK10*, *CCND1*), and overall, results suggest stage-specific lncRNA expression to be functionally correlated with coding genes regulating key aspects of CM differentiation.

3.5. Global DNA Methylation Patterns and CM Differentiation

To investigate both global and region specific DNA methylation features of CM differentiation, methyl binding domain (MBD)-seq was conducted for hESCs, PMESO, CMESO, and CMs. We find that CG island (CGI) methylation progressively increases during CM differentiation, with lncRNA genes harboring elevated CGI methylation relative to coding genes (Fig. 4a). DNA methylation was mapped to a composite model of all annotated transcripts, and both lncRNA and coding genes were observed to be hypomethylated at TSS-centered promoters (Fig. 4b). For both species, promoter methylation was inversely correlated with expression, but lncRNA genes on average were significantly more methylated through promoter and intragenic regions (Figs. 4b, S4c–d). Coding and lncRNA genes transiently lost methylation upstream of the TSS and within intragenic regions at mesoderm stages, notably more so for CMESO lncRNA genes. This observation extended into mesoderm methylation losses at SINE and Alu repetitive elements. By contrast, satellite elements increased methylation throughout differentiation (Fig. 4d). Likely reflecting CM cell cycle exit (Fig. 2a), *DNMT1* expression dropped significantly from CMESO-to-CM stages (Fig. 4e). *DNMT3A* expression approximately doubles from hESCs-to-PMESO and converges with pronounced *DNMT3B* downregulation. Although lower DNMT3B *de novo* methyltransferase activity may be responsible for transient mesoderm hypomethylation, it is striking that *TDG* DNA glycosylase is maximally expressed at PMESO and CMESO stages and *SMUG1* glycosylase at CMESO, accompanying dynamic changes in *TET* gene expression (Fig. 4f). These results suggest targeted DNA demethylation to be a natural feature of mesoderm and CM differentiation.

3.6. Differential DNA Methylation and CM Differentiation

Differentially methylated regions (DMRs) were identified between samples (see supplemental experimental procedures), annotated and quantified by genomic location, assessed for genome and gene ontology, and correlated with gene expression. DMRs numbered the highest between CMESO-to-CM stages (11,236,27 days), yet even the single day PMESO-to-CMESO transition resulted in 5380 DMRs. Thus DNA methylation at earlier stages of differentiation is particularly dynamic. DMR-based genome ontology results overall reflect CM differentiation (e.g., CMESO-CM = *cardiac muscle cell differentiation*, $p = 3.7E-14$; see Table S5 and Fig. S3a). Genes were quantified for DMR gains or losses within defined regions (promoters, exons, introns, etc), and consistent with global methylation trends, demonstrated a bias for PMESO-to-CMESO hypomethylation and CMESO-to-CM hypermethylation (Fig. S3b).

CM hypermethylated promoter DMRs, relative to preceding time points, were associated with decreasing or restrained gene expression and hypomethylated DMRs with increasing gene expression (Figs. 5a; S3c, OCT4, Nanog). By CM stage, the median expression of genes with promoter methylation gains was 5.04 fold lower than those with methylation loss ($p < 0.05$). Promoter regulation extended across functional gene groups involved in CM development, structure, and contraction (Fig. 5b). Both methylation gains ($p = 0.03$) and losses ($p = 0.018$)

