Essential and Unexpected Role of YY1 to Promote Mesodermal Cardiac Differentiation

Serge Gregoire¹, Ravi Karra², Derek Passer¹, Marcus-Andre Deutsch¹,³, Markus Krane¹,³, Rebecca Feistritzer¹, Anthony Sturzu¹, Ibrahim Domian¹,⁵, Yumiko Saga⁴, and Sean M. Wu¹,⁵

¹Cardiovascular Research Center, Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Boston, MA; ²Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, NC; ³Department of Cardiovascular Surgery, German Heart Center, Clinic at the Technische Universitaet Muenchen, Munich, Germany; ⁴Division of Mammalian Development, National Institute of Genetics, Mishima, Japan; and ⁵Harvard Stem Cell Institute, Cambridge, MA 02138.

Running title: YY1 is Required for Cardiac Differentiation

Subject codes:
[137] Cell biology/structural biology
[139] Developmental biology
[154] Myogenesis
[89] Genetics of cardiovascular disease

Address correspondence to:
Dr. Sean M. Wu
G1120A Lokey Stem Cell Building
265 Campus Drive
Stanford, CA 94305
Tel: (650)724-4498
Fax: (650)724-4689
smwu@stanford.edu

In December 2012, the average time from submission to first decision for all original research papers submitted to Circulation Research was 14.5 days.
ABSTRACT

**Rational:** Cardiogenesis is regulated by a complex interplay between transcription factors. However, little is known about how these interactions regulate the transition from mesodermal precursors to cardiac progenitor cells (CPCs).

**Objective:** To identify novel regulators of mesodermal cardiac lineage commitment.

**Methods and Results:** We performed a bioinformatic-based transcription factor binding site analysis on upstream promoter regions of genes that are enriched in embryonic stem cell (ESC)-derived CPCs. From 32 candidate transcription factors screened, we found that YY1, a repressor of sarcomeric gene expression, is present in CPCs *in vivo*. Interestingly, we uncovered the ability of YY1 to transcriptionally activate Nkx2.5, a key marker of early cardiogenic commitment. YY1 regulates Nkx2.5 expression via a 2.1 kb cardiac-specific enhancer as demonstrated by *in vitro* luciferase-based assays and *in vivo* chromatin immunoprecipitation (ChIP) and genome-wide sequencing analysis. Furthermore, the ability of YY1 to activate Nkx2.5 expression depends on its cooperative interaction with Gata4 at a nearby chromatin. Cardiac mesoderm-specific loss-of-function of YY1 resulted in early embryonic lethality. This was corroborated *in vitro* by ESC-based assays where we show that the overexpression of YY1 enhanced the cardiogenic differentiation of ESCs into CPCs.

**Conclusion:** These results demonstrate an essential and unexpected role for YY1 to promote cardiogenesis as a transcriptional activator of Nkx2.5 and other CPC-enriched genes.

**Keywords:** Progenitor cell, stem cell, cardiac development, congenital heart disease

**Non-standard Abbreviations:**
- ChIP-seq: chromatin immunoprecipitation sequencing
- CPCs: cardiac progenitor cells
- EMSA: electromobility shift assays
- TF: transcription factors
- TFBS: transcription factor binding site
- TSS: transcription start site
- YY1: Yin Yang 1

INTRODUCTION

Understanding the mechanisms of cardiac lineage commitment is important for identifying the pathophysiological basis of congenital heart disease, the most common cause of human birth defects. In addition, a detailed understanding of cardiac lineage commitment may also be important for successful cell-based therapies. Recently, we and others have described the isolation and characterization of embryo- and embryonic stem cell (ESC)-derived cardiac progenitor cells (CPCs). These cells give rise to a variety of different cells types in the developing heart. While the developmental potential of embryonic CPCs is an active area of investigation, limited progress has been made in understanding the genes and pathways that regulate the commitment of multipotent mesodermal precursor cells into committed CPCs.
Yin Yang 1 (YY1), a member of the GLI-Kruppel family of DNA binding zinc finger transcription factor\textsuperscript{10,11}, can activate or inhibit transcription in a context-dependent manner\textsuperscript{12}. Global deletion of the murine YY1 alleles results in peri-implantation embryonic lethality\textsuperscript{13}. Haploinsufficiency of a YY1 allele results in neural defect and cyanosis with subsequent fetal demise during late gestation. Mechanistically, YY1 has been described to anchor polycomb group (PcG) proteins to chromatin DNA and may substitute for polycomb protein pleiohomeotic during Drosophila development\textsuperscript{14}.

The role of YY1 in postnatal skeletal and cardiac sarcomeric gene transcription has been previously examined. Binding sites for YY1 have been identified within the regulatory regions of several cardiac and skeletal muscles genes, including muscle creatine kinase, α-skeletal actin, α-myosin heavy chain, myosin light chain 2, and α-cardiac actin\textsuperscript{15-19}. Despite these studies, the ability of YY1 to activate cardiac gene expression, particularly during early cardiac development in mice, has not been examined. YY1, in cooperation with Gata-4, activates the B-type natriuretic peptide promoter \textit{in vitro}\textsuperscript{20} and induce periostin enhancer expression in the cardiac outflow tract\textsuperscript{21}. In a chick Nkx2.5 enhancer, YY1 was shown to modulate Smad-mediated Nkx2.5 expression\textsuperscript{22}. Furthermore, YY1 binding to xenopus myosin light chain (xMLC) cardiac enhancer was necessary to induce full expression in transgenic embryos\textsuperscript{23}. While these studies provide suggestive evidence that YY1 may promote cardiac gene transcription during embryonic development, the requirement for YY1 to enhance cardiogenesis remains largely undefined. Furthermore, no study has addressed the role of YY1 during mammalian cardiac lineage commitment from multipotent mesodermal precursor cells.

