bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Nkx2.5-dependent alterations of the embryonic heart DNA methylome identify novel cis-regulatory elements in cardiac development

Bushra Gorsi[1] Timothy L. Mosbruger[1], Megan Smith[1], Jonathon T.Hill[1,2], H. Joseph Yost[1*]

1. Molecular Medicine Program, Eccles Institute of Human Genetics, University of Utah

15 North 2030 East, University of Utah, Salt Lake City, Utah, USA 84112-5330

2. current address: Department of Physiology & Developmental Biology, Brigham Young University, 3018 LSB, Provo, UT 84602

*Correspondence:

jyost@genetics.utah.edu

Running Title: DNA methylome maps of Nkx2.5 mutant hearts

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Transcription factor Nkx2.5 is frequently mutated in congenital heart disease, but the mechanisms by which Nkx2.5 regulates heart development are poorly understood. By generating comprehensive DNA methylome maps from zebrafish embryonic hearts in *nxk2.5* mutants and siblings, we discovered that Nkx2.5 regulates DNA methylation patterns during cardiac morphogenesis. We identified hundreds of Nkx-dependent heart-specific Differentially Methylated Regions (nhDMRs). A majority of the nhDMRs were hypomethylated in *nkx2.5*$^{-/-}$ hearts, correlating with changes in the mutant transcriptome, suggesting Nkx2.5 functions largely as a repressor. Distinct Nkx DNA-binding motifs were significantly enriched in subclasses of nhDMRs. Furthermore, nhDMRs were significantly associated with histone H3K4me1 and H3K27ac post-translational modifications, suggesting Nkx2.5 regulates gene expression by differential methylation of cis-regulatory elements. Using transgenics, we validated several nhDMRs with enhancer activities in the heart. We propose a novel role of Nkx2.5 mediated DNA methylation is integral in activating and repressing Nkx2.5 target genes during heart development.

Keywords: Heart; Nkx2.5; regulatory elements; Zebrafish; DNA methylation; transcriptome; epigenetics; cardiac; transcriptome; congenital heart disease

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Nkx2.5 is one of the most commonly mutated transcription factors in human Congenital Heart Disease (CHD), occurring in 1-4% of cardiac malformations (1). Nkx2.5 is a dual activator and repressor required for cardiac progenital cell (CPC) specification, chamber growth, and formation and maintenance of the conduction system (2). Although its function is critical during early cardiomyogenesis, our understanding of how Nkx2.5 can both activate and repress gene function is limited. In zebrafish, Nkx2.5 has a role in heart tube extension and chamber morphogenesis by maintaining normal ratios of ventricular and atrial cells (3,4). The ratio of atrial to ventricle cells is skewed to significantly more atrial cells, and the ventricle is diminished. Although *nkx2.5* is not necessary to induce ventricle initiation, it is required for the maintenance of ventricle identity from early to late stages of heart development (5). In the absence of *nkx2.5*, ventricular cells transdifferentiate to atrial cells (3), suggesting Nkx2.5 is functioning to repress the atrial gene program to maintain ventricle identity. Nkx2.5 also has been shown to be an important mediator of endothelial cell differentiation in pharyngeal arch (PAA) development, as Nkx2.5 actively represses angioblast differentiation early in the anterior lateral plate mesoderm but is required for PAA angioblast differentiation in later stages (6). The dual role of Nkx2.5 to function as a transcriptional activator and repressor in early lineage commitment decisions is not unique to zebrafish, but is also conserved in mice. Partial reduction or a complete loss of Nkx2.5 from an *Nkx2.5* hypomorphic mouse heart and Nkx2.5 knockout in ESC-derived cardiomyocytes, showed a large proportion of differentially genes are upregulated and have Nkx2.5 binding motifs at their loci, lending support to the hypothesis that Nkx2.5 functions as a repressor (7,8). In mouse ESC-derived Nkx2.5 knockout cardiomyocytes, CPCs differentiate quicker than the wildtype CPC, suggesting Nkx2.5 may act early to repress differentiation genes, such as Bmp2/Smad1, to maintain the CPC population (8,9). During stages of cardiac maturation in *Nkx2.5* hymorphic mice, there is a direct repression of cardiac fast troponins isoforms in the atrium, but not the ventricle, another example of Nkx2.5 mediated lineage specific repression (7). It is well established that Nkx2.5 cooperativity with other transcription factors such as Tbx5, Gata4, Tbx2, Tbx20,

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SRF and Meis1 facilitates its activating and repressing functions (7,10–12). This largely occurs through the conserved homeodomain, which serves as both a sequence-specific DNA binding domain and a protein-protein interaction domain. Molecular analysis of the homeodomain suggest a key regulatory role of Nkx2.5 in repressing gene functions through the direct interaction and/or recruitment of epigenetic modifiers such as Histone deacetylases (HDACs) that deacetylate histone (12). However HDACs are thought to not function alone and operate with pathways that involve DNA methyltransferases (DNMTs). Nkx2.5 can also repress transcription of target genes through direct interaction with histone H3K36me3 methyltransferase (13), this histone mark has also been shown to serve as a mark for HDACs (14) and DNMT3B (15). Moreover, DNMTs can recruit HDACs directly, suggesting methylation and deacetylation act together to establish a repressed state (16,17).

Direct epigenetic configuration of chromatin might be one of the mechanisms by which Nkx2.5 can function as a dual activator and repressor in embryonic heart development. We sought to address whether Nkx2.5 might serve as a functional repressor through the evolutionary conserved process of DNA methylation. Methylation of cytosines is thought to prevent gene transcription by directly interfering with the recognition sequence of some TFs, while specifically binding other TFs. Emerging studies have identified homeobox proteins including the Nkx family of proteins (Nkx2.2 and Nkx2.5) as interacting with methylated DNA via their homeodomian as well as binding to methylated CpGs in TF binding motif (18–20). Thus analysing the DNA methylome in Nkx2.5 mutants will provide a gateway for studying the interactions of Nkx2.5 with the DNA methylome.

Using a genomics approach, we generated comprehensive DNA methylome maps and RNA-seq profiles from isolated 48hpf zebrafish hearts and non-heart tissue from both *nkx2.5*$^{-/-}$ mutants and siblings. We discovered thousands of Nkx2.5-dependent heart-specific Differentially Methylated Regions (nhDMRs), and found significant enrichment of the nhDMRs at distal cis-regulatory regions. A majority of the nhDMRs were hypomethylated in the *nkx2.5*$^{-/-}$ hearts, positively correlating with a large number of

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

differentially upregulated genes in *nkx2.5*$^{-/-}$ mutants, corroborating the idea that Nkx2.5 acts largely as a repressor of gene function. nhDMRs neighboring the differentially genes were enriched in Nkx motifs. Our data shows for the first time a class of cardiac-specific distal and proximal regulatory elements in zebrafish that are dependent on Nkx2.5 regulation, and tests a selection of these elements for their ability to regulate cardiac expression in stable transgenic lines. The hundreds of Nkx2.5 dependent regulatory elements will serve as a resource for further functional exploration of coding and non-coding elements that contribute to cardiac disease. Furthermore, we propose that Nkx2.5 imparts a level of regulation by altering DNA methylation patterns in the heart.

RESULTS

nkx2.5 mutant hearts have differentially methylated DNA regions compared to sibling hearts

We performed whole genome-wide Bisulfite-seq and RNA-seq analysis at 48 hours post fertilization (hpf) to identify changes in the DNA methylome and transcriptome of *nkx2.5*$^{-/-}$ mutants that might contribute to the abberant heart development. We manually dissected *nkx2.5*$^{-/-}$ heart tissue and sibling heart tissue from whole embryos and collected DNA from three replicates of the following four samples: *nkx2.5*$^{-/-}$ heart, *nkx2.5*$^{-/-}$ body, sib heart and sib body. DNA samples had a successful bisulfite conversion rate of 99.9%. Using BisSeq (Useq v.8.8.5 package) we analyzed the changes in DNA methylation in *nkx2.5*$^{-/-}$ heart versus sib heart, compared to mutant and wildtype non-heart tissue. We imposed quality filters of ≥8 reads in a window and 5 ≥ CpGs in a region. Using a cutoff criterion of 1.5-fold change and $p < 0.05$, we found 3702 unique differentially methylated regions (DMRs) with 914 hypermethylated DMRs in the *nkx2.5*$^{-/-}$ heart and 2788 hypomethylated DMRs in the *nkx2.5*$^{-/-}$ heart (Figure 1A). Nkx dependent heart-specific DMRs (nhDMRs) were defined as DMRs that did not overlap with DMRs observed in the *nkx2.5*$^{-/-}$ heart v *nkx2.5*$^{-/-}$ body and sib heart v sib body comparisons. A large proportion of the nhDMRs were

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

located within 10kb of the nearest TSS for both hypermethylated (41%) and hypomethylated (37%) nhDMRs (Figure 1B) accompanied with the highest log2ratio fold change and percent changes in DNA methylation (Figure 1C). A moderate percent change in nhDMRs was observed at short and long range distances to TSS. By comparison, percent changes in DMRs observed in the *nkx2.5*$^{-/-}$ heart v *nkx2.5*$^{-/-}$ body and sib heart v sib body were smaller and a shorter range (Figure S1A-B). Average length of subclasses of *nkx2.5*$^{-/-}$ nhDMRs was 250bp, similar to DMRs in non-heart tissue (Figure 1D).

We next examined the nhDMRs in the context of hypomethylation, partial methylation and hypermethylation in the *nkx2.5*$^{-/-}$ heart. In order to define hypo and hyper methylation boundaries, we examined genes that encode cardiac transcription factors, ion channels, endocardial and epicardial genes that are expressed or silenced in our *nkx2.5*$^{-/-}$ RNA-seq dataset (summarized in heatmaps, Figure S2A-B). Methylation status of these genes showed negative correlation with gene expression. Therefore, we took the average fraction methylation at the TSS and defined hypomethylated regions as ≤ 0.3, partially methylated regions as ranging from ≥ 0.3 to ≤ 0.7, and hypermethylated regions as ≥ 0.7. Percentage of DNA methylation of individual CpGs showed clear separation between hypomethylation, hypermethylation and partial methylation (Figure 1E) of DMRs between *nkx2.5*$^{-/-}$ heart and sib hearts. High percent changes of methylation were associated with both classes of nhDMRs with high and low CpG densities, suggesting nhDMRs are not exclusively at promoters (Figure 1F).

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Distribution of DMRs show significant enrichment at the promoter in *nkx2.5* mutant heart

We next examined the genomic locations of *nkx2.5*-dependent nhDMRs in *nkx2.5*$^{-/-}$ mutant hearts. We defined filtered nhDMRs as nhDMRs that had greater than 20% methylation change. Using the Cis-regulatory Element Annotation System (CEAS), we found that genomic distribution of filtered nhDMRs was highly nonrandom. The largest percentage of filtered nhDMRs were observed in intergenic and intronic regions. We found a significantly lower percentage of intronic nhDMRs, but a greater percentage of intergenic nhDMRs than expected by random chance (Figure 2A, red asterisk). Approximately 2.6% of hypomethylated nhDMRs and 2.0% of hypermethylated nhDMRs were within the proximal promoter, greater than predicted by chance distribution throughout the genome (Figure 2B). The data suggests that promoters were enriched as main targets of DNA methylation, as one would expect since CpG islands are largely located at the TSS and that regulations of nhDMRs is likely to occur proximal to the promoters of genes.

Nkx2.5-dependent nhDMRs are highly conserved in vertebrates

To test whether both classes of nhDMRs, hypermethylated and hypomethylated, were evolutionarily conserved we obtained the average PhastCon scores based on an eight-way fish and vertebrate genome alignment with zebrafish using the UCSC browser. We found higher conservation scores for both hypermethylated and hypomethylated nhDMRs than for neighboring regions when aligned to multiple fish genomes (Figure 2C) and multiple vertebrate genomes (Figure 2D). In particular, a significant proportion of hypomethylated nhDMRs contained sequences that were evolutionarily conserved across vertebrates and fish (permutation test, p-value <0.001). The high level of significant conservation of sequences at the center of nhDMRs suggests these nhDMRs have functional significance across vertebrates, and will serve as a valuable resource for the searches in human whole genomes for non-coding variants that might be causative of CHD. However, lack of conservation in an nhDMRs does not rule out its regulatory function (21).

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Subsets of both classes of nhDMRs show regulatory function

To evaluate the functional potential of both classes of nhDMRs (hypomethylated and hypermethylated), we intersected our data with 48hpf whole embryo ChIP-seq data enriched for the following active histone post-translational modifications (PTMs): H3K4me1, H3K4me3 and H3K27ac(22). Distal and proximal regulatory elements have the capability of turning on gene expression independent of location and orientation relative to the gene of interest. In turn, this makes identifying distal and proximal regulatory elements particularly challenging. However, various experimental and computational approaches have identified signature histone PTMs, H3K4me1 and H3K27ac, that define regulatory elements (23–25). Using these criteria, we determined whether the nhDMRs were correlated with regulatory elements of nearby genes. Strikingly, we observed high ChIP-seq signals for H3K4me1, H3K4me3 and H3K27ac PTMs centered over hypomethylated nhDMRs (Figure 3A,C,E). More than half of the hypomethylated nhDMRs intersected with single, double and multiple active histone PTMs with a significant number associated with the triple histone H3K4me3; H3K4me1; H3K27ac (Fisher test, p-value <0.001) and double histone H3K4me1; H3K27ac active PTMs (Fisher test, p-value < 0.001), indicating these nhDMRs elements are correlated with active regulatory function. Conversely, the H3K4me3 signal was lower at the center of hypermethylated nhDMRs than in neighboring regions, but there was increased H3K4me1 and H3K27ac PTMs centered at the hypermethylated nhDMRs (Figure 3B,D). We identified a significant number of hypermethylated nhDMRs also associated with double histone PTMs (H3K4me1; H3K27ac, Fisher test, p-value 5.55×10^{-84}) (Figure 3F). Considering the evolutionary conservation of both classes of nhDMRs and the striking integration with histone PTM patterns, we identified a number of heart regulatory elements that are differentially methylated in *nkx2.5*$^{-/-}$ heart.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Hypomethylated nhDMRs are enriched for binding motifs of Nkx and other heart transcription factors

Several studies have shown DNA demethylation in intergenic regions of non-coding transcription can define active intergenic regulatory elements. Combining this with chromatin status can define functional classes of elements. Recent work comparing Nkx2.5 and p300 ChIP-seq data in mouse heart tissue identified Nkx2.5 as better predictor of cardiac enhancers (7). Nkx2.5 ChIP-seq studies have been carried out in immature and mature mouse cardiomyocyte cell lines, as well as mouse cadiac tissue (7,8,26,27). Several Nkx2.5 binding motifs have been proposed that show degeneracy in several base pairs amongst different studies. Zebrafish Nkx2.5 amino acid sequence closely resembles the human and mouse Nkx2.5 homologs, however, we find strong conservation of the homeodomain among all zebrafish Nkx family members, indicating all available Nkx motifs should be assessed (Figure S3A-B). To date no Nkx2.5 binding motif has been identified in zebrafish due to lack of availability of suitable ChIP-grade antibodies. To test for Nkx motifs in our subclasses of nhDMRs, we included all Nkx binding motif sequences from the Homer database, and loaded additional novel experimentally validated Nkx2.5 binding sites into the list of heart motifs. In addition, we compiled a list of all known Nkx2.5 binding partners from the literature (2,7,8,10,11,26) and extracted their known binding motifs from the Homer database (Figure S3C). We took all the filtered (>20% change in methylation) hypomethylated nhDMRs and used Homer to scan for enriched motifs. The predominant Nkx motif, observed in 1299 hypomethylated nhDMRs, had closer resemblance to Nkx6.1 motif than any other Nkx motif (Figure 4A). Since all Nkx family members share a conserved homeodomain, the most parsimonious explanation is that this is sequence of the more frequently used binding motifs for Nkx2.5 in zebrafish. Interestingly, for the hypermethylated class of nhDMRs, we observed a different Nkx motif with closer resemblance to Nkx2.1/2.5 motif and with a different combination of cardiac TF frequently co-clustering with this particular motif (Figure S4A-B).

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Our Homer analysis found a striking number of significant cardiac TF motifs present among the Top 10 enriched motifs (Figure 4A, for full list see Table S2). We found Isl1, Gata-4 and other Nkx motifs to be significantly co-clustered with Nkx6.1 motif in the hypomethylated nhDMRs (Figure 4B). In contrast, Isl1,Tbx5 and Meis1 motifs co-clustered with the Nkx2.1/2.5 motif in hypermethylated nhDMRs. The presence of Isl1 in both categories of nhDMRS suggest a unique role for Isl1 in Nkx2.5-mediated heart development. Nkx2.5 and Isl1 positive cardiac progenitor cells give rise to multiple cardiac lineages. Although Nkx2.5 is shown to repress Isl1 in mESCs (28). Our data suggests a level of cooperativity between these transcription factors in regulating methylation of the nhDMRs. Further clustering of the nhDMRs by these distinct motif patterns shows evidence of Nkx2.5 dependendent co-regulation of different nhDMRs most likely in different subpopulations of cardiomyocytes (Figure S5A-B). Future work detailing the characterisation of the different clusters of nhDMRs *in-vivo* will yield greater insight into Nkx2.5 driven gene regulatory networks.