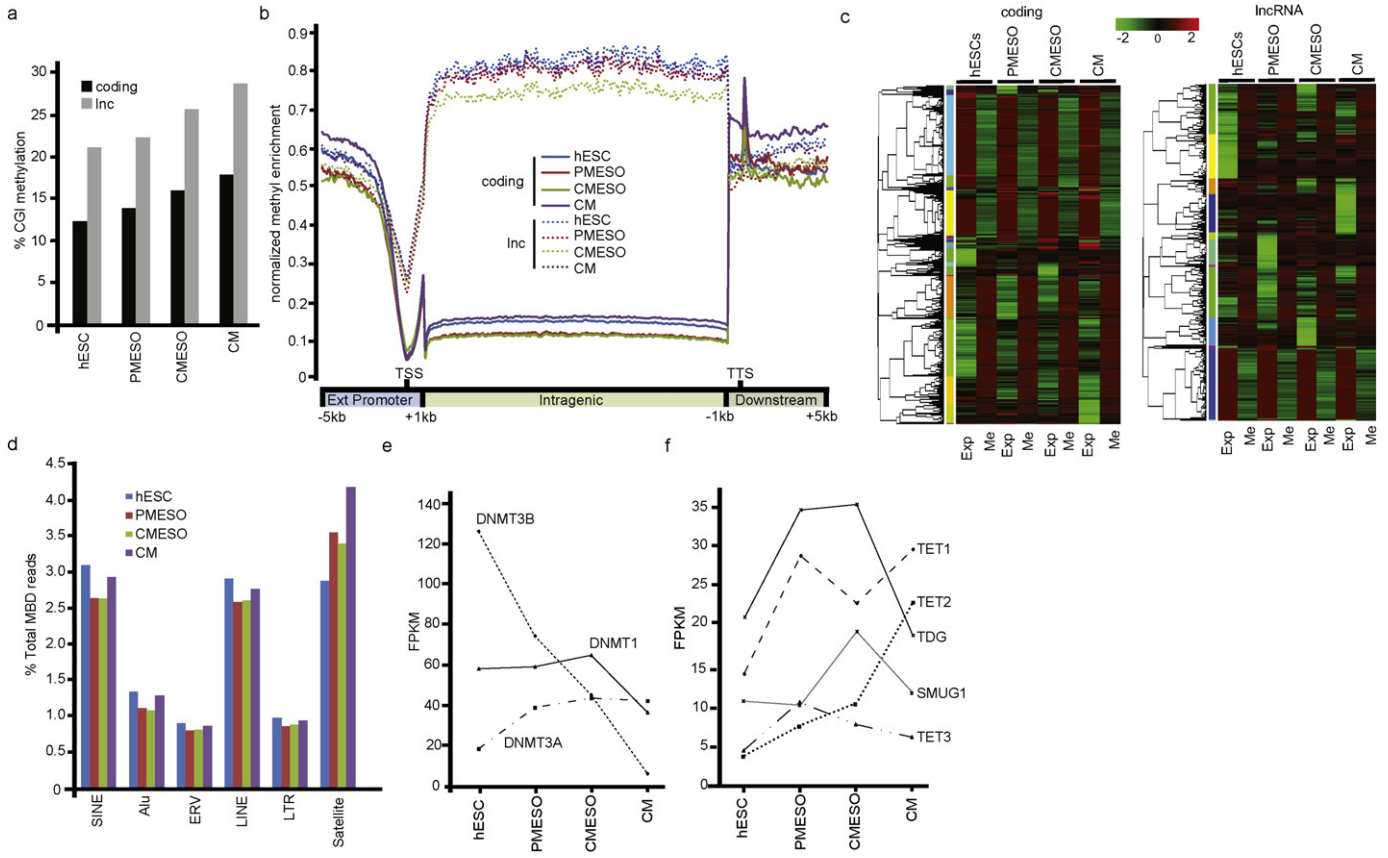


Fig. 4. Global methylation and CM commitment. a) Percent CGI methylation over differentiation. Methylated CGI = MBD/input ratio of greater than 1.5 across the island. b) Normalized methylation enrichment was plotted across promoter, gene body, and downstream bins representing a composite model of coding or lncRNA genes. c) Hierarchical clustering of promoter methylation and gene expression. Genes or lncRNAs differentially expressed over differentiation were clustered by expression (Exp) and matched MBD-seq values (Me) for promoter window −5 kb to +1 kb of the TSS. Scale = row z-score. d) Bar graph of total MBD-seq reads mapped to each repetitive element e) Expression of DNA methyltransferases and f) demethylation pathway members by differentiation stage.

were enriched for *heart development*; whereas, the top ranked GO term was *contractile fiber* (hypo, p = 6.75E−5), complementing *muscle cell contraction* (p = 6.52E−3) and correlating with massive CM stage up-regulation (Fig. 5b). CM hypermethylated promoters were also enriched for *basolateral plasma membrane* (p = 4.0E−3), hypomethylated promoters for *focal adhesion pathway* (p = 6.1E−3), and both gains and losses at *cell adhesion* promoters (Fig. 5b). CM Promoter DMRs therefore also reflect fundamental changes in cell architecture, including cell–cell and cell-extracellular matrix interactions that are known to occur over hESC-to-CM differentiation. CM promoter hypomethylation at cardiac muscle *MYH6* and *MHY7* is displayed in Fig. 5c.

Consistent with global expression results, several WNT, HH, and TGF-β genes also underwent promoter methylation changes, with the latter pathway being CM stage hypomethylated and upregulated (Fig. 5b). Differentially methylated WNT and HH genes tended to acquire promoter methylation post-transient mesoderm induction (Figs. 5b and S3d) and often harbor gene body DMRs as well (Fig. S2a–c). Methylation changes among pathway members may also interface with expected cell cycle (Table S6, CMESO hypo, *M phase*; p = 3.7E−2) and apoptotic regulation with connections to Rho-Rock signaling as cells differentiate and approach terminal states (Fig. S2a–c, expanded in Fig. S3e).

3.7. Gene Body Methylation and Expression

We and others have previously observed positive correlations between intragenic DNA methylation levels and gene expression, although this strong, positive correlation only holds up to mid-level expression and actually inverts at highly expressed genes (Tompkins et al., 2012; Jjingo et al., 2012). These observations are highly consistent with transcription coupled methylation activity. It remains unclear, however, whether elevated transcription activity and dense RNAPII occupancy inhibits linked methylation deposition (Jjingo et al., 2012) and/or if targeted demethylation is requisite to maximum transcription (Veloso et al., 2014; Baubec et al., 2015). Here, both CM intragenic methylation DMR gains and losses were positively correlated with differentiation associated increases in gene expression (Fig. 5d). Yet, consistent with transcription linked DNA methylation, intragenic methylation gains tend to occur at genes expressed at low levels during preceding differentiation stages, whereas methylation losses tend to occur at genes already expressed at moderate levels at prior stages. These observations in the context of transcription associated histone modifications are revisited within the discussion. GO results also suggest genes with intragenic methylation changes have roles in CM differentiation and function including 66 TFs (*regulation of transcription*, p = 2.7E−4, Table S6). Top ranked GO terms for CM exon methylation gains and losses were *cell adhesion* (p = 4.7E−9) and *induction of apoptosis*, respectively (1.4E-2). For introns, these terms were *ion transport* (hypermethylation, p = 3.7E−5) and *intracellular signaling cascade* (hypomethylation, 2.9E-5), respectively. Differential intragenic methylation at the *focal adhesion pathway* is also expanded (Fig. S3f–g).

In higher eukaryotes regulation of splicing is enigmatic, but recent evidence has implied a role for DNA methylation. Evidence suggests that splicing may be co-transcriptional and involve epigenetic interpretation, nucleosome positioning, and/or involve the recruitment of chromatin binding proteins including DNA methylation associated binding of MeCP2 (Maunakea et al., 2013; Lev Maor et al., 2015). In our study, we see clear evidence for both alternative splicing (Fig. S3h-l and

Table S2) and differential intragenic methylation changes over CM differentiation, and further, alternatively spliced genes are overrepresented among genes with DMRs (Fig. S3j). However, the overwhelming majority of genes with intragenic DMRs show no evidence for splicing, and exon methylation changes among alternatively spliced genes do not significantly occur (Fig. S3k). Although our results imply a connection between differential methylation and splicing, a direct methylation change to splicing event correlation was not identified.

3.8. Exon Methylation as a "Memory" of Developmental History

Similar to CM promoter methylation gains, exon hypermethylated DMRs were also enriched at *embryonic morphogenesis* ($p = 4.3E{-}8$) and *heart development* (9.9E-3) genes. As many CM exon methylated genes were observed to also have undergone promoter methylation (35.2% overall; 126/358), this presented an opportunity to investigate "dual methylation" events for correlations with gene expression (Fig. 5e). For *embryonic morphogenesis* genes with exon methylation gains, 11/22 also gained promoter methylation (Fig. 5f; Table S6; $p = 2.1E{-}3$). *TBX1* and *TBX2* provide clear examples. Both underwent progressive and extensive differentiation associated exon methylation and yet had widely divergent promoter and gene expression activity (Fig. 5g). As such, promoter hypermethylation with resulting reduction of transcription, appears to override otherwise positive intragenic methylation correlations with gene expression (Fig. 5e–f). This interpretation however, raises an important question regarding transcription coupled DNA methylation: How can one explain exon methylation gains in the absence of transcription? With this in mind, we re-examined gene expression at 7 additional time points between hESC and PMESO and CMESO and CM stages, focusing on gene sets undergoing exon hypermethylation by CM stage (Fig. 6a and b). Genes were re-sampled with increased stringency, requiring exon DMR gains in CMs versus hESCs and PMESO or CMESO. Of the resulting 48 exon methylated candidates, most were developmental TFs and most were transcriptionally induced late in differentiation, positively correlating CM exon methylation with transcription (Fig. 6a). Given these candidates, we also interrogated the aforementioned 66 TFs from CM exon hypermethylated DMR GO results (Table S6; Fig. 6b). These methylation changes were independent of gene splicing (Fig. S3l, legend) and although a significant portion of CM exon methylated genes were downregulated or silenced by this stage (~1/2 of TFs), virtually all (92%; 105/114), including TBX1, were expressed at FPKM > 1 during differentiation (by comparison, 16,873/41,495 or 40.7% of all annotated genes exceeded 1 FPKM). Merely the timing of expression had fallen between CMESO and CM methylation sampled time points. This collectively indicates that 1) exon methylation is generally correlated with elevated gene transcription during CM differentiation and 2) exon methylation can persist beyond the time of gene silencing marking developmental history. Residual exon methylation "memories" were enriched specifically among developmental TFs. Examples span from expected exon-expression correlations (Fig. 5g-*TBX2*, Fig. 6c-*TBX3*, *CBX4*, *FZD1*) to those silenced by CM stage and harboring transcriptional methylation "traces" (Fig. 5c-*SHH*, *SP6*, *GATA2*, *GATA3*, *FOXF2*). This "exon methylation memory" was also observed among the HOX genes, in which the HOXB cluster was uniquely expressed over CM differentiation and was the only HOX cluster with significant exon methylation, including persistent methylation beyond gene silencing (Fig. S4a and b).