In this study, we demonstrate that YY1 is required during early cardiac development. Genetic ablation of YY1 alleles in mesodermal cells resulted in early embryonic lethality. Using ESCs, we show that YY1 promotes cardiogenic differentiation. Detailed analysis of YY1-mediated transcriptional regulation at a cardiac-specific enhancer of Nkx2.5 revealed a direct requirement for YY1 DNA binding to activate gene expression. The \textit{in vivo} binding of YY1 to this enhancer in embryonic- and ESC-derived CPCs was heart cell-specific, as demonstrated by chromatin immunoprecipitation (ChIP) and genome-wide sequencing analysis. Taken together, these results illustrate an essential role for YY1 to promote cardiogenic commitment by transcriptionally activating Nkx2.5 and other cardiac genes.

METHODS

Detailed Materials and Methods can be found in Online Supplemental Information.

RESULTS

\textit{Screening for regulators of gene expression in CPCs.}

During cardiac development, DNA binding transcription factors that activate gene expression in CPCs are bound to the promoter/enhancers of genes enriched in CPCs\textsuperscript{25}. We postulated that by examining bioinformatically the occurrence of over-represented transcription factor binding sites (TFBS) in the promoter regions of a set of overexpressed genes in CPCs, we might be able to identify key transcription factors that regulate the expression of these genes collectively (Figure 1A). We first identified a list of CPC-enriched genes by performing genome-wide transcriptional profiling of early CPCs at days 4 and 5 of ES cell differentiation (Figure 1B). The global gene expression from each cell population was quantitated by microarray analysis (Figure 1C). After sample normalization and correction for multiple hypothesis testing, we identified a core set of 14 unique genes enriched in CPCs (FDR < 0.05) (Figure
Gene ontology analysis confirmed that these genes (e.g. Cnn1, Sema3C, Acta1, Acta2, TnnC1, and Actc1, SM22, and Myl9) are associated with “muscle contraction” (p < 0.0001) and “muscle development” (p < 0.0001). Of note, cardiac transcription factors (including Nkx2.5) were excluded from this list due to their relatively lower level of expression.

To identify upstream transcription factors that may drive the expression of these 14 CPC-enriched genes, we performed a TFBS analysis on the sequences 5 kb upstream from the transcription start site of these genes using rVISTA, a computational tool that allows for identification of over-represented TFBS. We found over-represented TFBS for 44 transcription factors (TFs) (Online Table I). Notably, several of these TFs including Smad, LEF1/TCF, Gli, Ets2 and Sp3 are involved in BMP, Wnt, Hedgehog, and other signaling pathways involved in normal cardiac development.

We further refined this list by finding over-represented TFBS in five regions (A, B, C, D, E) of highly conserved DNA sequence between the mouse and human Nkx2.5 promoter (Figure 1D). Analysis of TFBS in these five regions resulted in 13 over-represented TFBS (Online Table II). From the two lists of candidate TFs (Online Tables I and II), three factors - YY1, AP-2α, and myogenin - were present in both (Figure 1E). Immunostaining in the early developing heart tube showed that only YY1 has robust expression that coincides with the expression of Gata4 and Nkx2.5 (Figure 1E). Since cells from embryonic day 9.5 heart tubes have already matured beyond the CPC stage of development, the exact stage of development that YY1 plays a role in this process requires further investigation.

YY1 binds directly to an Nkx2.5 cardiac enhancer.

To examine in detail the potential binding of YY1 to Nkx2.5 enhancer, we performed an electrophoresis mobility shift assay (EMSA) using the 2.1 kb cardiac enhancer located within -9435 and -7353 bases upstream from the murine Nkx2.5 transcriptional start site. From 10 predicted YY1 binding sites (Figure 2A), we found that YY1 binds specifically to sites #1, #2, #3, #6, and #9 but not sites #4, #5, #7, #8 and #10 (Figure 2B). Binding of site #1 to YY1 was further confirmed using a DNA supershift assay (Figure 2C). To corroborate these in vitro data with their binding to the Nkx2.5 cardiac enhancer in vivo, we performed chromatin immunoprecipitation (ChIP) and PCR analysis using anti-YY1 antibodies on day 6 differentiated NK ESCs (Figure 2D), FACS-purified eGFP+ CPCs and eGFP- cells from day 6 differentiated NK ESCs (Figure 2E), and from embryonic day 9.5 hearts and body (Figure 2F). As shown in Figure 2D-F, YY1 binds to the Nkx2.5 cardiac enhancer in cardiac cells but not in eGFP- and non-cardiac cells from the e9.5 embryo.

Identification of YY1 target genes in CPCs by genome-wide ChIP-sequencing analysis.

Given the ability of YY1 to bind the Nkx2.5 cardiac enhancer both in vitro and in vivo, we postulated that YY1 may also bind to the promoter regions of other cardiac genes. We performed ChIP-sequencing (ChIP-seq) on YY1-bound chromatin in ES cell-derived CPCs and identified ~3700 unique binding peaks for YY1 (Figure 3A). Peak locations were analyzed with respect to Refseq gene models and to TSS (Figures 3B-3C). Gene ontology analysis of YY1 bound sites also identified genes belonging to the “plasma membrane”, “cytoskeleton”, “intermediate filaments”, and “extracellular region” (Figure 3D).

Interestingly, YY1 binding was highly enriched in the Nkx2.5 cardiac enhancer site (Figures 3E) and in other CPC enriched genes (Online Figure 1-A) in eGFP+ but not eGFP- cells. Among the 14 unique genes that are enriched in CPCs (Figure 1C), 6 had detectable YY1 binding peaks by ChIP-seq (Online Table III) and 4 were confirmed independently by ChIP-PCR (Online Figure I-B and I-D) and ChIP-qPCR (Online Figure I-C, I-E) in both eGFP+ cells and day 9.5 embryo hearts. Beyond the 14 genes that are enriched in CPCs, we also uncovered binding of YY1 to the enhancer regions of developmentally important cardiac transcription factors such as Tbx3, Tbx5, Mef2c, and Pitx2 (Figure...
as well as to the promoter/enhancer of cardiac sarcomeric genes such as titin, tropomysin, myosin 18B, and Mkl2 (Online Figure II-A). Analysis of over-represented TFBS within the YY1-bound genomic regions revealed an enrichment of Gata4, AP2, Foxc1, CEBPα and others (Online Figure II-B). Taken together, our in vitro studies and our genome-wide ChIP-seq analysis confirmed the binding of YY1 to the Nkx2.5 cardiac enhancer and to the promoter regions of several CPC-enriched genes. Additionally, the binding of YY1 to the enhancer regions of other key cardiac transcription factors and sarcomeric proteins (Figure 3E and Online Figure II-C) suggests that YY1 may be a critical transcriptional regulator of early cardiac development.