Nkx motif-containing nhDMRs neighboring cardiac genes are associated with active histone marks

We next asked whether the Nkx motif nhDMRs can act as regulatory elements to the nearby genes. We intersected all 1299 Nkx motif-containing hypomethylated nhDMRs with annotated active histone ChIP-seq marks and found significant enrichment of H3K4me1, H3K4me3 and H3K27ac centered around the Nkx motif nhDMRs (Figure 4C,D). Interestingly, we also observed H3K4me1 and H3K27ac centered around Nkx motif hypermethylated nhDMRs although with markedly less association of H3K4me3 PTM (Figure S4C-D).

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To evaluate the molecular pathways potentially regulated by the Nkx-motif hypomethylated nhDMRs, we examined the neighboring genes. Using Ingenuity Pathway Analysis (IPA)(29), we determined the pathways most enriched in the nearby genes. Two of the top networks were cardiovascular and muscle development (Figure S6A). Several unique pathways in cardiac development were enriched, including cardiac ventricle formation, looping morphogenesis and vascular smooth muscle formation (Figure 4E), strongly resonating with the *nkx2.5*$^{-/-}$ zebrafish phenotype (3,4,30). Similar to the hypermethylated nhDMRs, we found cardiovascular system to be one of the major pathways associated with hypomethylated nhDMRs. Interestingly we observed a strong association with the endocardial genes, suggesting Nkx2.5 dependent gene regulation plays an important role in vascular signalling pathways (Figure S4E-F). We then tested if any of the Nkx motif hypomethylated nhDMRs were associated with altered expression levels of genes nearby. We found the Nkx intergenic nhDMRs had moderate but significant negative correlation with expression of the nearby genes, in particular, histone marks associated with intergenic and intronic regions (Figure S6B-C). This suggests Nkx motif hypomethylated nhDMRs exert regulatory roles on the expression of nearby genes.

Nkx2.5 affects diverse signaling pathways during heart development

To determine whether nhDMRs contribute to the regulation of differential gene expression in Nkx2.5 mutants, we performed RNA-seq analysis on triplicate samples of *nkx2.5*$^{-/-}$ and sib hearts at 48 hpf. We observed altered gene expression patterns of a moderate number of genes (Figure 5A,B). We found that a majority of the differentially expressed genes were upregulated (112 genes, adj p-value <0.05) versus downregulated (29 genes, adj p-value <0.05), suggesting *nkx2.5* is an activator of a small subset of genes but acts predominantly as a repressor. Genes that were downregulated included *vmhcl* (ventricular myosin heavy chain) and genes encoding calcium ion binding proteins and intermediate filament proteins, in agreement with the proposed function of Nkx2.5 in the maintenance of ventricle identity. Genes that were upregulated were enriched in several pathways including myofibril assembly, cardiac and smooth muscle

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

contraction, consistent with the view that Nkx2.5+ progenitors cells can give rise to multiple cardiac cell lineages inlcuding conduction cell and smooth muscle cells (31). However these hypotheses are largely untested in zebrafish *nkx2.5*$^{-/-}$ mutants. Interestingly, we observed a cohort of upregulated genes not characterized in the context of Nkx2.5 heart development, including mical2b, tjp-1b, c-fos, junba and c-jun. Mical2b is a nuclear actin regulatory protein, morphants of which show looping defects and reduced expression of cardiac and smooth actin proteins (32). Although Tjp-1b is expressed in the heart (33), its functional role has not been elucidated. C-fos and Junba, proto-oncogenes that are components of the AP-1 TF complex, are upregulated in response to cardiac hypertrophy and negatively regulate atrial natriuretic factor (ANF) (34), a target of Nkx2.5. C-jun has shown to be essential for the development of the outflow tract (35). Motif analysis on Nkx2.5 differentially expressed genes, did not identify any significant overrepresentation of Nkx motifs. This prompted us to investigate whether the differentially expressed genes showed strong correlation to the DNA methylation status of their TSS. Average fraction methylation profiles of differentially expressed genes were plotted and clustered by gene body DNA methylation from TSS to TSE (Figure S7A,B). We found no distinct differences between *nkx2.5*$^{-/-}$ heart and sib heart in the overall gene body DNA methylation of these misexpressed genes. In addition gene expression was weakly correlated with respective DNA methylation level at their TSS (Figure S7C). This led us to ask whether neighboring nhDMRs could potentially regulate Nkx2.5-differentially expressed genes.

nhDMRs neighboring differentially expressed genes are enriched in Nkx motif

Nkx2.5 has been shown to bind and regulate genes via both proximal and distal promoters (36). In the developing mouse heart, selective expression of ANF is regulated by Nkx2.5 responsive cis-regulatory elements (36). To test if differentially expressed genes were near regulatory elements identified by our *nkx2.5*$^{-/-}$ DNA methylome map, we took each nhDMR nearest to each differentially expressed gene and categorized it as a DMR in promoter, exon, intron or intergenic region. Of the 95 genes that are upregulated (excluding mitochondrial genes, since we observed no nhDMRs on mitochondrial genome),

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

we identified 72 genes adjacent to a hypomethylated nhDMRs and 23 genes adjacent to an hypermethylated nhDMRs. Majority of the nhDMRs were intergenic. Using Homer, we found that Nkx(2.1/2.5) motifs were significantly enriched in a subset of the nhDMRs (Figure 6A, (p-val < 0.05)). Clustering of binding sites have successfully been used to scan the genome sequence to predict tissue specific enhancers; using these criteria we observed all Nkx motif-containing nhDMRs showed co-clustering of transcription factor binding motifs (Figure S8A). We calculated the average distance (bp) between Nkx enriched motif and other motifs in the nhDMRs. Using this list, we ranked the Nkx motif-containing nhDMRs in order of average distance (bp) from smallest to largest. We identified several other Nkx family members as the most frequently observed motifs that clustered with Nkx enriched motif (Figure S8B). This gave us a list of nhDMRs likely to be functionally responsive to Nkx2.5 regulation in-vivo.

To evaluate their functional potential, we identified a moderate number of nhDMRs that were significantly associated with either double (H3K4me1 and H3K27ac) or triple histone marks (H32k4me1, H3K4me3 and H3K27ac), with ChIP-seq signals enriched over the center of the nhDMRs (Figure 6B,C). To examine if the nhDMRs neighboring the upregulated genes showed potential regulatory function we looked for correlation with the differentially upregulated genes. Interestingly, we observed significant correlation with all promoter and intergenic nhDMRs with nearby gene expression (Figure 6D). Most notably, hypermethylated nhDMRs had strong correlation to nearby Nkx differentially expressed genes, suggesting a different mode of action for Nkx from that found in hypomethylated nhDMRs. Although we observed several strong negative correlations with the Nkx motif nhDMRs associated with a histone mark, these were not statistically significant due to the small number identified.

Nkx motif nhDMRs regulatory elements validated in vivo

We sought to test whether the nhDMRs that we identified could function as enhancers of heart development. We selected nhDMRs near genes associated with heart development (*nkx2.5, cited2, foxc1a*), genes encoding epigenenetic modification factors (*dnmt4,tet3*), and genes differentially

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expressed in the *nkx2.5*$^{-/-}$ mutants from our transcriptome analysis (*fosb, myh7bb, tjp1b,smtnl1,myocd,cacna1*). We cloned these nhDMRs sequences into reporter constructs with a minimal promoter driving green fluorescent protein (EGFP) and injected zebrasfish embryos with the enhancer reporter vector along with Tol2 transposase messenger mRNA, to drive integration and stable transgenesis. Each group of G_0 embryos injected with putative enhancer reporter constructs showed unique expression profiles in the embryo, with all showing expression in the heart. We established founder G_1 transgenic fish that had offspring embryos expressing GFP, indicative of stable germline transgenics. We compared the GFP expression profiles to the endogenous whole mount in-situ expression of the genes documented in Zebrafish Model Organism Database (ZFIN), and found that the transgenic GFP expression patterns (Figure 7A-D) recapitulated the expression patterns of the endogenous candidate gene from which the neighboring nhDMR was derived. The seven nhDMRs we tested all matched their nearby candidate gene expression profiles. Although we paired each nhDMRs to its closest putative target gene, this does not rule out the possibility that the identified nhDMRs might regulate additional target genes whose patterns differ slightly from their endogenous gene expression. All nhDMR transgenics showed moderate to strong GFP expression in the heart, in particular the ventricle, suggesting that Nkx2.5 may play an essential role in regulating these elements in heart development. Although we tested a small subset of our identified nhDMRs, this approach paves the way for additional testing of putative nhDMR enhancer elements and their role in Nkx2.5 dependent heart development.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Discussion

DNA methylation is altered in *nkx2.5*$^{-/-}$ mutant zebrafish embryonic hearts, indicating that Nkx2.5 is required for normal DNA methylation patterns in cardiac development. By generating comprehensive genome-wide DNA methylome maps at base-pair resolution in both wildtype and *nkx2.5*$^{-/-}$ mutant zebrafish embryonic hearts, comparing these to DNA methylation patterns in non-heart embryonic tissues, we have identified novel Nkx2.5-dependent heart-specific Differentially Methylated Regions (nhDMRs). Integrating the dataset with the RNA-seq transcriptomes in mutant hearts, and wildtype histone PTM ChIP-seq data, we created a list of putative enhancers dependent on Nkx2.5 regulation in heart development. We further validated these *in vivo* by creating stable transgenic lines that show expression in heart.

Does Nkx2.5 have a direct role in regulating the methylation status of nhDMRs? The most parsimoniuous explanation is that Nkx2.5 directly binds nhDMRs containing Nkx motifs and participates either in the upregulatulation or downregulation of the methylation status of the nhDMRs to influence the expression of neighboring genes. In support of this explanation, it is striking that each class of nhDMRs, hypomethylated and hypermethylated, were enriched for distinct combinations of Nkx binding motifs and potential co-regulatory factors. Hypomethylated nhDMRs showed strong enrichment of the Nkx 6.1 motif and subset of other core heart transcription factors motifs (Gata4, Tbx5, Meis1, and Isl1). In contrast, hypermethylated nhDMRs were enriched for Nkx2.1/2.5 motifs. These observations suggest hypermethylated and hypomethylated nhDMRs have distinct modes of regulation by Nkx2.5 and cofactors. Importantly, both classes of nhDMRs were found to be conserved across vertebrate taxa, strongly suggesting widespread functional potential in cardiac development.

Our RNA-seq data corroborated earlier findings of Nkx2.5 roles in ventricle formation (3,4). We also observed that several genes important in heart development, including *bmp4, vcana, vmhcl and gata4*, were misregulated in *nkx2.5*$^{-/-}$ hearts as previously reported. Intersection of the nhDMRs with ChIP-seq data identified significant association with H3K4me1 and H3K27ac PTM marks. Both marks are strongly

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

linked to active regulatory elements, such as enhancers. Our transgenic assay provides strong evidence that nhDMRs function as enhancers in the heart and other tissues and show that changes in DNA methylation of the enhancer elements play a role in Nkx2.5 mediated heart development.

Since we observed a significant association of a subset of nhDMRs with unique Nkx binding motifs, we hypothesize that Nkx2.5 transcription factor has two roles: First, in the context of hypomethylated nhDMRs observed in *nkx2.5*$^{-/-}$ mutants**,** we propose loss of Nkx2.5 acts to enhance expression of nearby differentially expressed genes most likely through a reduction in DNA methylation at the regulatory element (Figure 8A). Second, in the presence of Nkx2.5, the expression of the nearby gene is downregulated likely through increased DNA methylation of the regulatory element via the recruitment of cofactors (Figure 8B). Nkx2.5 comes from a family of homeobox proteins which can bind to methylated CpG via their homeodomain (19). So it is conceivable that Nkx2.5 could bind directly to CpG methylated DNA and foster the recruitment of co-factors to remove DNA methylation. In addition Nkx2.5 might also influence DNA methylation indirectly, by regulating the expression of DNA methylation machinery. Although Nkx2.5 has not been shown to directly bind at the promoter of HDAC1, we observe in multiple published Nkx2.5 ChIP-seq datasets (8,26) that Nkx2.5 binds at promoters of DNMT3a, TET1, TET2 and TET3 in several mouse derived cardiomyocyte cell lines. It is conceivable that Nkx2.5 binding events are amplified in a homogenous cardiomyocyte cell line which otherwise maybe masked in whole heart tissue experiment and makes it worthwhile investigating the DNA methylation profile of mouse derived cardiomyocyte cell line in response to Nkx2.5 knockout to further corroborate these hypotheses.

When does Nkx2.5 influence the methylation status of nhDMRs? DNA methylation is highly dynamic during cardiomyocyte development, postnatal maturation and disease. In mammals, demethylated regions in neonatal and adult cardiomyocytes are localized in cell type-specific enhancer regions and gene bodies of cardiomyocyte genes (37). Previous characterization of *nkx2.5*$^{-/-}$ zebrafish mutants suggested Nkx2.5

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

functions early, at 36hpf (5), Our RNA-seq datasets of sibling hearts detected high level of Nkx2.5 expression at 48hpf, consistent with our observation of strong enrichment of Nkx motifs in the nhDMRs. Furhermore, our recent Self Organizing Map (SOM) time course analysis of zebrafish wildtype heart development from linear heart tube (30 hpf) to fully looped heart (72 hpf) indicates that *nkx2.5* is expressed at steady state throughout these stages (Hill et al manuscript accepted). Since DNA methylation patterns can be inherited over several cell divisions, it is likely that the methylation status of nhDMRs is established by Nkx2.5 and cofactors early in development and that persistent methylation states in nhDMRs serve a critical function in maintenance of ventricle identity and downstream Nkx-mediated signaling pathways revealed by our analysis.

In summary, our results suggest that Nkx2.5 binding sites in the nhDMRs that we have identified participate in the methylation and demethylation of specific genomic regions. These observations provide a platform for further investigation of heart transcription factors and their participation in the regulation of DNA demethylation. The different transcription factor binding motifs identified in conjunction with Nkx motifs within subsets of hypermethylated and hypomethylated nhDMRs suggests interaction with different cardiac TFs in the regulation of DNA methylation of nhDMRs. This model provides an explanation for the dual nature of Nkx2.5 to act as an activator and repressor during cardiac development. Our study sets the foundation for addressing the functions of Nkx-dependent elements in the heart and to uncover Nkx driven gene regulatory networks in the heart. These nhDMRs co-localize with gene regulatory elements, particularly enhancers and transcription factor-binding sites, which will allow identification of key lineage-specific regulators. Therefore, the highly conserved Nkx motif nhDMRs that we have identified can serve as a starting point to explore new putative regulatory elements in congenital heart disease.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Methods and Materials

Zebrafish stocks

Adult *nkx2.5*$^{+/-}$ were outcrossed to *cmlc2*: GFP fish maintained on AB background. Adult carriers for nkx2.5 mutations were genotyped by High Resolution Melt Analysis (HRMA) (38) with the following primers: Fwd Primer; CAAACTCACCTCCACACAGG, Rev Primer; GCTGCCTCTTGCACTTGTATC. All adult fish were housed in the University of Utah Centralized Zebrafish Animal Resource facility. Animal research protocol 15-06004 was approved by University of Utah IACUC. Embryos were raised under standard conditions at 28ºC. Embryos raised from (*nkx2.5*$^{+/-}$*; cmlc2*: eGFP) crosses were screened at 48hpf when the *nkx2.5*$^{-/-}$ mutant phenotype is identifiable with an enlarged atrium and small ventricle. *nkx2.5*$^{-/-}$ mutants were segregated from siblings that consisted of mixed genotypes (*nkx2.5*$^{+/+}$ and *nkx2.5*$^{+/}$). This nkx2.5 allele has a stop codon in the homeodomain, and heterozytotes do not have a discernable cardiac phenotype. Genotypes were subsequently confirmed in RNA-Seq and whole genome Bisulfite-sequencing datasets.

nkx2.5$^{-/-}$ heart collections and RNA-sequencing

Zebrafish embryos were first segregated by phenotype into *nkx2.5*$^{-/-}$ mutants and sibs at 48 hours post fertilization (hpf) and then hearts were isolated; 150 hearts were collected per biological replicate for each genotype, using a protocol modified from previously published zebrafish heart isolation techniques (39). Isolated phenotyped hearts were washed twice with heart media and hand selected to further separate hearts from non-heart tissue and debris. Heart tissue was snap frozen in liquid nitrogen until sufficient samples were collected. RNA was purified from heart tissue following the QIAGEN Microplus Easy kit protocol. RNA samples were pooled together per replicate in order to collect an optimal amount for library preparation and sequencing. RNA quality and quantities were confirmed on the Bioanalyser RNA 6000 Pico Chip. cDNA libraries were prepared using the Illumina TruSeq kit at the University of Utah High Throughput Genomics Core Facility. All six samples, consisting of three matching *nkx2.5*$^{-/-}$ mutants

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and sibling pairs were run on a single lane and 50bp single end reads were generated on the HiSeq 2000 machine. Samples were aligned to zebrafish Zv9 Ensembl version 77 genome annotation build using Novoalign with default parameters.