4. Discussion

4.1. CM Differentiation Strategies for the Clinic

Pluripotent-derived-CMs promise an unlimited source of replacement cells for cardiac regeneration. In this report, we have utilized a fully defined, xeno-free, scalable GMP suspension hESC bank, and driven differentiation by small molecule WNT signal manipulation under defined and GMP compliant conditions. Such culture strategies are a pre-requisite to realizing clinical potential (Chen et al., 2012; Tompkins et al., 2012) and were built through manipulation of established regulators of cardiac development (e.g., Wnt signaling). In this study, scalability aided FACS purification of PMESO and CMESO cells for global transcription and DNA methylation studies and has provided a rich resource of candidate transcription factors, lncRNAs, and miRNAs, that represent a next generation of targets for improving the efficiency and perhaps maturity of derived CMs. As a next step for widespread clinical use, ongoing pre-clinical work is focused on a similar, but entirely suspension-based CM differentiation protocol.

4.2. Purification of Target Cell Populations for Genome-wide Studies

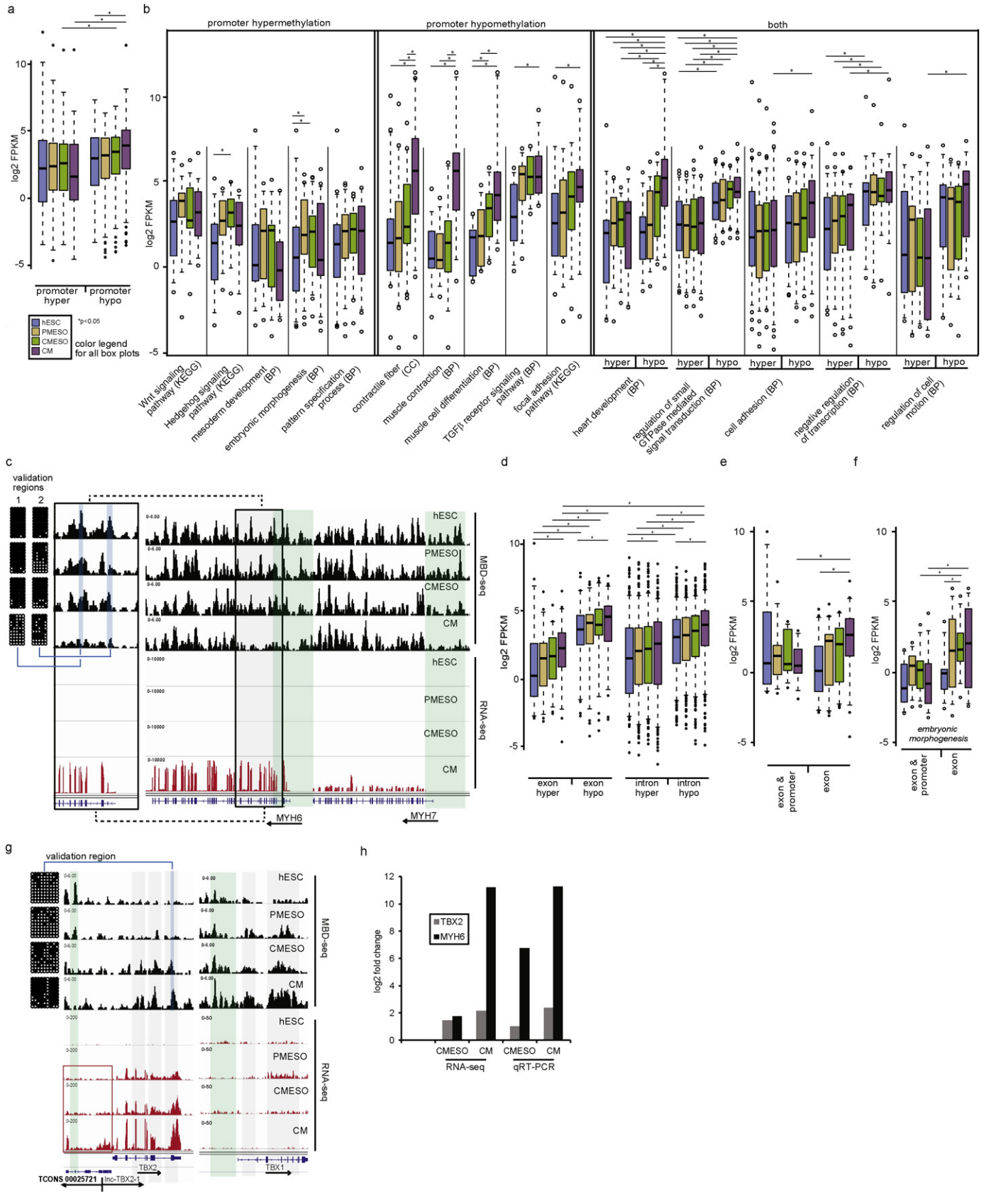
Mesoderm purification is a necessary step towards more clearly understanding human cardiomyogenesis and may also prove useful in the treatment of cardiac pathologies. For the latter, other groups have observed that relatively immature CMs or progenitors improve transplant engraftment, with evidence for in vivo maturation (Carpenter et al., 2012; Fujimoto et al., 2011; Tompkins and Riggs, 2015), and therefore, transplantation of PMESO or CMESO isolates might also improve cell retention and mediate pro-survival paracrine signaling in vivo.