To examine the potential role of cooperative interactions between YY1 and other cardiac transcription factors to regulate gene expression, we focused on the co-occupancy of YY1 and Gata4 to their genome targets since a recent study showed that YY1 binding was enriched within Gata4-bound sites. We extracted the data from this analysis performed on an HL1 cell line and cross-referenced with our YY1 ChIP-seq data from ESC-derived eGFP+ cells to determine the candidate list of genes that are co-occupied by Gata4 and YY1. Interestingly, 195 genes were bound by both factors (Figure 3G). Gene ontology analysis clustered these genes within “cell adhesion”, “cytoskeletal protein binding”, “cell migration” and “phosphoproteins” functional classes with an FDR rate of <0.05 (Figure 3H). While it should be emphasized that these two ChIP-seq data were from different cell sources and experimental contexts, the finding of a potential cooperative interaction between YY1 and Gata4 provided us with a compelling rationale for further exploring this interaction experimentally.

YY1 regulates Nkx2.5 cardiac enhancer activity in a Gata4-dependent manner.

To examine the biological consequence of YY1 binding to the Nkx2.5 cardiac enhancer and the involvement of Gata4 in this process, we generated a series of luciferase reporter constructs (Figure 4A) to probe the requirement of YY1 and Gata4 DNA binding on the transcriptional activity of the Nkx2.5 enhancer. Using the H9C2 cardiomyocyte cell line, we first confirmed that overexpression of Gata4 induces a dose-dependent transcriptional activation of the full-length enhancer (WT) (Figure 4B), as previously described. Without addition of exogenous YY1, we observed a considerable activation of the wild type reporter by Gata4 alone, likely due to endogenous YY1 activity (data not shown). Interestingly, a majority of this transcriptional activation appears to reside within the first 330 base pairs containing a critical Gata4 binding site (GS1) (Figure 4B).

To identify whether direct DNA binding of YY1 is required for the transcriptional activation of the Nkx2.5 cardiac enhancer, we generated reporter constructs in which either all five YY1 binding sites (#1, #2, #3, #6, #9) (5BS) or four of five (#2, #3, #6, #9) binding sites (4BS) have been mutated. Remarkably, in the absence of YY1 DNA binding, the transcriptional activity of this enhancer is completely abolished (Figure 4C). The presence of only one binding site in the 4BS mutant is insufficient to rescue the transcriptional activation of this enhancer (Figure 4C). We further explored the relative requirement of each YY1 binding site to the transcriptional activation of the Nkx2.5 cardiac enhancer and found that four of these sites (#1, #2, #3, and #9) are involved in Gata4-dependent transcriptional activation while one site (#6) may actually play a repressive role (Figure 4D). In support of this hypothesis, we found that the activity of the 5' enhancer domain (1-330) is abolished when this repressive YY1 binding site (#6) is present (Online Figure III-A, III-B). Of note, the overexpression of YY1 does not lead to an increase in Gata4 expression in H9C2 cells (and vice versa), eliminating a potential confounder of our reporter assay results (Online Figure III-C, III-D).

To determine if YY1 activity is dependent on DNA binding in the cardiac enhancer by Gata4, we generated mutant constructs in each of the two highly conserved Gata4 binding sites (GS1, GS2) (Figure 4A) and found that modification of either of these sites renders the Nkx2.5 cardiac enhancer inactive (Figure 4E). Finally, we confirmed the requirement of YY1 binding to Nkx2.5 cardiac enhancer in vivo.
by generating transient transgenic mouse embryos from pronuclear injection of control and 5BS mutant Nkx2.5 cardiac enhancer/base promoter-LacZ constructs. We were able to generate a total of 26 collected embryos containing the WT construct and 25 embryos containing the 5BS mutant construct, which we confirmed by PCR. From the 26 embryos carrying the wild type DNA construct, 6 showed LacZ+ signal in the heart. In contrast, from the 25 embryos carrying the 5BS mutant construct, none exhibited LacZ+ staining. A representative image of a LacZ stained heart is shown in Figure 4F.

To examine whether a direct physical interaction may exist between YY1 and Gata4 to mediate their synergistic effects, we performed a tandem YY1-Gata4 co-ChIP assay in day 6 in vitro differentiated NK ESCs (Figure 4G) and e9.5 hearts (Figure 4H) and found that YY1 and Gata4 occupy nearby chromatin sites. While it is possible that DNA-bound YY1 and Gata4 from two distant chromatin regions may be pulled down simultaneously in the presence of a third transcription factor capable of binding to both proteins, we believe a direct YY1/Gata4 interaction is the most likely mechanism given our data showing that YY1 and Gata4 proteins interact with one another in a co-immunoprecipitation (co-IP) assay (Figure 4I). Further studies using ES cell-derived eGFP+ cells or heart cells will be needed to validate this H9C2 cell-based binding interaction. Taken together, these results demonstrate a cooperative interaction between YY1 and Gata4 to regulate the Nkx2.5 cardiac enhancer activity.

YY1 loss-of-function results in embryonic lethality and the absence of cardiac lineage formation.