RNA-seq analysis

We used DESeq2 default settings to measure differential expression between *nkx2.5*$^{-/}$ heart and sib hearts, controlling for collection time. Gene read counts were generated using USeq "Defined Region Differential Seq" from the USeq package (40) .

Bisulfite-sequencing and analysis

Hearts and non-heart tissue were isolated as described above *nkx2.5*$^{-/-}$ mutants and siblings at 48hpf, and genomic DNA was extracted from each genotype. A total of 100ng was collected from 400 hearts, and corresponding non-heart tissue, using the Qiagen non-organic DNA clean up kit. DNA samples were collected in triplicate, sheared using COVARIS, and spiked with 1% unmethylated lambda. Samples underwent bisulfite treatment using Epitect bisulfite kit and a total of 12 samples were each sequenced on four lanes of HiSeq 2000 to generate 101bp paired end reads. Fastq files for each replicate per genotype and tissue were merged from each run to give one final fastq file per replicate, and aligned to Zv9 using Novoalign. Each replicate per condition had 3x-5x median coverage, and genotypes were confirmed for each sample. Bioinformatics analysis on the Genome wide Bisulfite sequencing used tools from the USeq package (http://useq.sourceforge.net) (40) to calculate differentially methylated regions on merged replicates from each genotype. Each merged replicate had 9x-15x coverage minimum. To test if merged samples provided good representation of the three replicates, we ran PCA on differentially hypomethylated and hypermethylated regions in each condition. Fraction methylation of defined regions in individual replicates clustered close to the merged replicate (averaged fraction methylation)(Figure S1A-B).

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Analysis of nhDMRs

Motif analysis and binding site prediction

Location of transcription factor binding sites and motif enrichments within DMRs were determined using default parameters in Homer (41). Additional Nkx2.5 motifs from recent published datasets (7,8,26) were also incorporated into the search. 50,000 random background sequences normalized for CpG content were chosen for the different subclasses of nhDMRs tested via Homer settings.

Cis-regulatory element analysis (CEAS), evolutionary conservation and GO analyses

The CEAS script build_genomeBG was used to create a custom CEAS annotation file(42). Coverage data was used from all four conditions: *nkx2.5*$^{-/-}$ heart, sib heart, *nkx2.5*$^{-/-}$ body and sib body. Bed files of Differentially Methylated Regions (DMRs) were loaded into the program. Genomic coordinates were downloaded from UCSC zv9 ref gene annotations. For evolutionary conservation of nhDMRs the fish and vertebrate PhastCons score based on the eight by eight way fish and vertebrate genome alignment was obtained from the UCSC Genomer Browser PhastCons conservation track (https://genome.ucsc.edu/). The data was generated by averaging phastcon scores across the DMRs and counting the number of DMRs that have average phastcon score of > 0.5. randomBed was used to create 10,000 random regions of matched size of each DMR. mapBed was used to calculate average phastcon scores across these random regions. P-value was calculated by testing how many of the random regions have more conserved regions than defined DMRs. Gene ontology (GO) analysis of the neighboring genes of nhDMRs was carried out using Igenuity pathway analysis (IPA)(29) and DAVID (https://david.ncifcrf.gov/).

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ChIP-seq analysis

Histone ChIP-seq data of 48hpf whole zebrafish embryos was retrieved from Bogdanovic *et al* 2009 (22) and re-processed using MACs 2.0 software with default settings (43). DMRs that intersected with all three histone peak H3Kme1, H3K4me3 and H3K27ac independently of each other were merged together as one single peak for identification.

Testing cis-regulatory elements in vivo

Eight nhDMR sequences with putative regulatory activity (S4 table) were PCR amplified with specific primers and TOPO-cloned into entry vector pCR8/GW/TOPO of the Gateway cloning system (Invitrogen). The cloned plasmid was then recombined with the destination vector pGW_cfosEGFP (21) to generate the desired reporter plasmid. Tol2 transposase mRNA was transcribed in vitro from pcs2+ using the mMessage mMachine Sp6 kit (Ambion). Embryos at one-cell or two-cell stages were injected with transposase mRNA. The reporter expression patterns were analysed at 24 hpf stage. If 10% of embryos exhibit the consistent GFP expression pattern, the construct was considered as a positive enhancer. The embryos with specific expression patterns were selected and raised to sexual maturity. Sexually mature G0 adults were crossed with AB wild-type strain to obtain germline transmission. Three independent F1 transgenic lines were established for each construct, unless otherwise indicated. F1 embryos were photographed at 48hpf stage.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Declarations

Data Access

All the data described in this paper can be downloaded from https://b2b.hci.utah.edu/gnomex/ RNA-seq experiment ID 246R, and Bisulfite seq experiment ID 324R.

Ethics Statement

All genome sequencing data are fully available without restriction, as indicated above.

Animal Research Protocols

Zebrafish animal research were reviewed and approved by the University of Utah IACUC, protocol number 15-06004.

Funding and Acknowledgements

BG was supported by AHA Postdoctoral Fellowship (12POST12030301). This study was funded by a NHLBI Bench-to-Bassinet Consortium (http://www.benchtobassinet.com) grant to HJY (2UM1HL098160) and sequencing was provided by a core facilities support grant to New England Research Institute (U01 HL098188). We thank Bradley Demarest and Darren Ames for performing Bisulfite seq alignments, Kimara Targoff for discussions and supplying *nkx2.5*$^{-/-}$ mutants and David Nix for bioinformatics discussions. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Reamon-Buettner SM, Borlak J. NKX2-5: an update on this hypermutable homeodomain protein and its role in human congenital heart disease (CHD). Hum Mutat [Internet]. 2010 Nov [cited 2011 Dec 7];31(11):1185–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20725931
2. Akazawa H, Komuro I. Cardiac transcription factor Csx/Nkx2-5: Its role in cardiac development and diseases. Pharmacol Ther [Internet]. 2005 Aug [cited 2011 Aug 9];107(2):252–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15925411
3. Targoff KL, Colombo S, George V, Schell T, Kim S-H, Solnica-Krezel L, et al. Nkx genes are essential for maintenance of ventricular identity. Development [Internet]. 2013;140(20):4203–13. Available from: http://dev.biologists.org/content/140/20/4203%5Cnhttp://dev.biologists.org/content/140/20/4203.abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/24026123
4. Targoff KL, Schell T, Yelon D. Nkx genes regulate heart tube extension and exert differential effects on ventricular and atrial cell number. Dev Biol [Internet]. 2008 Oct 15 [cited 2011 Jun 22];322(2):314–21. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2752039&tool=pmcentrez&rendertype=abstract
5. George V, Colombo S, Targoff KL. An early requirement for nkx2.5 Ensures first and Second heart field ventricular identity and cardiac function into adulthood. Dev Biol [Internet]. Elsevier; 2014 Dec 20 [cited 2014 Dec 29]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/25536398
6. Guner-Ataman B, Paffett-Lugassy N, Adams MS, Nevis KR, Jahangiri L, Obregon P, et al. Zebrafish second heart field development relies on progenitor specification in anterior lateral plate mesoderm and nkx2.5 function. Development [Internet]. 2013;140(6):1353–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23444361

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

7. Dupays L, Shang C, Wilson R, Wood S, Towers N, Mohun T, et al. Sequential Binding of MEIS1 and NKX2-5 on the Popdc2 Gene : A Mechanism for Spatiotemporal Regulation of Enhancers during Cardiogenesis Article Sequential Binding of MEIS1 and NKX2-5 on the Popdc2 Gene : A Mechanism for Spatiotemporal Regulation of Enhanc. CellReports [Internet]. The Authors; 2015;13(1):183–95. Available from: http://dx.doi.org/10.1016/j.celrep.2015.08.065

8. Luna-Zurita L, Stirnimann CU, Glatt S, Kaynak BL, Thomas S, Baudin F, et al. Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. Cell. 2016;164(5):999–1014.

9. Prall OWJ, Menon MK, Solloway MJ, Watanabe Y, Zaffran S, Bajolle F, et al. An Nkx2-5/Bmp2/Smad1 negative feedback loop controls heart progenitor specification and proliferation. Cell [Internet]. 2007 Mar 9 [cited 2011 Jul 11];128(5):947–59. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2092439&tool=pmcentrez&rendertype=abstract

10. Stennard FA, Costa MW, Elliott DA, Rankin S, Haast SJP, Lai D, et al. Cardiac T-box factor Tbx20 directly interacts with Nkx2-5, GATA4, and GATA5 in regulation of gene expression in the developing heart. Dev Biol. 2003;262(2):206–24.

11. Sepulveda JL, Vlahopoulos S, Iyer D, Belaguli N, Schwartz RJ. Combinatorial expression of GATA4, Nkx2-5, and serum response factor directs early cardiac gene activity. J Biol Chem. 2002;277(28):25775–82.

12. Elliott DA, Kirk EP, Schaft D, Harvey RP. NK-2 Class Homeodomain Proteins. Conserved Regulators of Cardiogenesis. Hear Dev Regen. 2010;I(1999):569–97.

13. Nimura K, Ura K, Shiratori H, Ikawa M, Okabe M, Schwartz RJ, et al. A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf-Hirschhorn syndrome. Nature [Internet]. Nature Publishing Group; 2009 Jul 9 [cited 2011 Jun 21];460(7252):287–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19483677

14. Joshi AA, Struhl K. Eaf3 chromodomain interaction with methylated H3-K36 links histone

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

deacetylation to pol II elongation. Mol Cell. 2005;20(6):971–8.

15. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature [Internet]. 2015;520(7546):243–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25607372%5Cnhttp://dx.doi.org/10.1038/nature14176%5Cnhttp://www.nature.com/doifinder/10.1038/nature14176

16. Fuks F, Burgers WA, Brehm A, Hughes-Davies L, Kouzarides T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nat Genet [Internet]. 2000;24(1):88–91. Available from: http://dx.doi.org/10.1038/71750

17. Fuks F, Burgers WA, Godin N, Kasai M, Kouzarides T. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. EMBO J. 2001;20(10):2536–44.

18. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science (80-) [Internet]. 2017;356(6337):eaaj2239. Available from: http://www.sciencemag.org/lookup/doi/10.1126/science.aaj2239

19. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, et al. Dynamic readers for 5-(Hydroxy)methylcytosine and its oxidized derivatives. Cell. 2013;152(5):1146–59.

20. Hu S, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, et al. DNA methylation presents distinct binding sites for human transcription factors. Elife. 2013;2013(2):1–16.

21. Fisher S, Grice EA, Vinton RM, Bessling SL, McCallion AS. Conservation of RET regulatory function from human to zebrafish without sequence similarity. Science (80-) [Internet]. 2006;312(5771):276–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16556802

22. Bogdanović O, Fernandez-Miñán A, Tena JJ, De La Calle-Mustienes E, Hidalgo C, Van Kruysbergen I, et al. Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. Genome Res. 2012;22(10):2043–53.

23. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

modification at human enhancers reflect global cell-type specific gene expression. Nature. 2009;459(7243):108–12.

24. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet [Internet]. 2007;39(3):311–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17277777

25. Rada-iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. developmental enhancers in humans. Nature [Internet]. Nature Publishing Group; 2011;470(7333):279–83. Available from: http://dx.doi.org/10.1038/nature09692

26. He A, Kong SW, Ma Q, Pu WT. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. Proc Natl Acad Sci U S A. 2011;108(14):5632–7.

27. Bouveret R, Waardenberg AJ, Schonrock N, Ramialison M, Doan T, de jong D, et al. NKX2-5 mutations causative for congenital heart disease retain functionality and are directed to hundreds of targets. Elife. 2015;4(JULY2015):1–30.

28. Klinik IM. P LURIPOTENT S TEM C ELLS Direct Nkx2-5 Transcriptional Repression of Isl1 Controls Cardiomyocyte Subtype Identity.

29. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. Bioinformatics. 2014;30(4):523–30.

30. Tu C-T, Yang T-C, Tsai H-J. Nkx2.7 and Nkx2.5 function redundantly and are required for cardiac morphogenesis of zebrafish embryos. PLoS One [Internet]. 2009 Jan [cited 2011 Dec 7];4(1):e4249. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2626283&tool=pmcentrez&rendertype=abstract

31. Wu SM, Fujiwara Y, Cibulsky SM, Clapham DE, Lien C ling, Schultheiss TM, et al. Developmental Origin of a Bipotential Myocardial and Smooth Muscle Cell Precursor in the Mammalian Heart. Cell. 2006;127(6):1137–50.

32. Lundquist MR, Storaska AJ, Liu TC, Larsen SD, Evans T, Neubig RR, et al. Redox modification

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of nuclear actin by MICAL-2 regulates SRF signaling. Cell [Internet]. Elsevier Inc.; 2014;156(3):563–76. Available from: http://dx.doi.org/10.1016/j.cell.2013.12.035

33. Kiener TK, Sleptsova-Friedrich I, Hunziker W. Identification, tissue distribution and developmental expression of tjp1/zo-1, tjp2/zo-2 and tjp3/zo-3 in the zebrafish, Danio rerio. Gene Expr Patterns. 2007;7(7):767–76.

34. Mcbride K, Robitaille L, Tremblay S, Argentin S, Nemer M. fos/jun Repression of Cardiac-Specific Transcription in Quiescent and Growth-Stimulated Myocytes Is Targeted at a Tissue-Specific cis Element. Mol Cell Biol. 1993;13(1):600–12.

35. Eferl R, Sibilia M, Hilberg F, Fuchsbichler A, Kufferath I, Guertl B, et al. Functions of c-Jun in liver and heart development. J Cell Biol. 1999;145(5):1049–61.

36. Warren SA, Terada R, Briggs LE, Cole-Jeffrey CT, Chien W-M, Seki T, et al. Differential role of Nkx2-5 in activation of the atrial natriuretic factor gene in the developing versus failing heart. Mol Cell Biol [Internet]. 2011;31(22):4633–45. Available from: http://mcb.asm.org/content/31/22/4633.long

37. Gilsbach R, Preissl S, Grüning BA, Schnick T, Burger L, Benes V, et al. Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease. Nat Commun [Internet]. 2014;5(May):5288. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4220495&tool=pmcentrez&rendertype=abstract

38. Parant JM, George SA, Pryor R, Wittwer CT, Yost HJ. A rapid and efficient method of genotyping zebrafish mutants. Dev Dyn. 2009;238(12):3168–74.

39. Burns CG, MacRae CA. Purification of hearts from zebrafish embryos. Biotechniques. 2006;40(3):274–82.

40. Nix DA, Courdy SJ, Boucher KM. Empirical methods for controlling false positives and estimating confidence in ChIP-Seq peaks. BMC Bioinformatics. 2008;9:523.

41. Gotea V, Visel A, Westlund JM, Nobrega MA, Pennacchio LA, Ovcharenko I. Homotypic clusters

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of transcription factor binding sites are a key component of human promoters and enhancers. Genome Res. 2010;20(5):565–77.

42. Shin H, Liu T, Manrai AK, Liu SX. CEAS: Cis-regulatory element annotation system. Bioinformatics. 2009;25(19):2605–6.

43. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol [Internet]. 2008;9(9):R137. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2008-9-9-r137

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figures

Figure 1: Distribution of nkx2.5-dependent heart-specific differentially methylated regions (nhDMRs).

(A) Venn diagram of DMRs in *nkx2.5* $^{-/-}$ heart compared to *nkx2.5* $^{-/-}$ body, sib heart and sib body. A larger number of DMRs are present (4710) in *nkx2.5* $^{-/-}$ heart when comparing to *nkx2.5* $^{-/-}$ non-heart tissue versus sib heart (3702) DMRs. DMRs that overlapped by 1bp or more were consolidated and redundancies removed from further analysis. (B) Percentage distribution of hypomethylated (red) and hypermethylated (blue) nhDMRs over distance to the nearest TSS. (C) Percent changes in methylation of nhDMRs from *nkx2.5* $^{-/-}$ v. sib heart plotted against distance to the nearest TSS. (D) Length of DMRs in mutants and siblings, heart and non-heart tissue have a median range of 250bp, comparable across different genotypes and tissue types. (E). nhDMRs in *nkx2.5*$^{-/-}$ v sib heart were categorized into hypermethylation (≥0.7), partial methylation (≥0.3 to <0.7) and hypomethylation (<0.3). The categories were designated based on tissue level expression of the known heart and non heart genes. (F) Hypermethylated and hypomethylated nhDMRs with the greatest changes in methylation are categorized comparably into high and low CpG density nhDMRs.