Two groups, Paige et al. (2012) (hESCs-to-CMs) and Wamstad et al. (2012) (mESCs-to-CMs) have profiled global histone modifications and gene expression in prior studies of CM differentiation. However, cellular heterogeneity, including final CM preparations that were ~25–30% cTnT(−) cells, significantly confounded epigenomic results (Parmacek and Epstein, 2013). Here, we can estimate that at minimum, 26% of D3 cells would be ROR2(−), 30% of D4 cells ROR2(+)/PDGFRα(−), and 21% of D31 cells (final CM time point) cTnT(−) (Fig. S1c–e). These potential contaminants were removed and found to contain both non-committing and multi-lineage differentiating cells (Fig. 1c; Table S1). By comparison, in hESC-to-CM studies, Paige et al. specifically noted the presence of smooth muscle and endothelial cells in CM preparations with a minimum of 50% cTnT+ CMs in sample preparations for genome-wide studies (Paige et al., 2012). Nonetheless, several important features of cardiomyogenesis were observed. Among key findings were the activation of Wnt, HH, and TGFβ gene families, similar to our results, as well as differentiation associated increases in transcription coupled with increased promoter H3K4me3 and TSS downstream H3K36me3. These modifications are well known to correlate with decreased promoter DNA methylation and increased intragenic DNA methylation (Baubec et al., 2015; Morselli et al., 2015; Paige et al., 2012), respectively, in agreement with results from our study.

Fig. 5. Differential DNA methylation and CM commitment. a) All CM promoter DMRs by hyper- or hypomethylation and b) those corresponding to selected enriched ontology terms (Table S6) were assessed for gene expression over CM differentiation. For all box plots in this figure, see Fig. 2f for methods. The color legend in panel a applies to all box plots. c) CM promoter hypomethylation at cardiac structural genes. Light green = promoter. *MYH6* is one of the most highly expressed CM genes and hypomethylation extends into the first several exons. Light blue = bisulfite sequencing validation region. Rows of circles represent consecutive CpG sites of individual sequences. Black circles = methylated. Additional validation at *TBX2* is shown in panel g. d–f) Expression distribution of genes with multiple types of CM gene body DMRs. Annotated 5′ and 3′ UTR exon DMRs were combined with coding DNA sequence exon DMRs for ontology and gene expression correlations. d) All CM exon or intron DMRs by methylation change. e and f). Dual promoter-exon methylation gains compared to those with CM exon gains, but lacking promoter hypermethylation DMRs for all genes and f) for CM exon methylated genes enriched for *embryonic morphogenesis* (Table S6). g) *TBX1* (restrictive promoter methylation gain) and *TBX2* (permissive methylation loss) exhibit differentiation associated intragenic hypermethylation. H) Validation of RNA-seq data at *MYH6* and *TBX2* genes, which undergo differential methylation. Both qRT-PCR and RNA-seq data are normalized to internal control *TPTI* and expressed as log2 fold change relative to PMESO time point. For these genes, there was no detectable amplification of transcripts from hESCs. qRT-PCR data for all RNA-seq samples is provided in Fig. S5c and d.

4.3. lncRNAs and Candidates for Functional CM Manipulation

Wamstad et al. (2012) observed that lncRNA expression in mESC-derived-CMs positively correlates with *cis* paired coding gene expression, and we extend this here for hESC-to-CM differentiation (Fig. 3c). Aided by PMESO and CMESO purification, we find that the overwhelming majority of examined lncRNAs are enriched unambiguously by differentiation stage, in line with known high tissue-specific



lncRNA expression (Fig. 3b) (Cabili et al., 2011). Unique expression within PMESO and CMESO populations suggests functional roles within CM development (e.g., *linc-MEIS1-1*, *linc-MEIS2-2*, *linc-BMP4-1*, etc) and this logic was also applied to identify key signaling pathway members (Fig. 2b–d), TF targets (Fig. S2d–e), and miRNAs (Fig. S2h–k). These results provide a diverse array of major candidate sets for future functional genetic and epigenetic manipulation studies of CM development and disease. Although functional lncRNA roles in murine cardiac development have been identified (Fatica and Bozzoni, 2014; Klattenhoff et al., 2013), the poor conservation between mouse and human lncRNAs (Ulitsky et al., 2011) and the dysregulation of lncRNA expression in human cardiomyopathies (Papait et al., 2013), strengthen further the importance of identifying lncRNAs specific to human cardiomyogenesis.

4.4. lncRNAs and DNA Methylation

Both lncRNA and coding genes exhibit a negative correlation between promoter methylation and gene expression and generally positive correlations between gene body methylation and expression (Figs. 4c and S4d). Interestingly, lncRNA genes exhibit higher methylation at TSS centered regions, at CGIs, and across intragenic regions (Figs. 4a–b, S4c–d). Higher TSS centered methylation relative to coding genes is consistent with the lower average expression of lncRNAs (Fig. 3a) and the well known inverse relationship between promoter methylation and transcription. With regards to intragenic methylation, lncRNAs have little overall expression range and therefore increases to gene body methylation with increasing expression percentiles, though evident, become muted relative to the wide intragenic methylation and expression range of coding genes (Figs. 3a and S4d). Recent identification and annotation strategies for describing lncRNAs have been necessarily improving, addressing complexities associated with handling diverse arrays of lncRNAs at a variety of overlapping genomic elements. In the future, our data can be revisited to identify novel lncRNAs, lncRNA subtype distributions, and functional correlations with distinct methylation events. Indeed, some lncRNAs are specifically associated with common features involving DNA methylation such as CGI, repetitive element, imprinting methylation, and MeCP2 binding (Forne et al., 1997; Chakraborty et al., 2014; Nan et al., 1997; Cabili et al., 2011) and physical associations may exist between DNMTs and lncRNAs (Wang et al., 2015).

4.5. DNA Methylation, Transcription, and Development

DNA methylation changes, both increases and decreases, occur throughout development and it is now well established that a combination of maintenance, *de novo*, and demethylating activities are required for proper methylation distribution both globally and specifically within promoters, enhancers, and gene bodies. Highlighting the essential role for DNA methylation in mammalian development, *Dnmt1* −/− mouse embryos die shortly after gastrulation and *Dnmt3b* −/− or *Dnmt3a* −/− mice die at embryonic day 9.5 and shortly after birth, respectively (Jurkowska et al., 2011). Deletion of *DNMT1* is not tolerated in hESCs (which are more epiblast like than mESCs), and *Dnmt1* −/− mESCs die upon differentiation. Interestingly, in mESCs (Jurkowska et al., 2011; Takebayashi et al., 2007), and recently demonstrated in hESCs (Liao et al., 2015), neither *DNMT3A* nor *DNMT3B* are required for ESC viability, but additional passaging of *Dnmt3a* or *Dnmt3b* knockout ESCs is associated with global hypomethylation and prevents differentiation. Both *de novo* transferases display considerable global binding redundancy and target SINE and LINE elements (Jin et al., 2011), but Liao et al. (2015) observed that DNMT3B knockout hESCs specifically lose satellite element methylation. However we find that despite virtual *DNMT3B* silencing by CM stage, satellite element methylation is highest in CMs where *DNMT3A* expression, but not *3B*, predominates (Fig. 4d–e). Although the specificity of DNA methyltransferases and demethylase enzymes is limited (e.g., CpG or 5mC), specificity for extended sequences is directed by chromatin structure and/or specific cofactors, which can include lncRNAs as well as transcription factors (Smith and Meissner, 2013). Future studies will continue to delineate the molecular details of DNA methylation regulation and will no doubt be aided by tracking methylation dynamics through cellular differentiation.