Given the ability of YY1 to induce cardiac gene expression in CPCs, we postulate that YY1 may be required for proper cardiac lineage differentiation. We interbred floxed YY1 mice with Mesp1-Cre mice and observed no surviving embryo carrying both homozygous YY1 floxed alleles and the Mesp1-Cre allele from e9.5 onwards (Online Table IV). At embryonic day 8.75-9.0, we found Mesp1-Cre+; YY1 Flox/Flox (F/F) embryos that exhibit developmental arrest (3 of 3) with failure to undergo midline fusion (Figure 5A). These knock-out embryos exhibit only 9 somites whereas the wild type and YY1 heterozygous embryos exhibit 11 somite (Figure 5A). By quantitative real-time PCR analysis of cells dissected from heart forming regions of wild type (WT), heterozygous (Het), and homozygous (Homo) YY1 mutant embryos at day 7.5 post coitum, we found a marked reduction in the expression of Nkx2.5 and Isl-1+ in homozygous embryos, suggesting the loss of CPC formation (Figure 5B). We further confirmed this by performing an immunohistochemical analysis on e7.5 heart sections and found an absence of Nkx2.5+ or Isl1+ cells in homozygous YY1 mutant embryos (yellow arrows) (Online Figure IV-B).

YY1 loss-of-function inhibits cardiogenesis of NK ESCs.

To gain access to cells that are most representative of differentiating cells in early embryos, we chose in vitro differentiated NK ESCs as a surrogate model. We first generated an YY1 inhibitor construct that contains the DNA binding domain of YY1 that acts in a dominant negative (DN) fashion. We then confirmed the ability of YY1-DN to silence the transcriptional activation of the Nkx2.5 cardiac enhancer using a luciferase assay in the H9C2 cell line (Figure 6B). We then further explored the effect of YY1-DN overexpression to inhibit cardiac lineage development in differentiating NK ESCs by generating doxycycline-inducible YY1-DN ES cell lines and characterized multiple single-cell clonally-derived cell lines (NK-YY1-DN ESC lines) for their ability to overexpress YY1-DN (Figure 6C). In one such line, doxycycline treatment resulted in ~20-fold increase in YY1-DN expression and was further studied (Figure 6D). We then investigated the effect of YY1-DN overexpression on cardiac differentiation (as measured by the abundance of eGFP+ cells) and found a ~40% (1.4 +/- 0.07% vs 0.85 +/- 0.06%) reduction in the percentage of eGFP+ cells with doxycycline treatment (Figures 6E-F). This was corroborated by a marked reduction in expression of CPC markers such as Nkx2.5, Isl-1, Gata4, and Tbx5 (Figure 6G).
YY1 gain-of-function promotes cardiogenesis of NK ESCs.

Since our loss-of-function studies above showed a requirement for YY1 in embryonic development and in cardiac differentiation of NK ESCs, we examined whether YY1 gain-of-function may lead to enhanced cardiogenesis. To address this, we generated a doxycycline-inducible YY1 overexpressing lentivirus using a previously described tet-on expression system (Online Figure V-A) and infected the NK ESC line with this virus to generate multiple single-cell clonally-derived inducible YY1 overexpressing NK ESC lines (NK-YY1) (Figure 7A). From the 32 lines examined, one of these was found to generate the highest increase in YY1 expression with doxycycline treatment (~60-fold) and was used for further study (Figure 7B).

To assess the effects of YY1 overexpression on cardiac differentiation of ESCs, the NK-YY1 ESCs were differentiated in vitro for 6 or 10 days in the presence or absence of doxycycline induction (Figure 7C). We confirmed that doxycycline treatment led to a 60- and 22-fold increase in the expression of YY1 on days 6 and 10 of differentiation, respectively (Online Figure 4V-B). As shown in Figure 7, the percentage of early (day 6) and late (day 10) eGFP+ cells increased in the presence of doxycycline treatment by approximately 4- and 3-fold, respectively, (Figure 7D, 7E, 7F). In support of the increase in cardiac cell population following YY1 overexpression, the mRNA levels of a number of key CPC-associated transcription factors such as Nkx2.5, Isl-1, Hand2, and Tbx5 were elevated by 2 to 10-fold at day 6 of differentiation (Figure 7G). The expression of vascular genes such as VE-cadherin and SMAα were increased as well (Figure 7H), supporting the role of YY1 to induce multipotent CPCs. On day 10, the increase in the percentage of cardiomyocytes with YY1 overexpression was corroborated by a 10 to 50-fold increase in cardiac transcription factor expression (Online Figure V-C) and a 20-30 fold increase in the expression of sarcomeric genes such as Mlc2v and cTnT (Figure 7I). While the increase in eGFP+ cell percentage (3 to 4-fold) by flow cytometry was significantly less than the increase in gene expression (10 to 50-fold), we believe this may be due to the differences in the method of quantification (flow cytometry vs qPCR) and the transcriptional activity of YY1 at these genes.

YY1 induces mesodermal stage precursors into CPCs.

To determine precisely the developmental window that YY1 acts to promote cardiac differentiation, we induced the overexpression of YY1 from either days 0 to 3 (i.e. analogous to pre-gastrulation stage cells) or 3 to 6 (i.e. analogous to germ layer stage cells) of in vitro differentiation of NK-YY1 ES cells (Figure 8A). When compared with untreated cells as control, doxycycline treatment from days 3 to 6 led to a >3-fold increase in the percentage of eGFP+ cells at day 6 of differentiation (Figure 8B, 8C). Interestingly, treatment of doxycycline from days 0 to 3 of differentiation resulted in no increase in eGFP+ cells (Figure 8C). The overexpression of YY1 from day 3-6 also induced a dramatic increase in the transcript levels of Nkx2.5 and Tbx5 (>150- and 40-fold, respectively) (Figure 8D) without a parallel increase in the expression of endoderm, ectoderm, and some non-cardiac mesoderm genes (Online Figure VI). These findings support a specific role for YY1 to enhance cardiogenic commitment from mesodermal precursor cells.