Figure 2: Genomic distribution *nkx2.5*-dependent hypomethylated and hypermethylated nhDMRs.

(A) Assignment of nhDMRs to genomic regions indicates increased representation in upstream promoters and distal intergenic regions. (B) Overrepresentation of nhDMRs in upstream promoter regions . Significant enrichment of the nhDMRs at the promoter in both hypermethylated and hypomethylated DMRs in the *nkx2.5* $^{-/-}$ heart . * $p < 0.01$ ** $p < 0.001$, *** $p < 0.0001$ (C, D) Sequence conservation of hypomethylated and hypermethylated nhDMRs in (C) multiple fish species and (D) other veterbrates;

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

average phastcon scores were plotted over 10kb of regions centered on hypomethylated nhDMRs (red), hypermethylated nhDMRs (blue), compared to random (black) DMRs.

Figure 3: nhDMRs significantly associate with histone PTM marks representative of enhancer elements.

Averaged ChIP-seq signals of active histone PTM marks from 48hpf embryos were plotted over +/-4kb regions centered on (A) hypomethylated nhDMRs and (B) hypermethylated nhDMRs. (C, D) Heatmaps of ChIP-seq signal over 2kb regions centered on individual (C) hypomethylated nhDMRs and (D) hypermethylated nhDMRs ordered by percent methylation change. (E,F) Venn diagram of the number of intersections of H3K4me1, H3K4me3, and H3K27ac PTM marks with (E) hypermethylated nhDMRs and (F) hypomethylated nhDMRs . *** $p < 0.001$, ** $p<0.01$ using Fisher exact test.

Figure 4: Hypomethylated nhDMRs associate strongly with Nkx family motifs, select cardiac TF motifs and with active histone PTM marks.

(A) Homer analysis on hypomethylated nhDMRs identified enrichment of Nkx motifs and motifs for a subset of other heart transcription factors. Motif identity is indicated in top right hand corner and q-val in the bottom of panel. For a full list of Motifs enriched see Supplementary table S2. (B) Frequency of motifs for heart TF and Nkx family members found in significant Nkx-motif containing hypomethylated nhDMRs. Hypergeometric test, ** p-val < 0.01,***p-val < 0.001. (C) Venn diagram indicating number of hypomethylated nhDMRs that associate with histone PTM marks for active transcription enhancers. Fisher test, ** p-val < 0.01,***p-val < 0.001. (D) ChIP-seq signals of active histone PTM marks from 48hpf embryos were plotted over +/-4kb regions centered on hypomethylated nhDMRs. (E) GO categories of the signaling molecules enriched in the cardiovascular network. Expanded GO analysis of

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the cardiovascular network identifies genes integral to ventricle, atrium and valve formation.

Figure 5: Vascular smooth and cardiac muscle contraction genes are strongly represented among differentially expressed genes in 48hpf *nkx2.5* $^{-/-}$ heart.

(A) Heat map of hierarchical clustering of differentially expressed genes. Genes in red denote upregulation and blue denotes downregulation in 48hpf *nkx2.5* $^{-/-}$ hearts. GO clusters of differentially expressed genes identified by IPA analysis. Cardiac muscle contraction the most enriched GO pathways in gene upregulated in *nkx2.5* $^{-/-}$ heart. (B) Volcano plot of *nkx2.5* differentially expressed genes. Red outlined box highlights the upregulated genes. (C) GO categories of genes upregulated in *nkx2.5* $^{-/-}$ hearts, pertinent to nkx mutant phenotype.

Figure 6: Nkx motif hypomethylated nhDMRs regulate nearby differentially expressed genes.

(A) Top Motifs and consensus sequences observed in nhDMRs neighboring differentially expressed genes upregulated in *nkx2.5* $^{-/-}$ hearts. Nkx motif was observed in 40 of the 95 nhDMRs. (B) Venn diagram intersecting the hDMRs nearest genes upregulated in *nkx2.5* $^{-/-}$ hearts with the indicated H3K4me1, H3K4me3, and H3K27ac PTM marks. Fisher test, *** p-val < 0.001. (C) Heatmaps of ChIP-seq signal and CpG density mapped over 2kb regions centered on nhDMRs nearest to upregulated genes. (D) Percent change in methylation of nhDMRs plotted against the log2fold change in expression of the nearest differentially expressed gene.

Figure 7: *In vivo* validation of Nkx putative enhancers.

(A) Few examples of Nkx motif hypomethylated nhDMRs that serve as regulatory elements of nearby differentially expressed genes. First two rows from top denote DNA sequencing coverage in both samples. Third and fourth rows denote methylation tracks (5 reads or more per CpG) of sib and *nkx2.5* $^{-/-}$

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

heart. Fifth through seventh rows indicate histone PTM ChiP-seq enrichment. Eight row indicates region of nearest Nkx motif hypomethylated nhDMRs. Ninth row lists calculated distance of nhDMRs to nearest differentially expressed gene and last row is a depiction of the direction of transcription of the gene. GFP expression driven by the the nhDMRs that we identified adjacent to the following genes: Nkx2.5, (B, B') DNMT4, (C, C') and Fos-b (D,D'). These nhDMRs appear to function as enhancer elements driving expression in the heart (B'-D'), stronger expression can be detected in the ventricle. V; ventricle, OFT; outflow tract.

Figure 8: Proposed roles of Nkx2.5 in heart-specific DMRs neighboring differentially expressed genes.

Hypomethylated nhDMRs containing Nkx motif are represented by red rectangular boxes located proximally (upstream and downstream) and distally relative to gene transcription start site. (A) In the context of hypomethylated nhDMRs observed in the *nkx2.5*$^{-/-}$ mutant**,** we propose loss of Nkx2.5 acts to enhance expression of nearby differentially expressed gene. (B) In the presence of Nkx2.5 in wildtype embryos, the nhDMRs would increase in DNA methylation either through recruitment of "hypomethylated nhDMRs cofactors" by Nkx2.5 or by a direct role in allowing access of DNA methylation machinery.

Supporting Information

Figure Supplement 1: Clustering of individual replicates per condition and percent methylation changes in Nkx and Sib nhDMRs compared to non-heart tissue. PCA plots on the Fraction methylation of (A) hypomethylated DMR and (B) hypermethylated DMRs in each replicate (black outline circle) and averaged fraction methylation in merged sample (red outline circle). The first two PCA components were plotted. (C) Percent changes in methylation of DMRs from *nkx2.5* $^{-/-}$ heart v. *nkx2.5* $^{-/-}$ body plotted against distance to the nearest TSS. (D) Percent changes in methylation of DMRs

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

from sib heart v. sib body plotted against distance to the nearest TSS. Less than 20% change is indicated in orange/pink and greater than 20% in cyan.

Figure Supplement 2: Gene body methylation of core cardiac and epigenetic markers in *nkx2.5*$^{-/-}$ heart and sib heart. (A) Average fraction methylation summed over entire gene body lengths (normalized) and 100bp beyond each end of gene body. A complete list of core cardiac and epigenetic markers is in the S3 file. (B) Hierarchical clustering of core cardiac and epigenetic markers based on average fraction methylation from TSS to TSE. Blue represents regions of low CpG methylation and orange high CpG methylation. GO categories of genes in several clusters are listed in blue boxes.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Supplement 3: Nkx2.5 protein structural domains and homeodomain (A) A schematic of zebrafish structural protein domains compared to mouse Nkx2.5 protein. YRD domain and carboxy-terminal Nkx2.5 box are missing in zebrafish. *TGA indicates site of mutation in zebrafish homeodomain. (B) Amino acid sequence of zebrafish, mouse and human Nkx family members homeodomains. Yellow indicates conserved amino acids. (C) Transcription factor binding motifs used in Homer analysis for TFs known to interact with Nkx2.5 and all zebrafish Nkx family members. References indicate sources for motif sequences.

Figure Supplement 4: Hypermethylated nhDMRs associate strongly with Nkx family member motifs and with active histone PTM marks. (A) Homer analysis on hypermethylated nhDMRs identified enrichment of Nkx motifs and other heart TF motifs. Significant enriched motif is indicated in the top right hand corner and q-val in the bottom of each panel. For a full list of Motifs enriched see Supplementary table S2. (B) Frequency of heart TF motifs and other Nkx family member motifs found in significant Nkx motif hypermethylated nhDMRs. Hypergeometric test,* p-val<0.05, ** p-val < 0.01,***p-val < 0.001. (C) Venn diagram of number of Nkx motif hypermethylated nhDMRs associated with histone PTM marks associated with active enhancers, Fisher test, ** p-val < 0.01,***p-val < 0.001. (D) ChIP-seq signals of active histone PTM marks from 48hpf embryos were plotted over +/- 4kb regions centered on Nkx motif hypermethylated nhDMRs. IPA analysis of the neighboring genes of Nkx motif hypermethylated nhDMRs. (E) Top networks associated with neighboring genes in hypermethylated DMRs. (F) Expansion of one of the top networks associated with cardiovascular function.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Supplement 5: Distinct clusters of hypomethylated and hypermethylated nhDMRs based on motif patterns

K-means clustering on the motif patterns of hypomethylated nhDMRs (A) and hypermethylated nhDMRs (B) . Each column is a cluster of nhDMRs, where the values is the average occurrence for a motif. A value of 1 (red) means all the nhDMRs in that cluster have that motif, while 0 (blue) means none of the nhDMRs have that motif. The motifs are in represented in rows. The number of clusters (k) being 8.

Figure Supplement 6: Nkx motif containing hypomethylated nhDMRs negatively regulate expression of neighboring cardiac genes (A) Top networks represented with cardiovascular development being among the top networks in this subset of genes. (B) Categorization of the Nkx motif containing hypomethylated nhDMRs categorized into different genomic regions. Nkx motif hypomethylated nhDMRs are predominantly located in intergenic regions. (C) Genomic regions of Nkx motif hypomethylated nhDMRs and their correlation with nearest gene expression. Pearson correlation and p-value is reported on plot.

Figure Supplement 7: Gene body methylation of the differentially expressed genes in *nkx2.5*$^{-/-}$ hearts

(A) Average fraction methylation summed over entire gene body lengths (normalized) and 100bp beyond each end of gene body. TSS and TSE denote transcription start and end. (B) K-means clustering of differentially expressed genes based on average fraction methylation across gene body. Blue represents regions of low CpG methylation and orange high CpG methylation. GO categories listed on the right of heatmaps. (C) Negative correlation of TSS methylation of differentially upregulated genes and their expression in *nkx2.5*$^{-/-}$ heart. Positive correlation of TSS methylation of differentially downregulated genes and their expression in *nkx2.5*$^{-/-}$ heart.

bioRxiv preprint doi: https://doi.org/10.1101/186395; this version posted September 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Supplement 8: Nkx motif patterns in nhDMRs neighboring genes upregulated in *nkx2.*5 mutants.

(A) Heatmap of the motifs arrangement in each of nhDMRs neighboring each differentially upregulated gene in nkx2.5 mutant hearts. All nhDMRs were scaled to 100bp and average distance between motifs was calculated per nhDMRs and scored in a table and plotted on a heatmap (B) Most frequently associated motifs in the Nkx containing nhDMRs *$p < 0.05$.

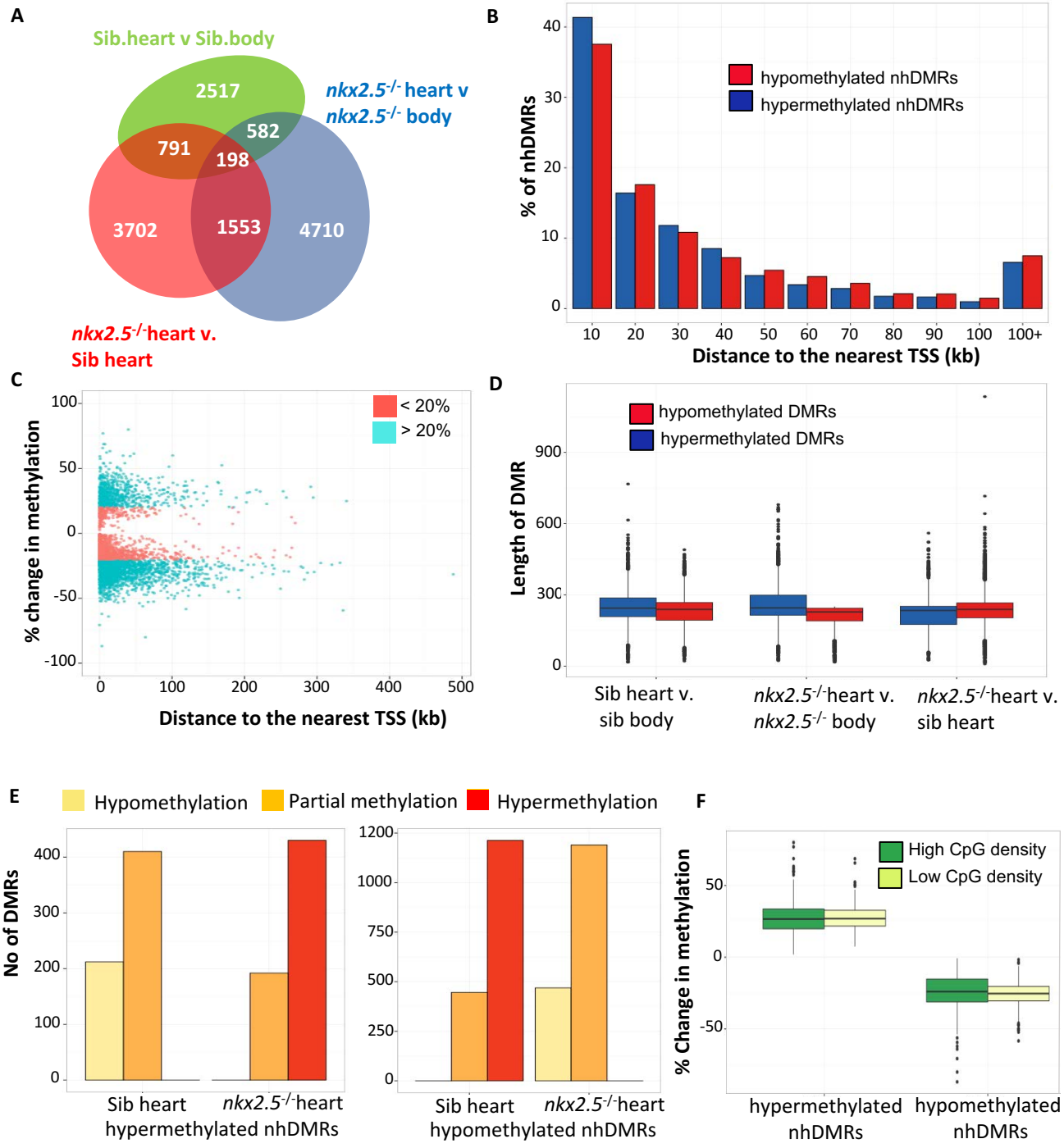

A
Sib.heart v Sib.body
2517
582
791
198
3702
1553
4710
nkx2.5-/- heart v
nkx2.5-/- body
nkx2.5-/-heart v.
Sib heart
B
% of nhDMRs
hypomethylated nhDMRs
hypermethylated nhDMRs
Distance to the nearest TSS (kb)
C
% change in methylation
< 20%
> 20%
Distance to the nearest TSS (kb)
D
Length of DMR
hypomethylated DMRs
hypermethylated DMRs
Sib heart v.
sib body
nkx2.5-/-heart v.
nkx2.5-/- body
nkx2.5-/-heart v.
sib heart
E
Hypomethylation
Partial methylation
Hypermethylation
No of DMRs
Sib heart
hypermethylated nhDMRs
nkx2.5-/-heart
Sib heart
hypomethylated nhDMRs
nkx2.5-/-heart
F
% Change in methylation
High CpG density
Low CpG density
hypermethylated
nhDMRs
hypomethylated
nhDMRs