4.6. Histone Methylation, Transcription, and Gene Body Methylation

Recently a clear connection between histone methylation and intragenic DNA methylation has been established in which *de novo* methyltransferase DNMT3B binds H3K36me3, which marks regions of active transcription (Morselli et al., 2015; Baubec et al., 2015), and this may help explain positive correlations between gene body methylation gains and increasing gene transcription (Fig. 5d). Morselli et al. (2015) have shown in yeast, which normally has no 5mC, that H3K36me3 introduced by SET2 histone methyltransferase dictates the location of DNA methylation by recruiting DNMT3B (exogenously introduced). This same study shows that H3K4me3, which is located preferentially at active promoters, strongly prevents DNMT3B recruitment. Thus, DNMT3B activity and resulting methylation patterns are strongly influenced by intragenic H3K36 methylation patterns. Our working hypothesis is that transcription of both coding and lncRNA genes leads to an increase in H3K36me3 by action of SET2 histone methyltransferase and this in turn leads to an increase in DNA methylation over the transcribed region. Therefore, transcription coupled DNA methylation requires preceding H3K36me deposition. At high transcription levels the transcribed region can become less methylated (ex: MYH6, but not MYH7 in Fig. 5c), but we assume that this is by a different mechanism that applies to relatively rare, highly transcribed regions (Fig. S5f). Reduced exon methylation in highly expressed genes might be due to tandem RNAPII blocking of maintenance methylation activity (Jjingo et al., 2012), or increased active demethylation. Indeed, the TET enzymes, TDG, and SMUG1 glycosylases are generally induced upon CM differentiation with the glycosylases transiently peaking at CMESO stage and corresponding to transient hypomethylation within gene bodies (Fig. 4b and f).

4.7. Exon Methylation, Transcriptional "Memory Traces" and Development

The linking of intragenic methylation with transcription likely explains why gene body DNA methylation is a better predictor of cell identity than promoter methylation (Illingworth et al., 2010). In this study, we observed CM hypermethylated DMRs to be enriched at exons of developmental TFs, including many TBOX (TBX1-5) and many HOXB cluster genes. Of the 4 HOX clusters, only HOXB became hypermethylated within the CM lineage and, strikingly, aside from low level *HOXA1* transcription, we observe only the HOXB cluster to be expressed (Figs. S4a–b; 6b). These and other developmental genes (Fig. 6) were observed to maintain exon methylation post-gene silencing, in effect establishing a transcriptional "memory." These transcription factor genes may be uniquely protected during re-establishment of the methylome in the inner cell mass of early embryo and the establishment of the cell line. Further, hESCs are known to be more epiblast-like than naïve mESCs (Liao et al., 2015), so it remains possible that an examination of a more naïve hESC line would yield additional candidate genes exhibiting transcriptional "traces" at earlier developmental windows.

How persistent is transcription linked exon methylation after transcription is reduced or stopped? Why do we see this phenomenon enriched at developmental TFs? These are important questions for future studies, however, study of evolutionary sequence changes have already demonstrated that "pro-epigenetic" selection has functioned to preferentially preserve CG sites in coding regions of HOX and other master developmental genes, as well as the retention of CGI clusters near these genes (Branciamore et al., 2010). Indeed, exon methylation, not promoter or CGI methylation, is over evolutionary time the most highly conserved feature of the DNA methylation system, conserved in