DISCUSSION

To gain insights into early cardiac development, we employed a combination of expression and bioinformatic analysis that led to the identification of YY1 as a novel candidate gene in early cardiogenesis. We confirmed the in vivo binding of YY1 to an Nkx2.5 cardiac enhancer in both ESC and embryo-derived cardiac cells by chromatin immunoprecipitation. We also found that YY1, in cooperation with
Identification of novel transcription factors involved in early cardiogenesis.

While the transcription network driving cardiac gene expression during embryonic development has been intensely studied\textsuperscript{37-41}, there has been relatively few efforts focusing on the developmental transition from mesodermal precursors to committed CPCs. During gastrulation, mesodermal precursors that are destined for a cardiovascular fate express Mesp1\textsuperscript{42-44}. Mesp1 overexpression enhances the commitment of multipotent mesodermal cells into CPCs inducing the expression of key cardiac transcription factors such as Nkx2.5, Isl-1, and Tbx5\textsuperscript{45,46}. Here, we identified YY1 as a regulator of cardiogenesis by its regulation of CPC gene expression. We also found that other TFs such as AP2-\(\alpha\), Ets2, and Gli may play important roles in early cardiac development as well. In support of this, signaling by sonic hedgehog has been shown to regulate the number of myocardial progenitor cells in the arterial pole during zebrafish heart development\textsuperscript{47,48}. The formation of cardiac precursors in ascidian \textit{Ciona intestinalis} require activation of Ci-Ets2\textsuperscript{49}. Further investigation into the biological requirement of these and other factors during mesodermal to CPC transition in mice will be informative.

Regulation of NKX2.5 expression by YY.

We performed an in-depth analysis of the regulation of Nkx2.5 expression by YY1 since Nkx2.5 is one of the earliest genes expressed in heart forming precursors and plays an essential role in cardiac development\textsuperscript{50,51}. We found that YY1 functions cooperatively with Gata4 on nearby chromatin. Indeed, recent genome-wide target analysis in the HL1 cell line by ChIP sequencing for these cardiac transcription factors have found co-occupancy of SRF, Tbx5, and Gata4 along with p300 within the YY1 binding region on the Nkx2.5 enhancer\textsuperscript{36}. This implies that YY1 is able to regulate cardiac-specific gene expression by its association with other cardiac TFs. In support of this, we identified an over-representation of binding sites for Gata4, CEBP\(\alpha\), and Foxc1 in YY1 ChIP-seq target regions (Figure 3D). These genes are all known to play important roles during early cardiac development\textsuperscript{52-55}.

Regulation of other cardiac gene expression by YY1.

Our genome-wide analysis has also found also many other cardiac genes (e.g. Tbx3, Tbx5, Pitx2, and Mef2c, etc) that are potentially regulated by YY1. Of interest, Tbx3 is expressed in the outflow tract, the AV canal, the central conduction system and is required for normal alignment of the outflow tract and for normal ventricular septation\textsuperscript{56}. Pitx2 plays a critical role in regulating left-right asymmetry during heart development and is a direct down-stream effector of sonic hedgehog signaling\textsuperscript{48}. The occupancy of YY1 in their promoter/enhancer regions during the mesodermal stage of development may be required for their proper expression during later stages of cardiac development.

Context-dependent regulation of cardiac transcriptional program by YY1.

A surprising aspect of our studies here is the role of YY1 to transcriptionally activate genes in early cardiac development despite YY1’s role as a transcriptional repressor of sarcomeric genes\textsuperscript{15-19,57}. While a few in \textit{vivo} studies have found correlation between increased YY1 expression and decreased sarcomeric gene expression during pathological cardiac hypertrophy\textsuperscript{19}, most of the existing data are based on \textit{in vitro} luciferase assays. In the adult heart, the expression of YY1 is attenuated but increases in
response to pathological hypertrophy. The repressive function of YY1 in this context appears to involve the recruitment of HDAC5 to sarcomeric gene loci. In support of a context dependent role of YY1 in cardiac lineage, we found that the ability of YY1 to induce Gata4-dependent Nkx2.5 cardiac enhancer expression is abolished in differentiated H9C2 cells (Gregoire and Wu, unpublished data).

Conclusion

In summary, we describe here an essential and unexpected role of YY1 in early cardiac development. Our in vivo data to genetically ablation YY1 alleles and our transcriptional reporter assays in vitro as well as our ESC-based gain- and loss-of-function studies all point to a requirement for YY1 to regulate early cardiac development. Taken together, these results demonstrate a critical role for YY1 to regulate mesodermal cardiac commitment by regulating cardiac gene expression.

ACKNOWLEDGEMENTS

We thank Drs. Edward Seto and William Pu for the YY1 and Gata4 expression plasmids, respectively, Laura Pickett-Rice for her assistance with the flow cytometry, Erica Mazaika for her assistance with Illumina ChIP-sequencing, Mark Borowsky for bioinformatics support, Yuh-Shin Chang, Sek Won Kong, and Annie Tremblay for manuscript critiques.

DISCLOSURES

None.

SOURCES OF FUNDING

Financial support was provided by the MGH Fund for Medical Discovery and the American Heart Association (S.G), and by the German Research Foundation (M.K.), the German Heart Foundation (M-A. D), and by NIH Director’s New Innovator Award, the NHLBI, and the Harvard Stem Cell Institute (S.M.W.)

REFERENCES


DOI: 10.1161/CIRCRESAHA.113.259259

DOI: 10.1161/CIRCRESAHA.113.259259

FIGURE LEGENDS

Figure 1. Identification of YY1 as a putative regulator of CPC gene expression. (A) Schematic of NK ESC differentiation microarray analysis. (B) Hierarchical clustering of genes associated with eGFP+ CPCs. Color key – fold difference in expression. (C) Highly conserved regions (A,B,C,D,E) between mouse and human NKX2.5 locus were further analyzed for their over-represented TFBS. (D) Venn diagram representation of the transcription factors (TFs) identified from bioinformatics analysis. (E) Immunostaining of E9.5 mouse heart sections with antibodies against Nkx2.5, Gata4, YY1, myogenin, and AP-2α.