Figure 1

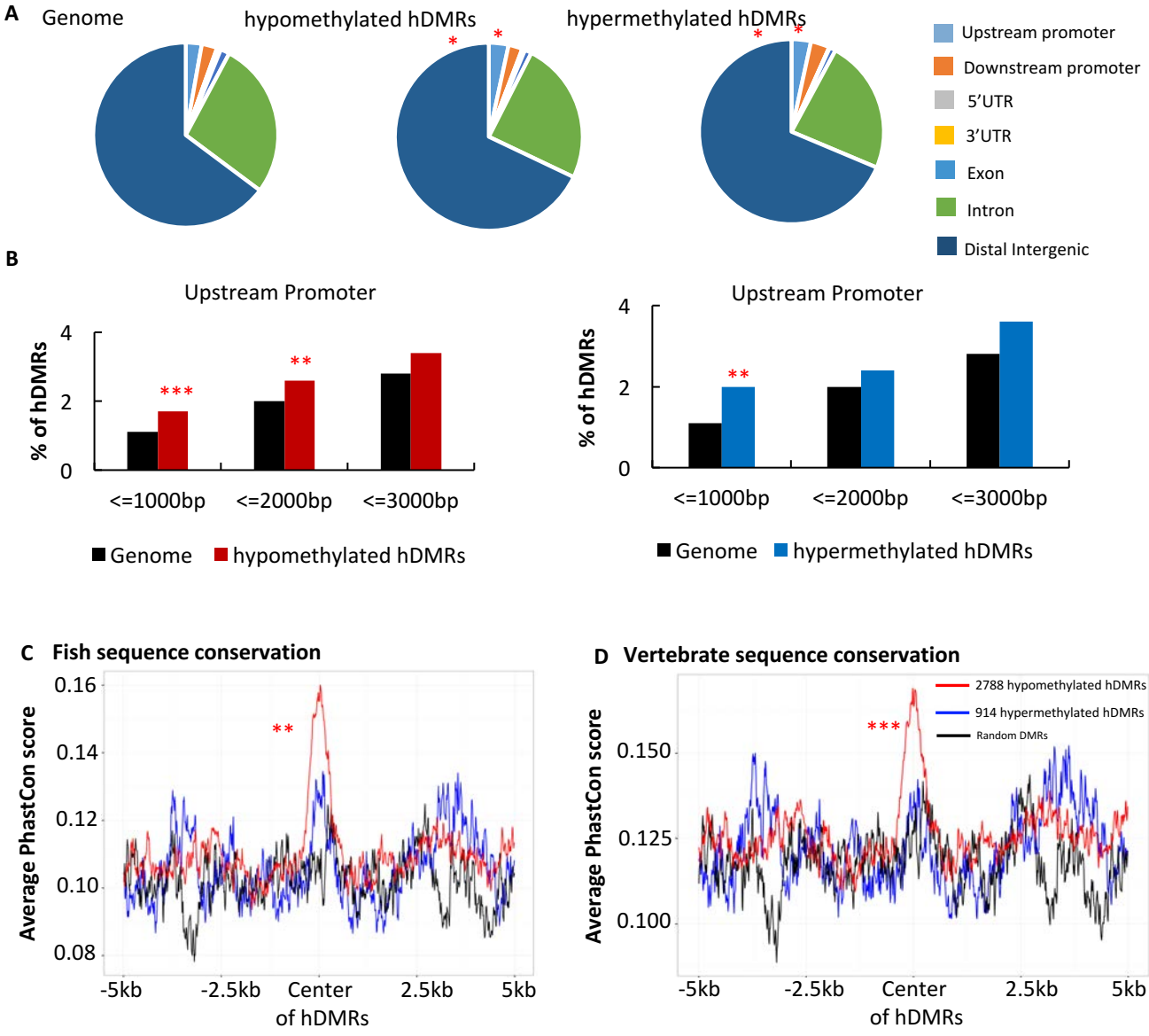

A
Genome
hypomethylated hDMRs
hypermethylated hDMRs
Upstream promoter
Downstream promoter
5'UTR
3'UTR
Exon
Intron
Distal Intergenic
B
Upstream Promoter
% of hDMRs
<=1000bp
<=2000bp
<=3000bp
Genome
hypomethylated hDMRs
Upstream Promoter
% of hDMRs
<=1000bp
<=2000bp
<=3000bp
Genome
hypermethylated hDMRs
C Fish sequence conservation
Average PhastCon score
-5kb
-2.5kb
Center of hDMRs
2.5kb
5kb
D Vertebrate sequence conservation
2788 hypomethylated hDMRs
914 hypermethylated hDMRs
Random DMRs
Average PhastCon score
-5kb
-2.5kb
Center of hDMRs
2.5kb
5kb

Figure 2

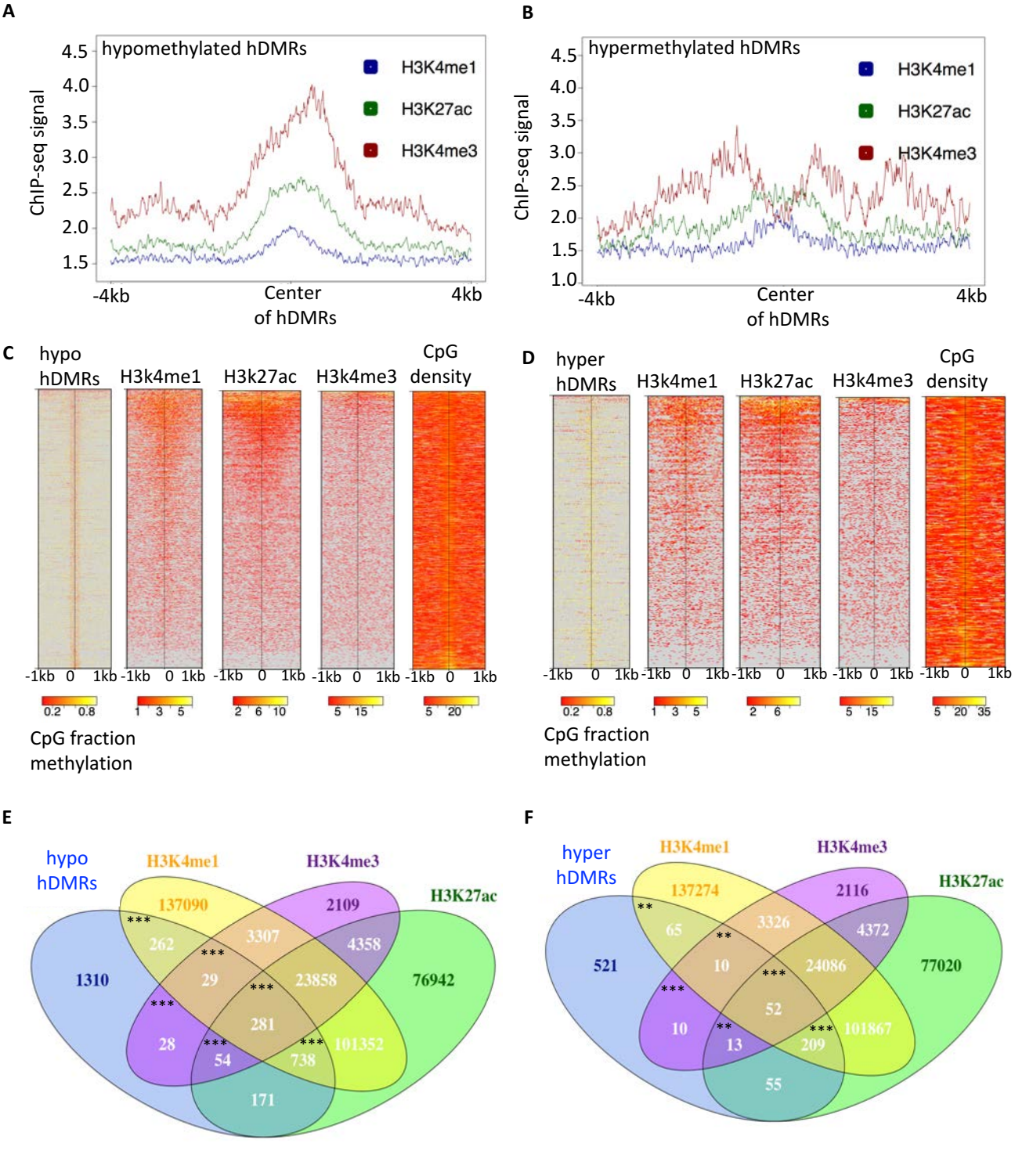
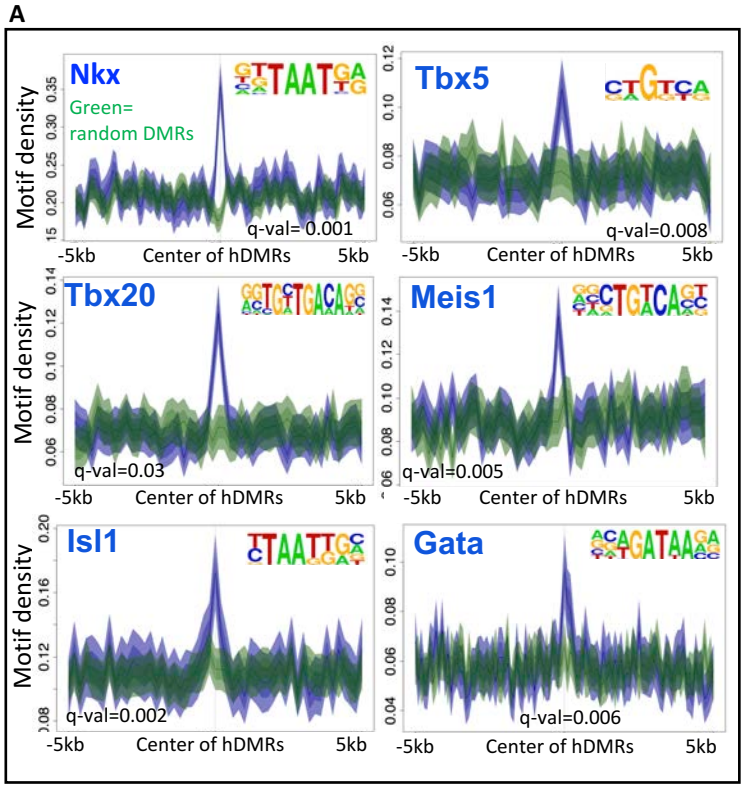
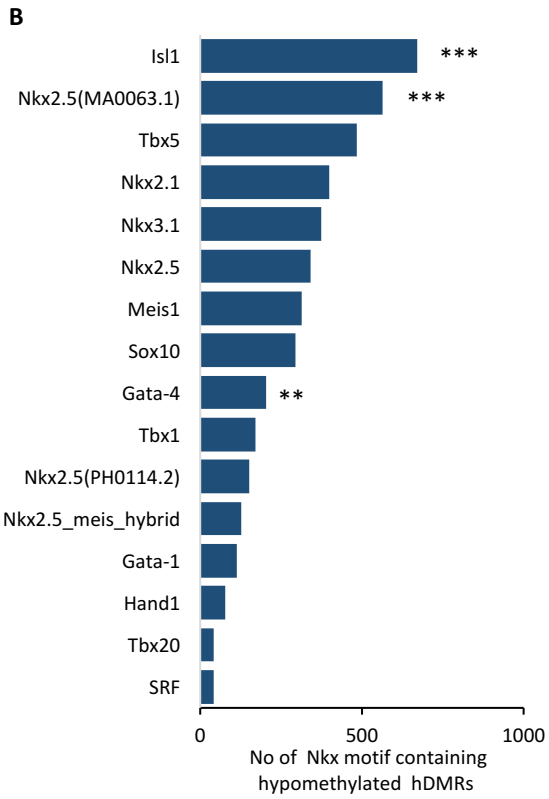
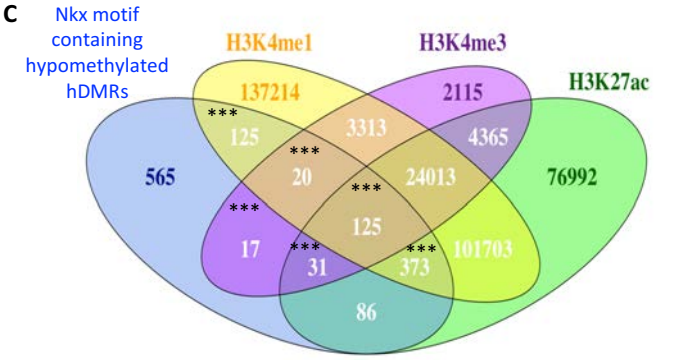
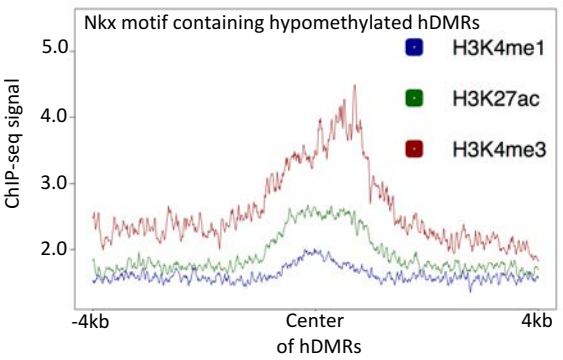

A
hypomethylated hDMRs
H3K4me1
H3K27ac
H3K4me3
ChIP-seq signal
4.5
4.0
3.5
3.0
2.5
2.0
1.5
-4kb
Center
of hDMRs
4kb
B
hypermethylated hDMRs
H3K4me1
H3K27ac
H3K4me3
ChIP-seq signal
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
-4kb
Center
of hDMRs
4kb
C
hypo
hDMRs
H3k4me1
H3k27ac
H3k4me3
CpG
density
-1kb 0 1kb
0.2 0.8
1 3 5
2 6 10
5 15
5 20
CpG fraction
methylation
D
hyper
hDMRs
H3k4me1
H3k27ac
H3k4me3
CpG
density
-1kb 0 1kb
0.2 0.8
1 3 5
2 6
5 15
5 20 35
CpG fraction
methylation
E
hypo
hDMRs
H3K4me1
H3K4me3
H3K27ac
1310
137090
2109
76942
262
3307
4358
29
23858
281
28
54
738
101352
171
***
F
hyper
hDMRs
H3K4me1
H3K4me3
H3K27ac
521
137274
2116
77020
65
3326
4372
10
24086
52
10
13
209
101867
55
**
***

Figure 3


A
Nkx
Green=
random DMRs
Tbx5
Tbx20
Meis1
Isl1
Gata
Motif density
q-val= 0.001
q-val=0.008
q-val=0.03
q-val=0.005
q-val=0.002
q-val=0.006
-5kb
Center of hDMRs
5kb


B
Isl1
Nkx2.5(MA0063.1)
Tbx5
Nkx2.1
Nkx3.1
Nkx2.5
Meis1
Sox10
Gata-4
Tbx1
Nkx2.5(PH0114.2)
Nkx2.5_meis_hybrid
Gata-1
Hand1
Tbx20
SRF
***
***
**
0
500
1000
No of Nkx motif containing
hypomethylated hDMRs


C
Nkx motif
containing
hypomethylated
hDMRs
H3K4me1
H3K4me3
H3K27ac
137214
2115
565
3313
4365
125
20
24013
76992
125
17
31
373
101703
86
***


D
Nkx motif containing hypomethylated hDMRs
H3K4me1
H3K27ac
H3K4me3
ChIP-seq signal
5.0
4.0
3.0
2.0
-4kb
Center
of hDMRs
4kb

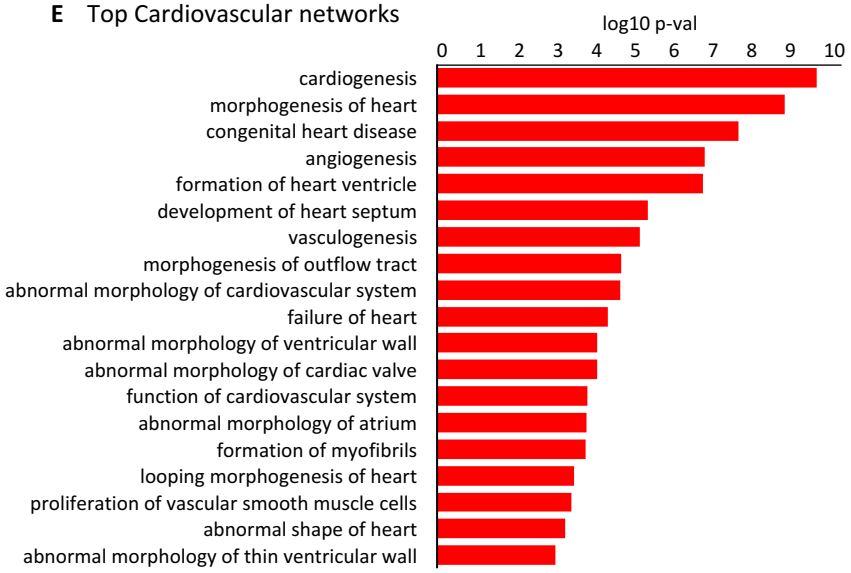

E
Top Cardiovascular networks
log10 p-val
0 1 2 3 4 5 6 7 8 9 10
cardiogenesis
morphogenesis of heart
congenital heart disease
angiogenesis
formation of heart ventricle
development of heart septum
vasculogenesis
morphogenesis of outflow tract
abnormal morphology of cardiovascular system
failure of heart
abnormal morphology of ventricular wall
abnormal morphology of cardiac valve
function of cardiovascular system
abnormal morphology of atrium
formation of myofibrils
looping morphogenesis of heart
proliferation of vascular smooth muscle cells
abnormal shape of heart
abnormal morphology of thin ventricular wall