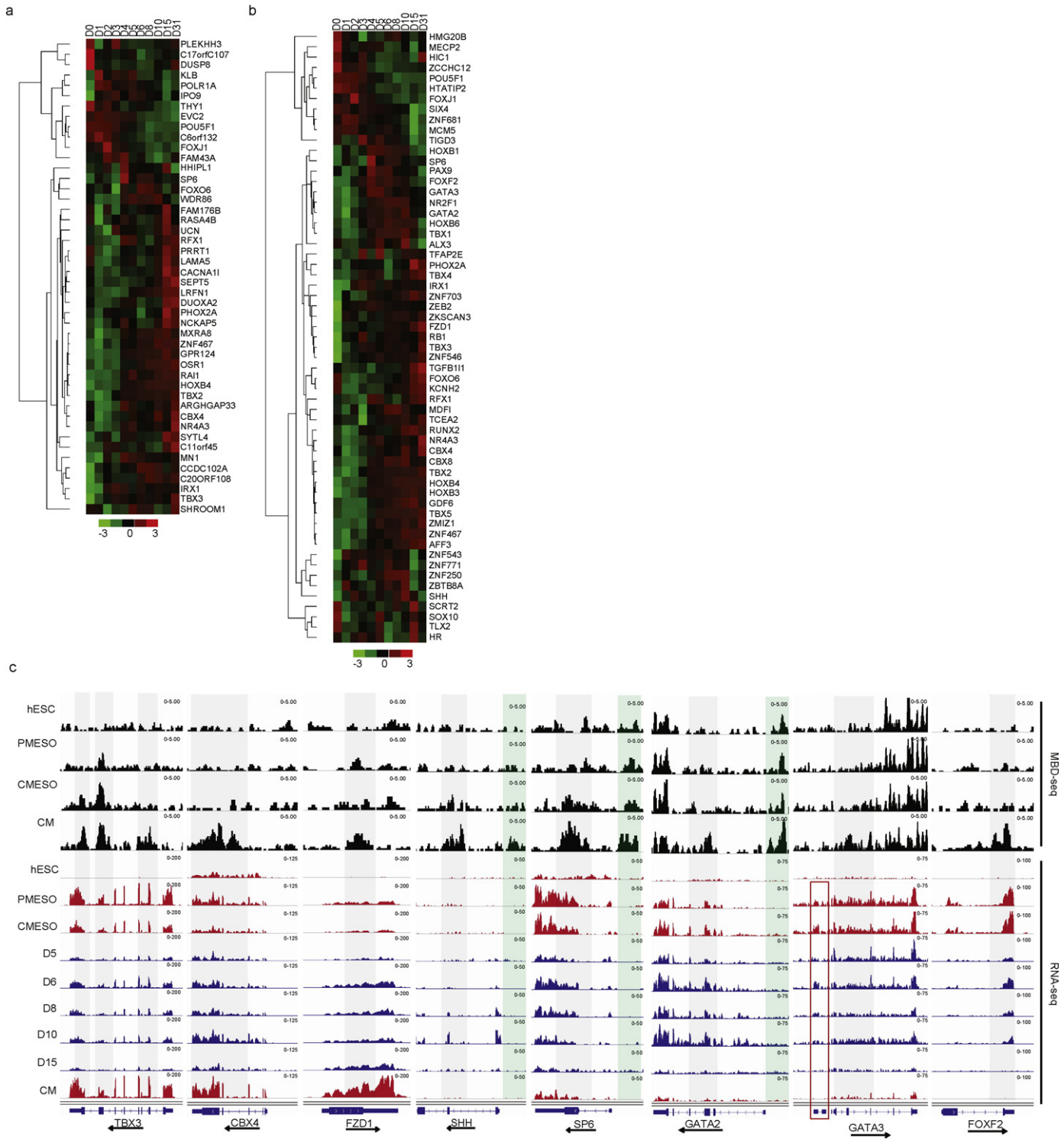


Fig. 6. Exon methylation as a transcriptional "trace." a) 46 high stringency CM exon methylated genes were clustered by expression. 2/48 candidate genes had FPKM values <1 at every sampled time point and were removed. b) CM exon hypermethylated DMRs functionally enriched for *regulation of transcription*. 7/66 were filtered out prior to clustering (1 pseudogene and 6 FPKM < 1 at all sampled time points). Hierarchical clustering by correlation with centroid linkage, gene ordering by peak time (dChIP, 2010.01; (Li, 2008)). Scale = standardized expression level. c) Examples of exon methylation as a "memory" of developmental transcription history. *GATA3*, *FOXF2*, *SHH*, *SP6*, and *GATA2* are repressed by CM stage, the latter 3 being promoter hypermethylated. Gray = regions of intragenic methylation gains; green = promoter regions. Blue RNA-seq data = CMESO-CM intervening time points. Red box = *GATA3* associated lncRNA expression.

both plants and animals, even among most insects, including honey bees where it is involved in the control of queen bee development (Lyko et al., 2010). Further, TF binding sites are rich within gene bodies, tend to be GC rich, often overlap CGIs, and many TFs display distinct methyl DNA binding affinities (Deaton and Bird, 2011). This raises the possibility that exon methylation "memories" serve to maintain global TF distributions, support chromatin networks, and additionally regulate cryptic promoters or enhancers. For the latter possibility, these elements are also enriched among CG-rich gene bodies and highlight a long-standing idea that intragenic DNA methylation may reduce

transcriptional noise, including antisense transcription (Deaton and Bird, 2011; Illingworth et al., 2010). Although this does not explain all intragenic methylation (Jjingo et al., 2012), like conventional gene promoters, gene body promoters are subject to DNA methylation silencing. Therefore, an attractive possibility is that developmentally important switches can be stabilized by the exon methylation system. For example, transcription through an intragenic, potentially active promoter, will result is methylation of that promoter and thereby prevent inappropriate antisense and sense transcription. Such a system seems to have value for establishing stable cell states.

DNA methylation epigenetic "memories" have been described previously, but the focus was on promoter methylation. For example, during iPSC generation the incomplete erasure of somatic epigenetic marks, including promoter DNA methylation, constitute a residual "memory" of somatic tissue origins. As these "memories" are lost with additional iPSC passaging, it is likely that they simply reflect incomplete iPSC reprogramming (Kim et al., 2010). Well before these concepts, however, and still widely accepted today, DNA methylation is thought to "lock-in" stably silent states at promoters and repetitive elements (Razin and Riggs, 1980). Though we again observe promoter methylation to be associated with gene silencing (Figs. 4c, 5a–c, e–g, 6c) this is not a reflection of preceding transcriptional activity. Importantly, our results suggest that one must now consider exon methylation as a likely player in cellular memory. These "exon memories" reflect preceding transcription and as such raises the possibility that developmental or pathological history might be predicted from exon methylation patterns. Indeed, heart failure is marked by the reactivation of improper fetal developmental gene networks (reviewed by Tompkins and Riggs, 2015) and the hijacking of developmental programs is a common feature of many cancer types. For example, HOX cluster CGI hypermethylation marks patients with coronary heart disease (Nazarenko et al., 2015) and is frequently observed in cancer (Shah and Sukumar, 2010). Insight into the transcriptional disease etiology through "exon memory" studies may prove to be quite powerful in understanding and treating these conditions.

In closing, the replacement of lost CMs will be required for cardiac regenerative medicine strategies and the scalability of pluripotent cell sources, including suspension cultures, offers an unlimited cell replacement source in cardiac medicine. This naturally hinges on researcher controlled CM differentiation, which continues to improve, and as highlighted in this report also provides an essential in vitro model of human cardiac muscle development. Here, the multi-stage global analysis of hESC differentiation provides a comprehensive understanding of CM generation through the lenses of transcription and epigenetic patterns and led to observations of transcription "memories" in CM products.

Data generated from this study is available through NCBI's Gene Expression Omnibus (GEO) through GEO series accession number GSE76525 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE76525). Authors declare no conflict of interest.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ebiom.2016.01.021.

Acknowledgments

We would like to thank City of Hope's Analytical Cytometry and Integrative Genomics Core. This work was supported in part by CIRM TG2-01150.

References

Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., Schubeler, D., 2015. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature 520, 243–247.

Branciamore, S., Chen, Z.X., Riggs, A.D., Rodin, S.N., 2010. CpG island clusters and pro-epigenetic selection for CpGs in protein-coding exons of HOX and other transcription factors. Proc. Natl. Acad. Sci. U. S. A. 107, 15485–15490.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., Rinn, J.L., 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 25, 1915–1927.

Carpenter, L., Carr, C., Yang, C.T., Stuckey, D.J., Clarke, K., Watt, S.M., 2012. Efficient differentiation of human induced pluripotent stem cells generates cardiac cells that provide protection following myocardial infarction in the rat. Stem Cells Dev. 21, 977–986.

Chakraborty, S., Deb, A., Maji, R.K., Saha, S., Ghosh, Z., 2014. LncRBase: an enriched resource for lncRNA information. PLoS One 9, e108010.

Chen, V.C., Couture, S.M., Ye, J., Lin, Z., Hua, G., Huang, H.I., Wu, J., Hsu, D., Carpenter, M.K., Couture, L.A., 2012. Scalable GMP compliant suspension culture system for human ES cells. Stem Cell Res. 8, 388–402.