Figure 2. Binding of YY1 to the Nkx2.5 cardiac enhancer in vitro. (A) Putative YY1 bindings sites within the 2.1 kb Nkx2.5 cardiac enhancer. The arrows indicate the location of the primer sequences used for chromatin immunoprecipitation (ChIP). (B) Electromobility shift assays of in vitro translated YY1 to radiolabeled double-stranded oligonucleotide probes. The sequences of the wild type and mutant YY1 competitor probes are listed in red. (C) Anti-YY1 antibody supershifts (SS) the YY1/probe complex. YY1 ChIP using anti-YY1 (α-YY1) and control (α-IgG) antibodies on chromatin derived from day 6 in vitro differentiated NK ESCs (D), FACS-purified eGFP+ and eGFP- cells (E), or mouse embryonic day 9.5 heart and body cells (F). Three independent ChIP experiments were performed for each chromatin material with at least 2 different primer pairs.

Figure 3. Identification of YY1 target genes in CPCs by ChIP-sequencing. (A) ChIP peak summary. (B) Peak locations with respect to TSS. (C) Peak locations with respect to Refseq gene models. (D) Gene ontology term analysis of genes enriched in the YY1 ChIP-seq. The actual value for Benjamini-Hochberg false discovery rate (FDR) are shown. (E) Genomic plot of YY1 binding peaks at the indicated cardiac transcription factor loci. Vertical dashes indicate 5kb intervals. Y-axis denotes peak strength. (F) Expanded view of the ChIP-seq peaks located within the Nkx2.5 cardiac enhancer region. (G) Venn diagram representation of the set of genes bound by YY1 or Gata4 or both by ChIP-seq. (H) Gene ontology analysis of set of genes bound by both YY1 and Gata4.

Figure 4. YY1 regulation of the Nkx2.5 cardiac enhancer is Gata4-dependent. (A) Diagrams of the luciferase reporter constructs. Red – confirmed YY1 binding site. Pink – unconfirmed YY1 binding site. Blue – Gata4 binding site. (B) Luciferase activity of full length (WT) or truncated (1-330) reporters in the presence of Gata4 (100 ng) and YY1 (100 ng or 200 ng) or both. (C) Luciferase activity of full length (WT) or 4 or 5 binding site mutant (4BS or 5BS) reporters. (D) Luciferase activity of full length (WT) or single mutant reporters. (E) Luciferase activity of full length (WT) or Gata4-binding sites (GS1 or GS2)
mutant reporters. (F) WT and 5BS constructs were injected to generate embryonic day 9.5 transgenic embryos. Representative Xgal-stained embryos are shown. Sequential ChIP (Co-ChIP) using antibodies against YY1 followed by antibodies against Gata4 on chromatin derived from day 6 NK ESCs (G) or E10.5 embryonic hearts (H). (I) Co-immunoprecipitation of endogenous YY1 and Gata4 using α-YY1 or α-IgG antibodies and nuclear extract from H9C2 cells. Bound proteins immunoblotted with antibodies against YY1 and Gata4. (***) P<0.001, (**) P<0.05.

**Figure 5.** Mesodermal expression of YY1 is required for proper development in vivo. (A) Whole mount microscopy of day 9.0 mouse embryos from the intermating of Mesp1-Cre+; YY1 +/Flox (+/F) males with YY1 +/F females. WT and Het YY1 embryos exhibit 11 somites while Homo embryos exhibit 9 somites. Yellow bar = 10 micrometer. (B) Quantitative real-time PCR analysis of different cardiac markers from dissected embryonic day 7.5 heart region. (*** P<0.001, **) P<0.05.

**Figure 6.** YY1 is required for the Gata4-dependent transcription of Nkx2.5. (A) Basal expression level of YY1 in H9C2 cells by quantitative PCR analysis. (B) Luciferase activity of full length Nkx2.5 cardiac enhancer in the presence of Gata4 (100 ng), YY1-DN (50 to 200 ng), and YY1 (100 or 200 ng) expression constructs. (C) Generation of NK ESC lines expressing YY1-DN in a doxycycline-dependent manner. (D) Quantitative PCR analysis of YY1-DN expression in undifferentiated NK-YY1-DN ESCs in the presence or absence of 1 ug/mL doxycycline. (E) Diagram of NK-YY1-DN ESC line differentiation in the presence or absence of doxycycline. (F) The effect of YY1-DN overexpression on the emergence of eGFP+ cells at day 6 of in vitro differentiation. (G) Quantitative PCR analysis or early cardiac transcription factor expression in the presence/absence of doxycycline. (***) P<0.001, (**) P<0.05.

**Figure 7.** YY1 overexpression enhances in vitro cardiac differentiation of NK ES cells. (A) Generation of doxycycline-inducible YY1-overexpressing ESC lines (NK-YY1). (B) Quantitative PCR analysis of YY1 overexpression in undifferentiated NK-YY1 ESC line (Clone #YY5-1). (C) Diagram of in vitro cardiac differentiation of the NK-YY1 ESC line. (D) Flow cytometric analysis of YY1-induced eGFP+ cells at day 10 of in vitro differentiation. (E) Quantification of the effect of YY1 overexpression on cardiac differentiation. (F) Fluorescence microscopy of embryoid bodies at days 6 and 10 of in vitro differentiation. Note the increase in eGFP+ cell areas with doxycycline treatment. (G) Quantitative PCR analysis of the level of cardiac transcription factor expression at day 6 of differentiation. The expression of endothelial (VE-cadherin) and smooth muscle (SMAα) genes (H) as well as cardiac sarcomeric genes (I) were measured at day 10 of in vitro differentiation. Myosin light chain 2V – MLC2V; cardiac Troponin T – cTnT. (***) P<0.001, (**) P<0.05.