Figure 4

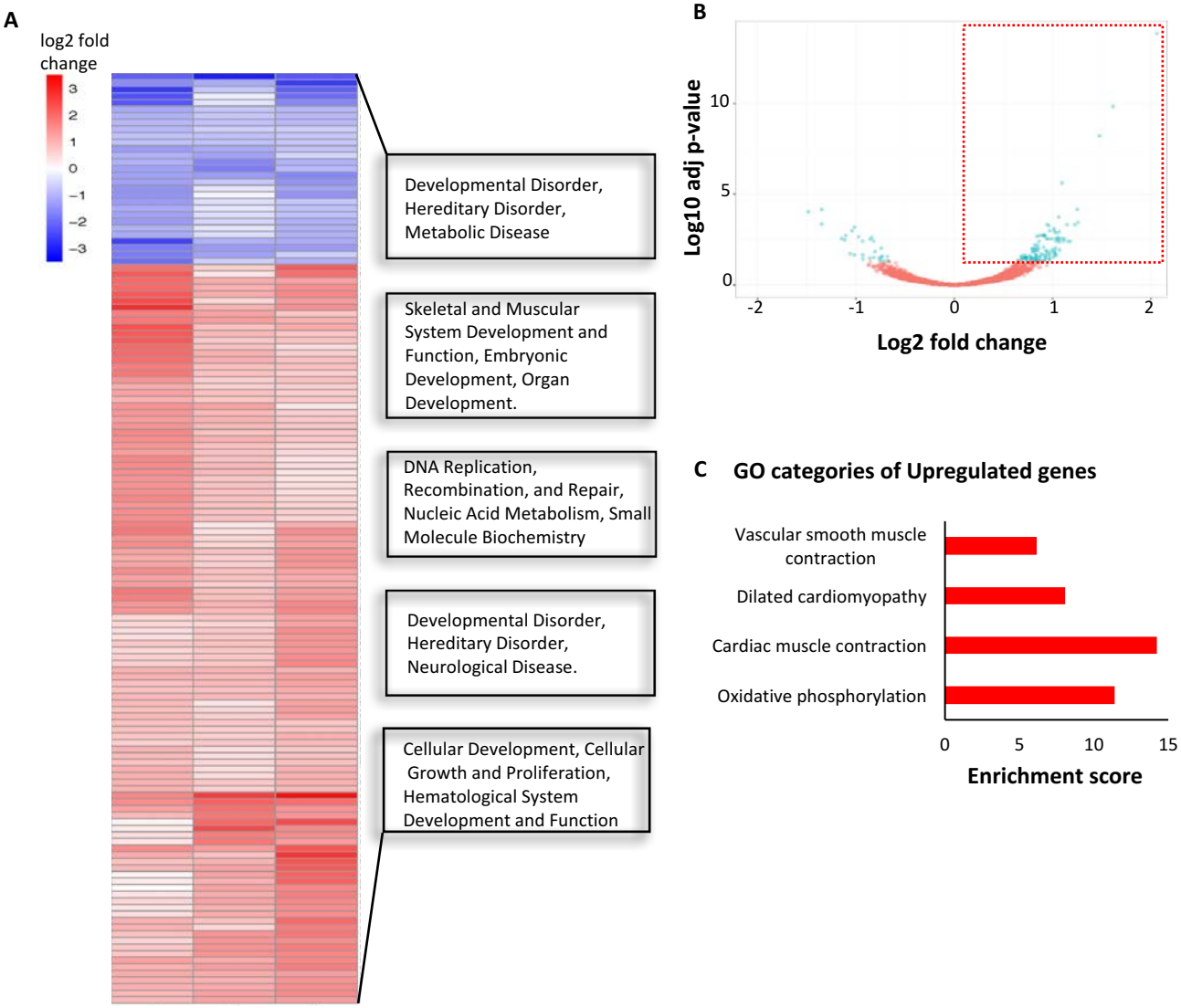
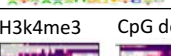
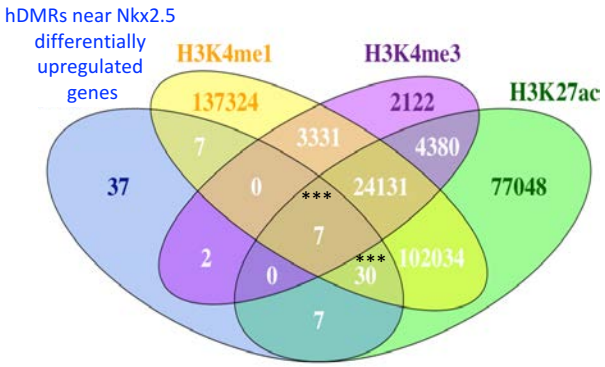

A
log2 fold change
3
2
1
0
−1
−2
−3
Developmental Disorder, Hereditary Disorder, Metabolic Disease
Skeletal and Muscular System Development and Function, Embryonic Development, Organ Development.
DNA Replication, Recombination, and Repair, Nucleic Acid Metabolism, Small Molecule Biochemistry
Developmental Disorder, Hereditary Disorder, Neurological Disease.
Cellular Development, Cellular Growth and Proliferation, Hematological System Development and Function
B
Log10 adj p-value
10
5
0
-2
-1
0
1
2
Log2 fold change
C
GO categories of Upregulated genes
Vascular smooth muscle contraction
Dilated cardiomyopathy
Cardiac muscle contraction
Oxidative phosphorylation
0
5
10
15
Enrichment score

Figure 5

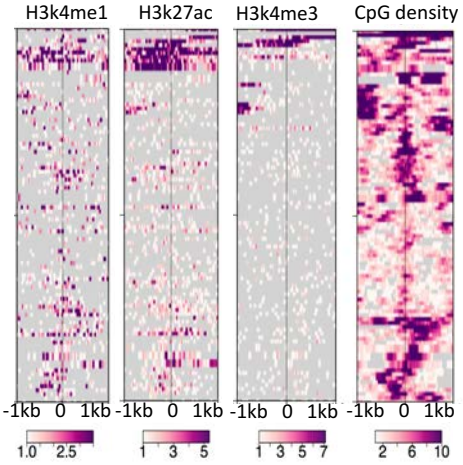
A

Motif	q-value	
Nkx	0.04	
PPARE	0.07	
Pax7	0.12	
Cdx2	0.12	
FoxA1	0.12	
Fli1	0.24	
ERG	0.24	
Erra	0.24	
RxR	0.24	

B



C



D

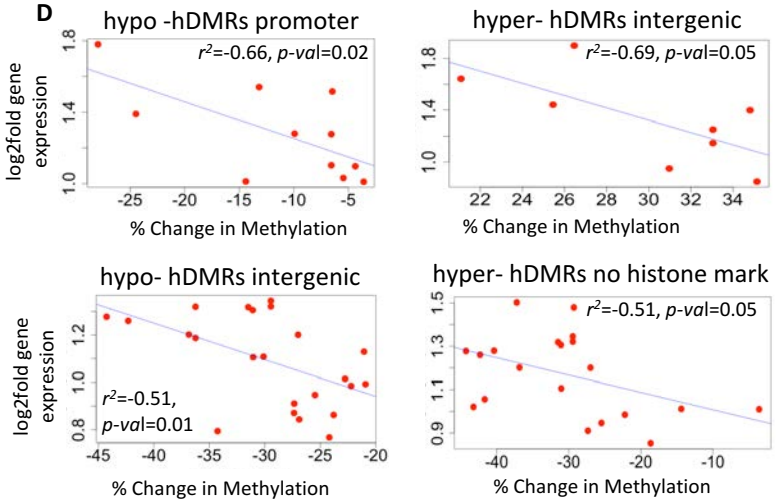
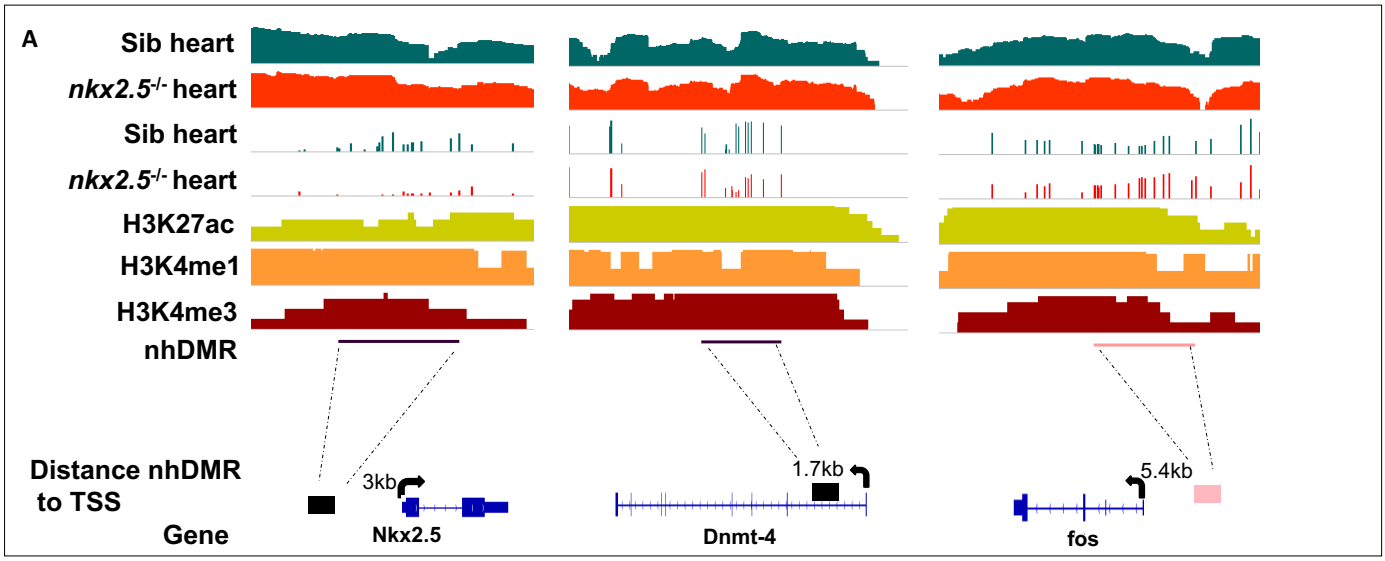


Figure 6


A
Sib heart
nkx2.5-/- heart
Sib heart
nkx2.5-/- heart
H3K27ac
H3K4me1
H3K4me3
nhDMR
Distance nhDMR to TSS
Gene
3kb
Nkx2.5
1.7kb
Dnmt-4
5.4kb
fos

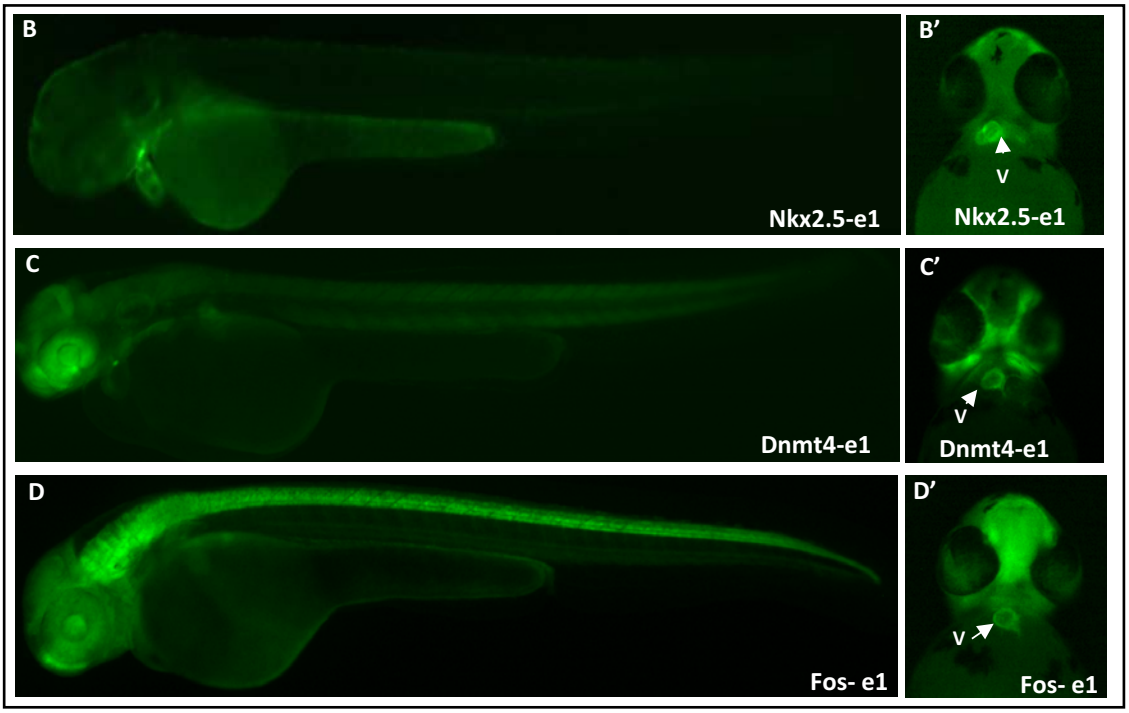
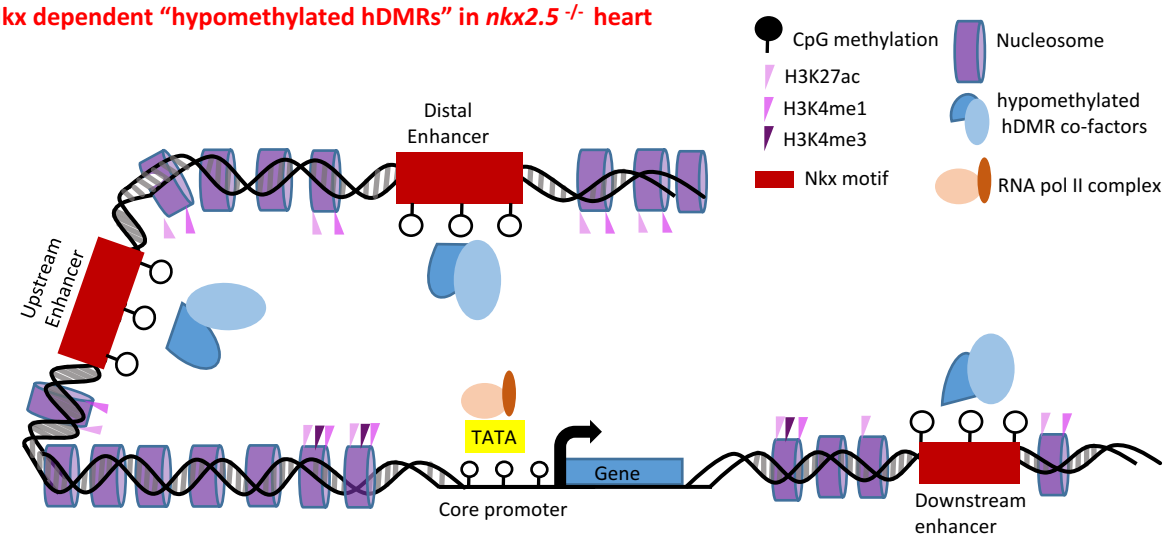