Chinchilla, A., Lozano, E., Daimi, H., Esteban, F.J., Crist, C., Aranega, A.E., Franco, D., 2011. MicroRNA profiling during mouse ventricular maturation: a role for miR-27 modulating Mef2c expression. Cardiovasc. Res. 89, 98–108.

Chong, J.J., Yang, X., Don, C.W., Minami, E., Liu, Y.W., Weyers, J.J., Mahoney, W.M., van Biber, B., Cook, S.M., Palpant, N.J., Gantz, J.A., Fugate, J.A., Muskheli, V., Gough, G.M., Vogel, K.W., Astley, C.A., Hotchkiss, C.E., Baldessari, A., Pabon, L., Reinecke, H., Gill, E.A., Nelson, V., Kiem, H.P., Laflamme, M.A., Murry, C.E., 2014. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. Nature 510, 273–277.

Chung, S., Dzeja, P.P., Faustino, R.S., Perez-Terzic, C., Behfar, A., Terzic, A., 2007. Mitochondrial oxidative metabolism is required for the cardiac differentiation of stem cells. Nat. Clin. Pract. Cardiovasc. Med. 4 (Suppl. 1), S60–S67.

Deacon, D.C., Nevis, K.R., Cashman, T.J., Zhou, Y., Zhao, L., Washko, D., Guner-Ataman, B., Burns, C.G., Burns, C.E., 2010. The miR-143-adducin3 pathway is essential for cardiac chamber morphogenesis. Dev. 137, 1887–1896.

Deaton, A.M., Bird, A., 2011. CpG islands and the regulation of transcription. Genes Dev. 25, 1010–1022.

Drukker, M., Tang, C., Ardehali, R., Rinkevich, Y., Seita, J., Lee, A.S., Mosley, A.R., Weissman, I.L., Soen, Y., 2012. Isolation of primitive endoderm, mesoderm, vascular endothelial and trophoblast progenitors from human pluripotent stem cells. Nat. Biotechnol. 30, 531–542.

Evseenko, D., Zhu, Y., Schenke-Layland, K., Kuo, J., Latour, B., Ge, S., Scholes, J., Dravid, G., Li, X., Maclellan, W.R., Crooks, G.M., 2010. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 107, 13742–13747.

Fatica, A., Bozzoni, I., 2014. Long non-coding RNAs: new players in cell differentiation and development. Nat. Rev. Genet. 15, 7–21.

Forne, T., Oswald, J., Dean, W., Saam, J.R., Bailleul, B., Dandolo, L., Tilghman, S.M., Walter, J., Reik, W., 1997. Loss of the maternal H19 gene induces changes in Igf2 methylation in both cis and trans. Proc. Natl. Acad. Sci. U. S. A. 94, 10243–10248.

Freire, A.G., Resende, T.P., Pinto-Do-Ó, P., Tua, 2014. Building and repairing the heart: what can we learn from embryonic development? BioMed. Res. Int. 2014, 8.

Fujimoto, K.L., Clause, K.C., Liu, L.J., Tinney, J.P., Verma, S., Wagner, W.R., Keller, B.B., Tobita, K., 2011. Engineered fetal cardiac graft preserves its cardiomyocyte proliferation within postinfarcted myocardium and sustains cardiac function. Tissue Eng. A 17, 585–596.

Goll, M.G., Bestor, T.H., 2005. Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. 74, 481–514.

Gu, Y., Liu, G.H., Plongthongkum, N., Benner, C., Yi, F., Qu, J., Suzuki, K., Yang, J., Zhang, W., Li, M., Montserrat, N., Crespo, I., del Sol, A., Esteban, C.R., Zhang, K., Izpisua Belmonte, J.C., 2014. Global DNA methylation and transcriptional analyses of human ESC-derived cardiomyocytes. Protein Cell 5, 59–68.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., Glass, C.K., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589.

Horie, T., Ono, K., Nishi, H., Iwanaga, Y., Nagao, K., Kinoshita, M., Kuwabara, Y., Takanabe, R., Hasegawa, K., Kita, T., Kimura, T., 2009. MicroRNA-133 regulates the expression of GLUT4 by targeting KLF15 and is involved in metabolic control in cardiac myocytes. Biochem. Biophys. Res. Commun. 389, 315–320.

Huang, Z.-P., Wang, D.-Z. miR-22 in cardiac remodeling and disease. Trends in Cardiovascular Medicine.

Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R., Bird, A.P., 2010. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet. 6, e1001134.

Jin, B., Li, Y., Robertson, K.D., 2011. DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes Cancer 2, 607–617.

Jjingo, D., Conley, A.B., Yi, S.V., Lunyak, V.V., Jordan, I.K., 2012. On the presence and role of human gene-body DNA methylation. Oncotarget 3, 462–474.

Jurkowska, R.Z., Jurkowski, T.P., Jeltsch, A., 2011. Structure and function of mammalian DNA methyltransferases. Chembiochem 12, 206–222.

Kehat, I., Kenyagin-Karsenti, D., Snir, M., Segev, H., Amit, M., Gepstein, A., Livne, E., Binah, O., Itskovitz-Eldor, J., Gepstein, L., 2001. Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. J. Clin. Invest. 108, 407–414.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I., Yabuuchi, A., Takeuchi, A., Cunniff, K.C., Hongguang, H., Mckinney-Freeman, S., Naveiras, O., Yoon, T.J., Irizarry, R.A., Jung, N., Seita, J., Hanna, J., Murakami, P., Jaenisch, R., Weissleder, R., Orkin, S.H., Weissman, I.L., Feinberg, A.P., Daley, G.Q., 2010. Epigenetic memory in induced pluripotent stem cells. Nature 467, 285–290.

Klattenhoff, C.A., Scheuermann, J.C., Surface, L.E., Bradley, R.K., Fields, P.A., Steinhauser, M.L., Ding, H., Butty, V.L., Torrey, L., Haas, S., Abo, R., Tabebordbar, M., Lee, R.T.,