**Figure 8.** YY1 promotes cardiac differentiation during mesodermal transitions into CPCs. (A) Diagram of in vitro cardiac differentiation of the NK-YY1 ESC line (Clone #YY5-1) and the doxycycline treatment strategy. (B) Flow cytometric analysis of embryoid body cells at day 6 of differentiation following doxycycline treatment from days 3 to 6. The percentage of eGFP+ cells present is labeled. (C) Quantification of eGFP+ cells present at day 6 of differentiation following doxycycline treatment during days 0 to 3 (0-3) or days 3 to 6 (3-6) of differentiation. (D) Quantitative PCR analysis of Nkx2.5 and Tbx5 expression at day 6 of differentiation in cells treated with doxycycline from days 0-3 or 3-6. (***) P<0.001.
Novelty and Significance

What Is Known?

- YY1 is a context-dependent transcriptional regulator.
- YY1 inhibits sarcomeric gene expression in mature cardiomyocytes.

What New Information Does This Article Contribute?

- YY1 was found to bind and transcriptionally activate the cardiac enhancer of Nkx2.5 in vivo.
- By chromatin immunoprecipitation and genome-wide sequencing, YY1 was found to bind to the promoter/enhancer regions of other cardiac genes during early development.
- Embryos with mesodermal-specific loss of YY1 exhibit developmental arrest and do not form cardiac progenitor cells.
- In vitro gain- and loss-of-function studies in ES cells confirm that YY1 is required to promote the commitment of mesodermal precursors to cardiac progenitor cells.

The mechanism that regulates the transition of mesodermal precursors to cardiac progenitor cells (CPCs) is poorly understood. In this study, we show that YY1 is expressed in cardiac mesoderm and that it transcriptionally activates key cardiac genes such as Nkx2.5. The ability of YY1 to regulate Nkx2.5 expression is dependent on its cooperative interaction with Gata4 at a nearby chromatin. In addition, we show that cardiac differentiation of ES cells into CPCs is enhanced by YY1 overexpression and suppressed when YY1 is functionally abolished. We confirmed the in vivo requirement of YY1 by showing that mesodermal-specific deletion of YY1 results in developmental arrest and failure of heart formation. These results demonstrate an essential and unexpected role for YY1 in promoting cardiogenesis by regulating the transition of mesodermal precursors to CPCs.
Figure 1. Identification of YY1 as a putative regulator of CPC gene expression. (A) The NK ES cell line was in vitro differentiated and the gene expression profile was obtained by microarray analysis at days 4 and 5 of differentiation. (B) Hierarchical clustering of genes associated with eGFP+ cells at days 4 and 5 of differentiation. Two biological replicates of four samples were tested. The fold difference in gene expression is indicated by the color key. (C) Identification of highly conserved regions between mouse and human genome sequences at the NKX2.5 locus. The labeled (A,B,C,D,E) areas within 10 kb upstream of the transcription start site (TSS) of NKX2.5 were further analyzed for their over-represented TFBS. (D) Venn diagram representation of the transcription factors (TFs) identified from bioinformatics analysis. (E) Immunostaining of E9.5 mouse heart sections with antibodies against Nkx2.5, Gata4, YY1, myogenin, and AP-2α.
Figure 2. Binding of YY1 to the Nkx2.5 cardiac enhancer in vitro. (A) A diagram representation of the putative YY1 binding sites within the 2.1 kb Nkx2.5 cardiac enhancer. The arrows indicate the location of the primer sequences used for chromatin immunoprecipitation (ChIP) studies. (B) Binding of in vitro translated YY1 to radiolabeled double-stranded oligonucleotide probes containing the indicated YY1 binding sequence as determined by electromobility shift assays. Three independent experiments were performed in triplicate. The sequences of the wild type and mutant YY1 competitor probes are listed in red. (C) Anti-YY1 antibody supershifts (SS) the YY1/probe complex. (D) YY1 ChIP using anti-YY1 (α-YY1) and control (α-IgG) antibodies on chromatin derived from day 6 in vitro differentiated NK ESCs. (E) YY1 ChIP on chromatin derived from FACS-purified eGFP+ and eGFP- cells from day 6 in vitro differentiated NK ESCs. (F) YY1 ChIP on chromatin derived from mouse embryonic day 9.5 heart and body cells. Three independent ChIP experiments were performed for each chromatin material with at least 2 different primer pairs.
Figure 3. Identification of YY1 target genes in CPCs by genome-wide ChIP-sequencing analysis. (A) ChIP peak characteristics. (B) Peak locations with respect to TSS. (C) Peak locations with respect to Refseq gene models. (D) Gene ontology term analysis of genes enriched in the YY1 ChIP-Seq. The actual value for Benjamini-Hochberg false discovery rate (FDR) are shown. (E) Cardiac transcription factor with YY1 binding in eGFP+ but not eGFP- cells. Arrow indicates TSS. Vertical dashes indicate 5kb intervals. 0 corresponds to the TSS. The right Y-axis represents the peak strength. (F) Expanded view of the Nkx2.5 enhancer ChIP-Seq peaks along with the corresponding YY1 binding sites. (G) Venn diagram representation of common set of genes from ChIP-Seq for both YY1 and Gata4. (H) Gene Ontology term analysis of common set of genes from the YY1 and Gata4 ChIP-Seq bioinformatics analysis.
Figure 4. YY1 regulation of the Nkx2.5 cardiac enhancer is Gata4-dependent. (A) A schematic diagram of the luciferase reporter constructs. Red - confirmed YY1 binding site. Pink - unconfirmed YY1 binding site. Blue - Gata4 binding site. (B) Luciferase activity of full length (WT) or truncated (1-330) reporters that has been transfected with expression vectors for Gata4 (100 ng) and YY1 (100 ng or 200 ng). (C) Luciferase activity of full length (WT) or mutant (4BS or 5BS) reporters where four or five putative YY1 binding sites have been mutated, respectively. (D) Luciferase activity of full length (WT) or mutant (BS mutant) reporters where each confirmed YY1 binding site was mutated individually. (E) Luciferase activity of full length (WT) or mutant reporters where Gata4 binding sites (GS1 and GS2) have been individually mutated. (F) Nkx2.5 enhancer-promoter-lacZ WT and 5BS constructs were used to generate E9.5 transgenic embryos. Representative Xgal-stained embryos are shown. (G, H) Sequential ChIP (Co-ChIP) using antibodies against YY1 followed by antibodies against Gata4 on chromatin derived from day 6 in vitro differentiated NK ES cells (G) or E10.5 embryonic hearts (H). The co-occupancy of YY1 and Gata4 on the same chromatin region within the Nkx2.5 cardiac enhancer is indicated by PCR analysis. (I) Co-immunoprecipitation of endogenous YY1 and Gata4. Antibodies against YY1 (α-YY1) or control (α-IgG) were incubated with nuclear extract from H9C2 cells. Bound proteins were resolved by SDS-PAGE and immunoblotted with antibodies against YY1 and Gata4. (*** P < 0.001, ** P < 0.05)
Figure 5. Mesodermal expression of YY1 is required for proper development in vivo. (A) Whole mount microscopy of day 9.0 mouse embryos from the intermating of Mesp1-Cre+;YY1+/Flox (+/F) males with YY1+/F females. Note the developmental arrest and absence of midline fusion (yellow arrows) in Mesp1-Cre+;YY1 Flox/Flox embryo. This phenotype was present in 3 out of 3 day 9.0 Mesp1-Cre+;YY1 Flox/Flox embryos and 0 out of 8 littermate embryos with other genotypes. The wild type (WT) and heterozygous (Het) embryos show 11 somites. The homozygous (Homo) embryos show 9 somites. Yellow bar = 10 micrometer. (B) Quantitative real-time PCR of different cardiac markers from E7.5 hearts. (*** \( P < 0.001 \), (**) \( P < 0.05 \))
Essential and Unexpected Role of YY1 to Promote Mesodermal Cardiac Differentiation
Serge Gregoire, Ravi Karra, Derek Passer, Marcus-Andre Deutsch, Markus Krane, Rebecca Feistritzer, Anthony Sturzu, Ibrahim Domian, Yumiko Saga and Sean M. Wu