B
Nkx2.5-e1
B'
V
Nkx2.5-e1
C
Dnmt4-e1
C'
V
Dnmt4-e1
D
Fos- e1
D'
V
Fos- e1

Figure 7

A **Nkx dependent "hypomethylated hDMRs" in *nkx2.5* $^{-/-}$ heart**



B **Nkx dependent "hypomethylated hDMRs" in wildtype heart**

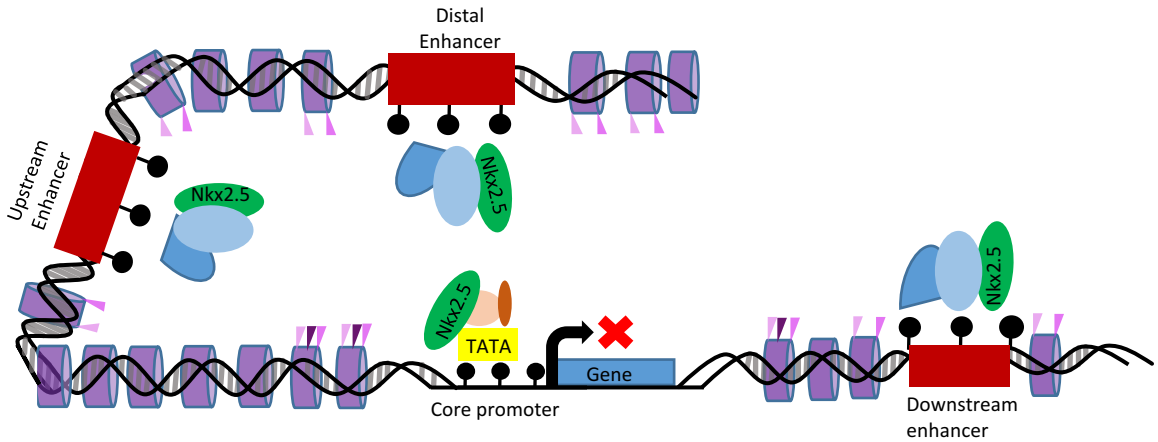


Figure 8

Figure S1

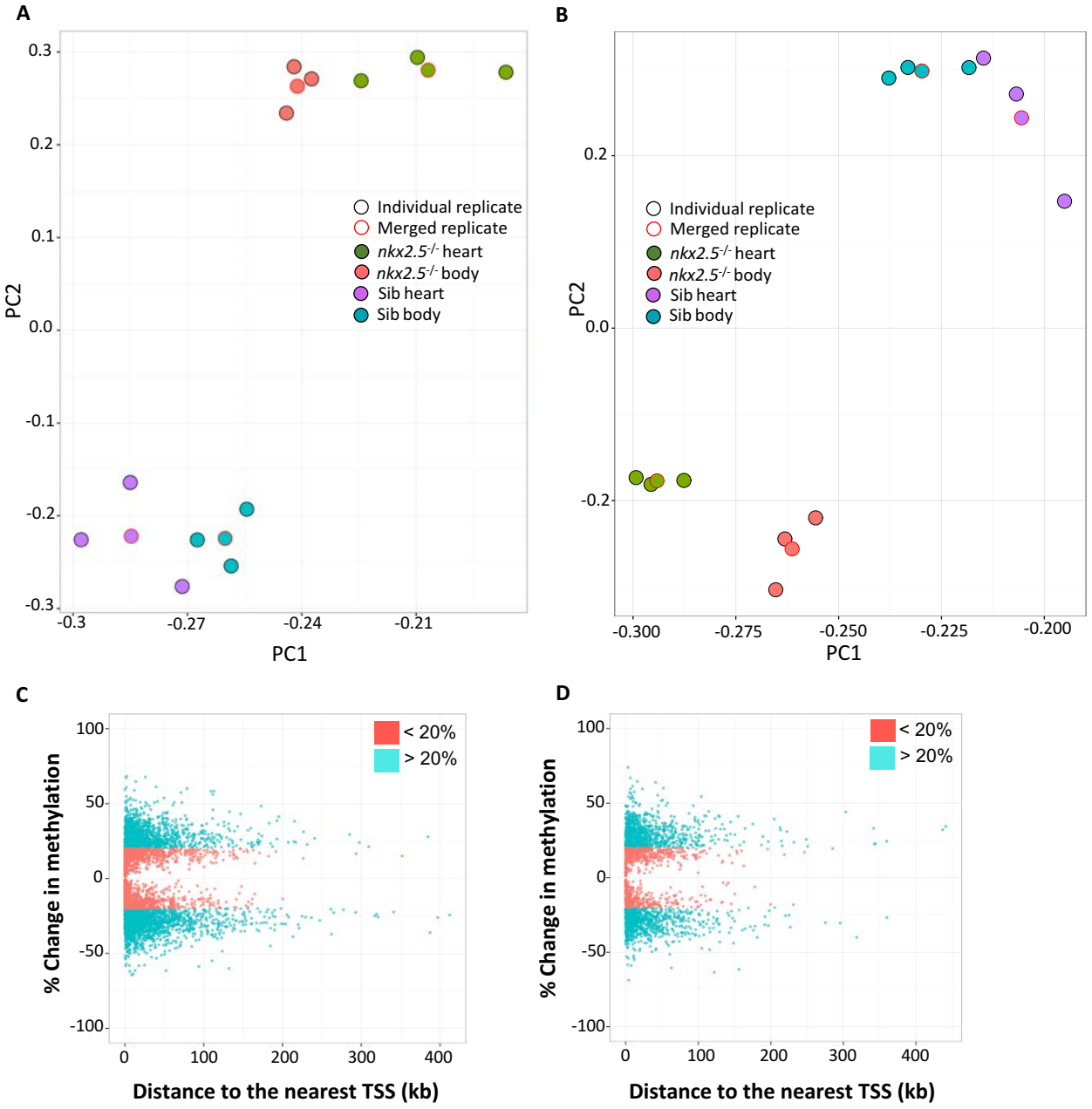

A
B
C
D
PC1
PC2
Individual replicate
Merged replicate
nkx2.5-/- heart
nkx2.5-/- body
Sib heart
Sib body
% Change in methylation
Distance to the nearest TSS (kb)
< 20%
> 20%

Figure S2

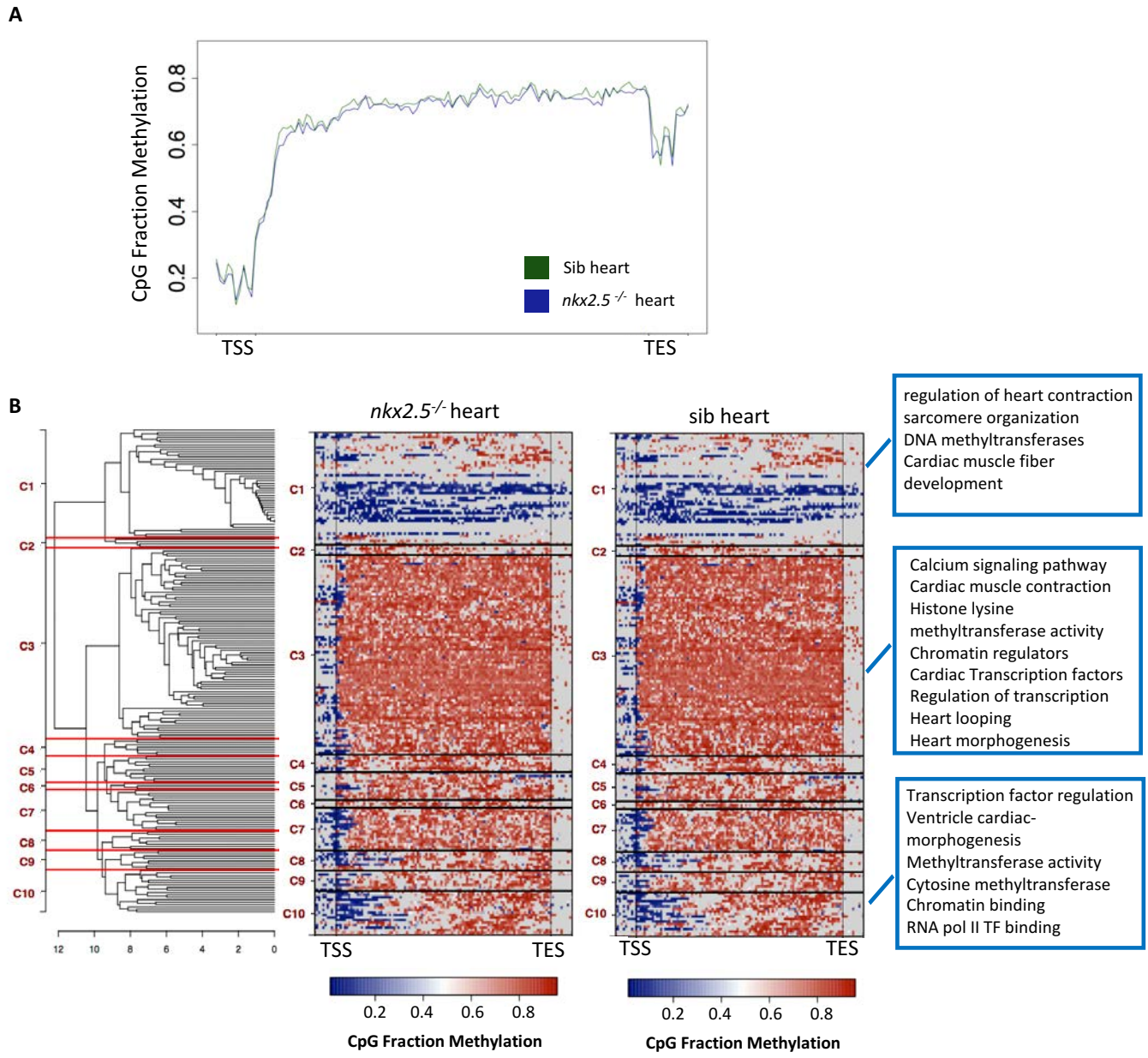
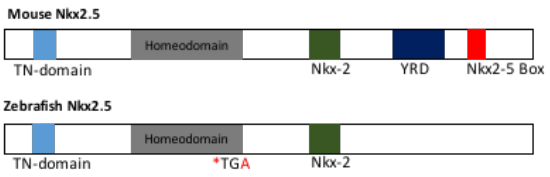
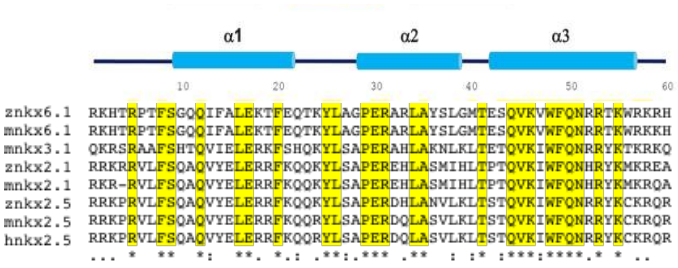

A
CpG Fraction Methylation
0.2
0.4
0.6
0.8
Sib heart
nkx2.5 -/- heart
TSS
TES
B
nkx2.5-/- heart
sib heart
C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
12
10
8
6
4
2
0
TSS
TES
0.2
0.4
0.6
0.8
CpG Fraction Methylation
regulation of heart contraction
sarcomere organization
DNA methyltransferases
Cardiac muscle fiber development
Calcium signaling pathway
Cardiac muscle contraction
Histone lysine methyltransferase activity
Chromatin regulators
Cardiac Transcription factors
Regulation of transcription
Heart looping
Heart morphogenesis
Transcription factor regulation
Ventricle cardiac-morphogenesis
Methyltransferase activity
Cytosine methyltransferase
Chromatin binding
RNA pol II TF binding

Figure S3

A



B



C

Motifs	Sequence	Reference
Nkx2.5		Homer (He *et al* 2011)
Nkx2.5_Meis hybrid		Dupays *et al* 2015
Nkx2.5 (MA0063.1)		Homer
Nkx2.5 (PH0114.2)		Homer
Nkx2.1		Homer
Nkx3.1		Homer
Nkx6.1		Homer
Tbx5 (1)		Homer
Tbx5 (2)		Luna-Zurita *et al* 2016
Tbx5 (3)		Luna-Zurita *et al* 2016
Tbx20		Homer
SRF		Homer
Gata-4		Homer
Gata-3		Homer
Gata-2		Homer
Gata-1		Homer
Hand1		Homer
Tbx1		Homer
Meis1		Homer

Figure S4

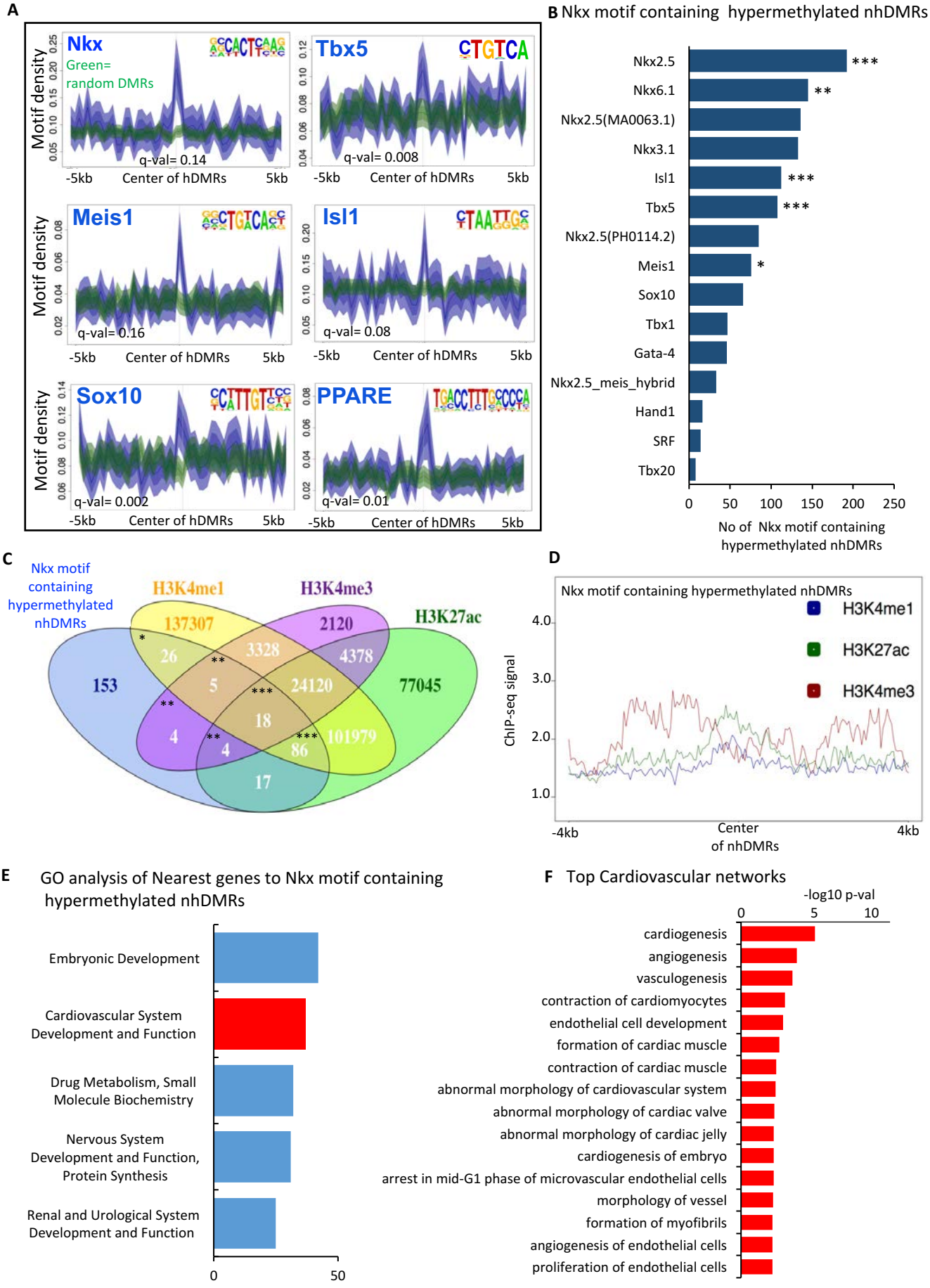
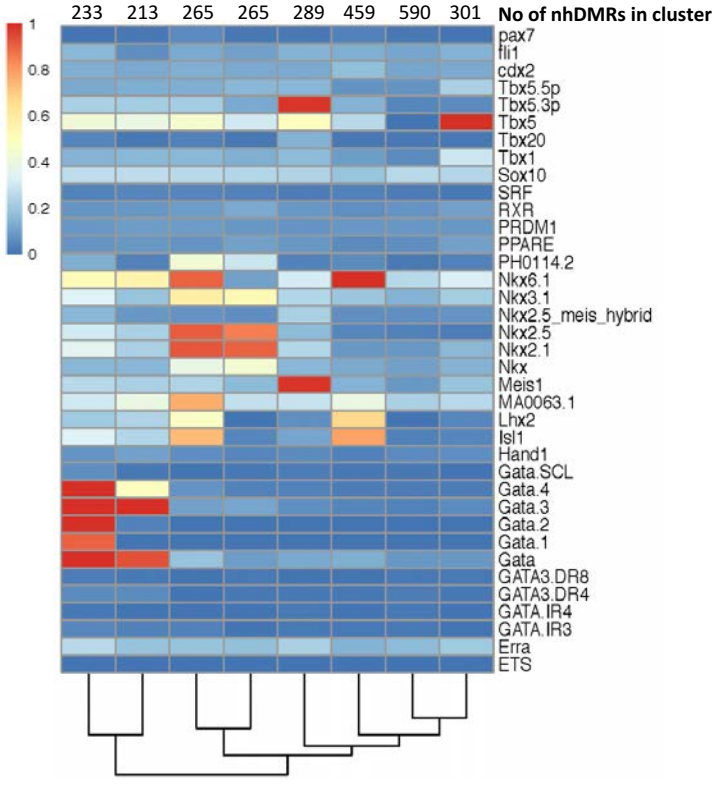