Burge, C.B., Boyer, L.A., 2013. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell 152, 570–583.
Kohli, R.M., Zhang, Y., 2013. TET enzymes, TDG and the dynamics of DNA demethylation. Nature 502, 472–479.
Kovacic, J.C., Mercader, N., Torres, M., Boehm, M., Fuster, V., 2012. Epithelial-to-mesenchymal and endothelial-to-mesenchymal transition: from cardiovascular development to disease. Circulation 125, 1795–1808.
Lev Maor, G., Yearim, A., Ast, G., 2015. The alternative role of DNA methylation in splicing regulation. Trends Genet. 31, 274–280.
Li, C., 2008. Automating dChip: toward reproducible sharing of microarray data analysis. BMC Bioinform. 9, 231.
Lian, X., Zhang, J., Azarin, S.M., Zhu, K., Hazeltine, L.B., Bao, X., Hsiao, C., Kamp, T.J., Palecek, S.P., 2013. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/beta-catenin signaling under fully defined conditions. Nat. Protoc. 8, 162–175.
Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., Pop, R., Reyon, D., Tsai, S.Q., Mallard, W., Joung, J.K., Rinn, J.L., Gnirke, A., Meissner, A., 2015. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat. Genet. 47, 469–478.
Lindskog, C., Linne, J., Fagerberg, L., Hallstrom, B.M., Sundberg, C.J., Lindholm, M., Huss, M., Kampf, C., Choi, H., Liem, D.A., Ping, P., Varemo, L., Mardinoglu, A., Nielsen, J., Larsson, E., Ponten, F., Uhlen, M., 2015. The human cardiac and skeletal muscle proteomes defined by transcriptomics and antibody-based profiling. BMC Genomics 16, 475.
Loirand, G., Guerin, P., Pacaud, P., 2006. Rho kinases in cardiovascular physiology and pathophysiology. Circ. Res. 98, 322–334.
Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., Maleszka, R., 2010. The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol. 8, e1000506.
Macarthur, B.D., Lachmann, A., Lemischka, I.R., Ma'ayan, A., 2010. GATE: software for the analysis and visualization of high-dimensional time series expression data. Bioinformatics 26, 143–144.
Marco, A., Macpherson, J.I., Ronshaugen, M., Griffiths-Jones, S., 2012. MicroRNAs from the same precursor have different targeting properties. Silence 3, 8.
Maunakea, A.K., Chepelev, I., Cui, K., Zhao, K., 2013. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res. 23, 1256–1269.
Morselli, M., Pastor, W.A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A.T., Ottonello, S., Jacobsen, S.E., Pellegrini, M., 2015. In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. Elife 4, e06205.
Mummery, C.L., Zhang, J., Ng, E.S., Elliott, D.A., Elefanty, A.G., Kamp, T.J., 2012. Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. Circ. Res. 111, 344–358.
Nan, X., Campoy, F.J., Bird, A., 1997. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell 88, 471–481.
Nazarenko, M.S., Markov, A.V., Lebedev, I.N., Freidin, M.B., Sleptcov, A.A., Koroleva, I.A., Frolov, A.V., Popov, V.A., Barbarash, O.L., Puzyrev, V.P., 2015. A comparison of genome-wide DNA methylation patterns between different vascular tissues from patients with coronary heart disease. PLoS One 10, e0122601.
Paige, S.L., Thomas, S., Stoick-Cooper, C.L., Wang, H., Maves, L., Sandstrom, R., Pabon, L., Reinecke, H., Pratt, G., Keller, G., Moon, R.T., Stamatoyannopoulos, J., Murry, C.E., 2012. A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. Cell 151, 221–232.
Papadopoulos, G.L., Reczko, M., Simossis, V.A., Sethupathy, P., Hatzigeorgiou, A.G., 2009. The database of experimentally supported targets: a functional update of TarBase. Nucleic Acids Res. 37, D155–158.
Papait, R., Kunderfranco, P., Stirparo, G.G., Latronico, M.V., Condorelli, G., 2013. Long noncoding RNA: a new player of heart failure? J. Cardiovasc. Transl. Res. 6, 876–883.
Parmacek, M.S., Epstein, J.A., 2013. An epigenetic roadmap for cardiomyocyte differentiation. Circ. Res. 112, 881–883.
Razin, A., Riggs, A.D., 1980. DNA methylation and gene function. Science 210, 604–610.
Robertson, C., Tran, D.D., George, S.C., 2013. Concise review: maturation phases of human pluripotent stem cell-derived cardiomyocytes. Stem Cells 31, 829–837.
Salomonis, N., Nelson, B., Vranizan, K., Pico, A.R., Hanspers, K., Kuchinsky, A., Ta, L., Mercola, M., Conklin, B.R., 2009. Alternative splicing in the differentiation of human embryonic stem cells into cardiac precursors. PLoS Comput. Biol. 5, e1000553.
Sdek, P., Zhao, P., Wang, Y., Huang, C.J., Ko, C.Y., Butler, P.C., Weiss, J.N., Maclellan, W.R., 2011. Rb and p130 control cell cycle gene silencing to maintain the postmitotic phenotype in cardiac myocytes. J. Cell Biol. 194, 407–423.
Shah, N., Sukumar, S., 2010. The Hox genes and their roles in oncogenesis. Nat. Rev. Cancer 10, 361–371.
Smith, Z.D., Meissner, A., 2013. DNA methylation: roles in mammalian development. Nat. Rev. Genet. 14, 204–220.
Synnergren, J., Ameen, C., Lindahl, A., Olsson, B., Sartipy, P., 2011. Expression of microRNAs and their target mRNAs in human stem cell-derived cardiomyocyte clusters and in heart tissue. Physiol. Genomics. 43, 581–594.
Takebayashi, S., Tamura, T., Matsuoka, C., Okano, M., 2007. Major and essential role for the DNA methylation mark in mouse embryogenesis and stable association of DNMT1 with newly replicated regions. Mol. Cell. Biol. 27, 8243–8258.
Tompkins, J., Riggs, A., 2015. An epigenetic perspective on the failing heart and pluripotent-derived-cardiomyocytes for cell replacement therapy. Front. Biol. 11–17.
Tompkins, J.D., Hall, C., Chen, V.C., Li, A.X., Wu, X., Hsu, D., Couture, L.A., Riggs, A.D., 2012. Epigenetic stability, adaptability, and reversibility in human embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 109, 12544–12549.
Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., Bartel, D.P., 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 147, 1537–1550.
Veloso, A., Kirkconnell, K.S., Magnuson, B., Biewen, B., Paulsen, M.T., Wilson, T.E., Ljungman, M., 2014. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. Genome Res. 24, 896–905.
Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., Erwin, G., Kattman, S.J., Keller, G.M., Srivastava, D., Levine, S.S., Pollard, K.S., Holloway, A.K., Boyer, L.A., Bruneau, B.G., 2012. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. Cell 151, 206–220.
Wang, L., Zhao, Y., Bao, X., Zhu, X., Kwok, Y.K., Sun, K., Chen, X., Huang, Y., Jauch, R., Esteban, M.A., Sun, H., Wang, H., 2015. LncRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. Cell Res. 25, 335–350.
Zhang, Z., Xiang, D., Heriyanto, F., Gao, Y., Qian, Z., Wu, W.S., 2013. Dissecting the Roles of miR-302/367 Cluster in Cellular Reprogramming Using TALE-based Repressor and TALEN. Stem Cell Rep. 1, 218–225.