A
Nkx
Green=
random DMRs
q-val= 0.14
Tbx5
q-val= 0.008
Meis1
q-val= 0.16
Isl1
q-val= 0.08
Sox10
q-val= 0.002
PPARE
q-val= 0.01
Motif density
-5kb
Center of hDMRs
5kb
B Nkx motif containing hypermethylated nhDMRs
Nkx2.5 ***
Nkx6.1 **
Nkx2.5(MA0063.1)
Nkx3.1
Isl1 ***
Tbx5 ***
Nkx2.5(PH0114.2)
Meis1 *
Sox10
Tbx1
Gata-4
Nkx2.5_meis_hybrid
Hand1
SRF
Tbx20
0 50 100 150 200 250
No of Nkx motif containing
hypermethylated nhDMRs
C
Nkx motif
containing
hypermethylated
nhDMRs
H3K4me1
H3K4me3
H3K27ac
137307
2120
153
26
3328
4378
5
24120
77045
18
4
4
86
101979
17
D
Nkx motif containing hypermethylated nhDMRs
H3K4me1
H3K27ac
H3K4me3
ChIP-seq signal
4.0
3.0
2.0
1.0
-4kb
Center
of nhDMRs
4kb
E
GO analysis of Nearest genes to Nkx motif containing
hypermethylated nhDMRs
Embryonic Development
Cardiovascular System
Development and Function
Drug Metabolism, Small
Molecule Biochemistry
Nervous System
Development and Function,
Protein Synthesis
Renal and Urological System
Development and Function
0
50
F Top Cardiovascular networks
-log10 p-val
0
5
10
cardiogenesis
angiogenesis
vasculogenesis
contraction of cardiomyocytes
endothelial cell development
formation of cardiac muscle
contraction of cardiac muscle
abnormal morphology of cardiovascular system
abnormal morphology of cardiac valve
abnormal morphology of cardiac jelly
cardiogenesis of embryo
arrest in mid-G1 phase of microvascular endothelial cells
morphology of vessel
formation of myofibrils
angiogenesis of endothelial cells
proliferation of endothelial cells

Figure S5

A **Hypomethylated nhDMRs**



B **Hypermethylated nhDMRs**

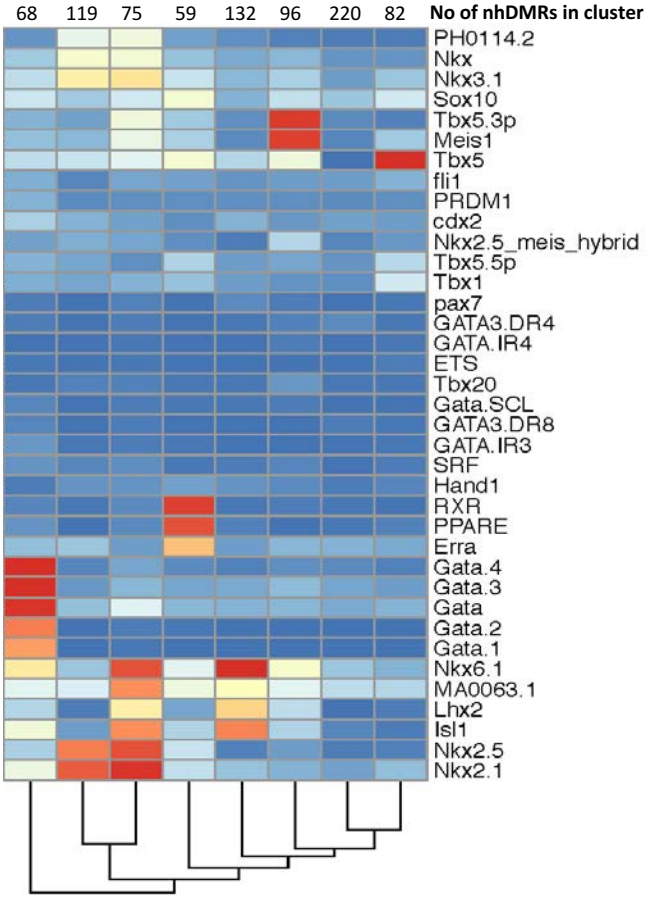
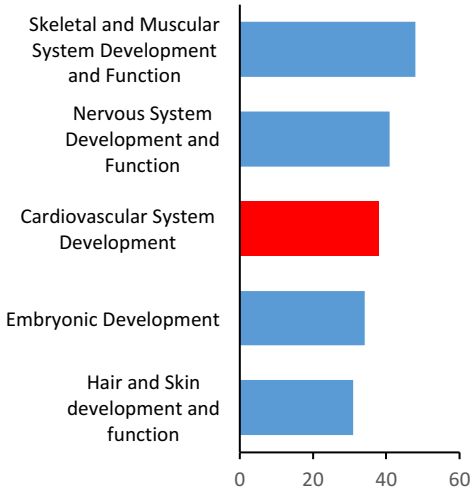
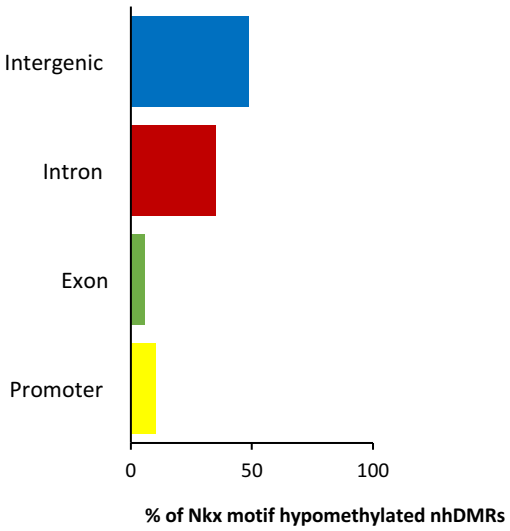


Figure S6

A **GO analysis of Nearest genes to Nkx motif containing hypomethylated nhDMRs**



B **Nkx motif containing hypomethylated nhDMRs**



C

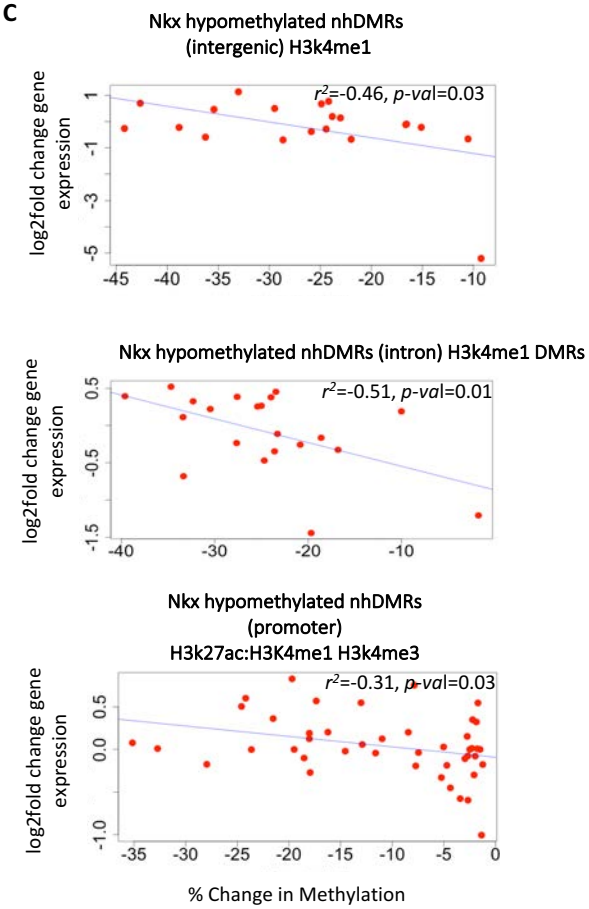
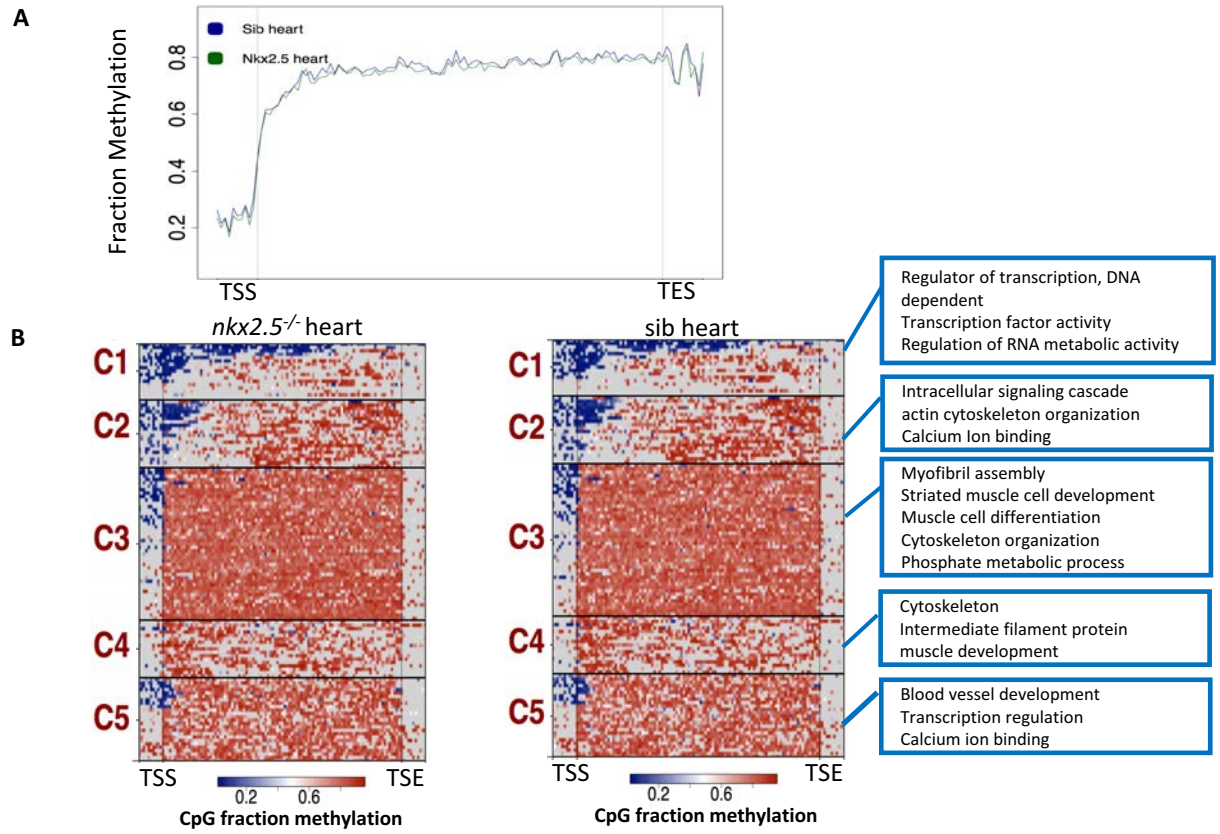


Figure S7


A
Sib heart
Nkx2.5 heart
Fraction Methylation
0.2
0.4
0.6
0.8
TSS
TES
B
nkx2.5-/- heart
sib heart
C1
C2
C3
C4
C5
TSS
TSE
0.2
0.6
CpG fraction methylation
Regulator of transcription, DNA dependent
Transcription factor activity
Regulation of RNA metabolic activity
Intracellular signaling cascade
actin cytoskeleton organization
Calcium Ion binding
Myofibril assembly
Striated muscle cell development
Muscle cell differentiation
Cytoskeleton organization
Phosphate metabolic process
Cytoskeleton
Intermediate filament protein
muscle development
Blood vessel development
Transcription regulation
Calcium ion binding

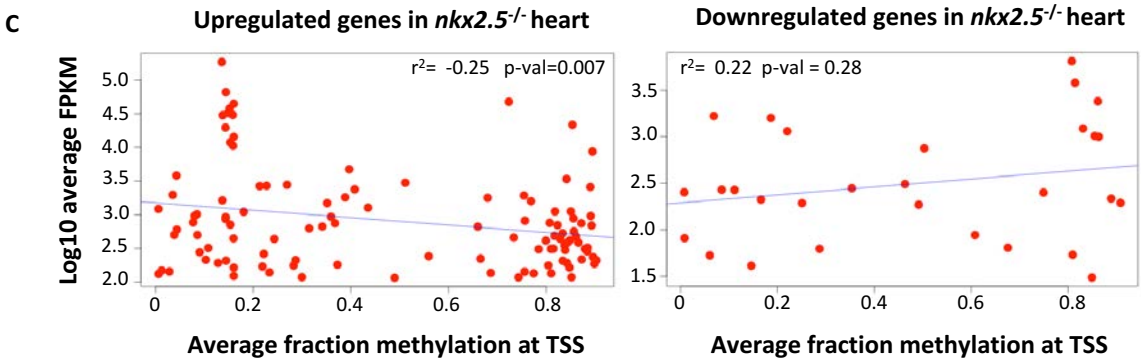
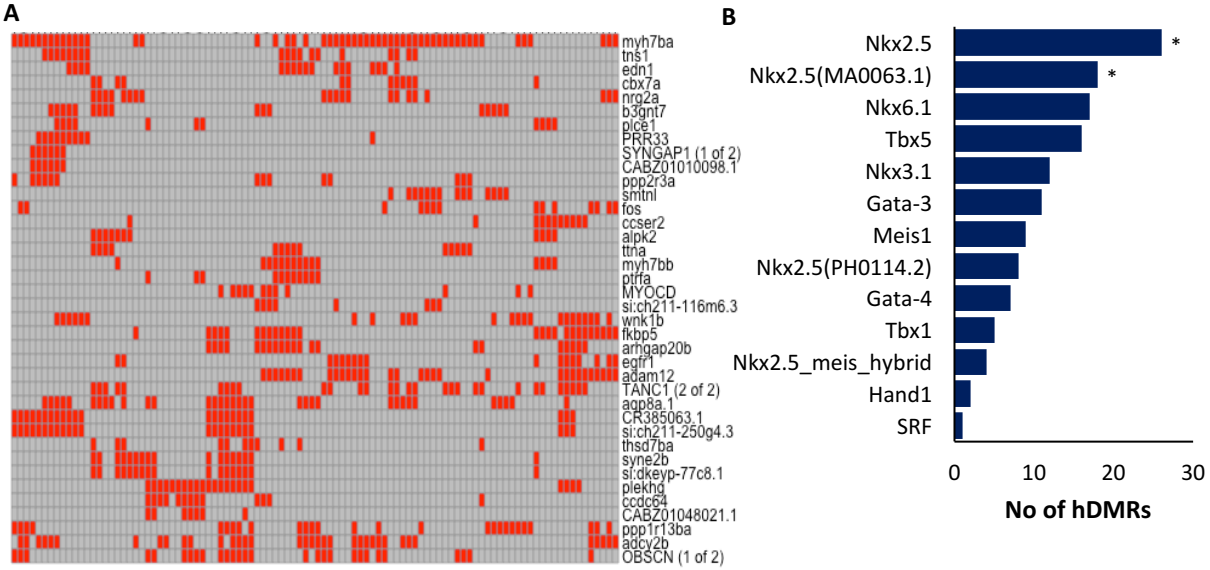

C
Upregulated genes in nkx2.5-/- heart
Downregulated genes in nkx2.5-/- heart
r2= -0.25 p-val=0.007
r2= 0.22 p-val = 0.28
Log10 average FPKM
Average fraction methylation at TSS
Average fraction methylation at TSS

Figure S8


A
myh7ba
tns1
edn1
cbx7a
nrg2a
b3gnt7
plce1
PRR33
SYNGAP1 (1 of 2)
CABZ01010098.1
ppp2r3a
smtnl
fos
ccser2
alpk2
ttna
myh7bb
ptrfa
MYOCD
si:ch211-116m6.3
wnk1b
fkbp5
arhgap20b
egfr1
adam12
TANC1 (2 of 2)
aqp8a.1
CR385063.1
si:ch211-250g4.3
thsd7ba
syne2b
si:dkeyp-77c8.1
plekhg
ccdc64
CABZ01048021.1
ppp1r13ba
adcy2b
OBSCN (1 of 2)
B
Nkx2.5
Nkx2.5(MA0063.1)
Nkx6.1
Tbx5
Nkx3.1
Gata-3
Meis1
Nkx2.5(PH0114.2)
Gata-4
Tbx1
Nkx2.5_meis_hybrid
Hand1
SRF
*
*
0
10
20
30
No of hDMRs