Circ Res. published online January 10, 2013;
Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/early/2013/01/10/CIRCRESAHA.113.259259

Data Supplement (unedited) at:
http://circres.ahajournals.org/content/suppl/2013/01/10/CIRCRESAHA.113.259259.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at:
http://circres.ahajournals.org//subscriptions/
SUPPLEMENTAL INFORMATION

Supplemental Materials and Methods

Plasmid constructs

The mammalian expression plasmids for YY1 and Gata4 were kindly provided by Dr. Edward Seto at Moffitt Cancer Center, Florida and Dr. William Pu at Boston Children’s Hospital, respectively. Both YY1 and Gata4 were amplified by PCR with the KOD Hot Start DNA polymerase (Invitrogen) and subcloned into pcDNA3.1 (Invitrogen). In addition, the YY1 sequence was introduced into a tet-inducible lentiviral vector (pLV-tetO)\(^1\). The Nkx2.5-Luciferase reporter was generated by replacing the MCK promoter from the MCK-luciferase reporter (generous gift of Dr Xiang-Jiao Yang, Molecular Oncology Group, McGill University, Canada) with the 2.5 kb enhancer/base promoter element of murine NKX2.5\(^2\). The reporter plasmid containing mutant YY1 binding sites in the Nkx2.5 cardiac enhancer was generated by PCR and subcloned in the luciferase reporter construct. The dominant negative YY1 mutant was generated by PCR cloning of fragment 119\(-\)414 of YY1 as previously described\(^3\). The Nkx2.5-LacZ construct was generated by replacing the Luciferase reporter with the LacZ reporter from pCALL vector.

ESC and H9C2 cell culture

ES cells were cultured on a monolayer of mitomycin-inactivated mouse embryonic fibroblasts and maintained in media containing DMEM with high glucose, 15% FCS, 5000 i.u./mL penicillin/streptomycin, 2 mM L-glutamine, 0.1 mM non-essential amino acid solution, 0.1 mM beta-mercaptoethanol, and 10\(^{3}\) u/mL leukemia inhibitory factor as previously described\(^2\). H9C2 cells were obtained from ATCC (Virginia) and maintained in Dulbecco’s Modified Eagle’s Medium supplemented with 10% FBS and 5000 i.u./mL penicillin/streptomycin according to vendor supplied protocol.

Microarray analysis of gene expression and identification of transcription factor binding sites

Total RNA was isolated from FACS-purified eGFP+ and eGFP- cells from days 4 and 5 of in vitro differentiation of Nkx2.5-eGFP ESCs (NK ESCs) using the SV Total RNA purification kit (Promega, Madison, WI). mRNA was linearly amplified into cRNA and hybridized onto Affymetrix oligonucleotide microarrays (MOE 430 v2). Gene expression was modeled to control for days of differentiation and for eGFP status using the following algorithm:

\[
y_{ij} = \mu_j + \beta_{GFP} x_{GFPi} + \beta_{TIME} x_{TIMEi} + \beta_{GFP\cdot TIME} x_{GFP\cdot TIMEi} + \epsilon_{ij}
\]

where \(\mu_j\) is the basal expression level, \(\beta_{GFP}\) is the effect of GFP status on gene expression, \(\beta_{TIME}\) is the effect of length of differentiation on gene expression, \(\beta_{GFP\cdot TIME}\) is the combined effect of GFP status and length of differentiation on gene expression, and \(\epsilon_{ij}\) is the error term or residual of the model\(^4\). Multiple hypothesis testing correction was performed using the FDR method with a false discovery rate of 5%. All gene expression analysis and gene ontology analysis was performed using the affy and limma package of Bioconductor. (http://www.bioconductor.org\(^5\)). Over-represented transcription factor binding sites (TFBS) for genes associated with cardiogenic fate were identified using the whole genome rVista tool (http://genome.lbl.gov/vista/index.shtml) using the assembled alignment of the Human March 2006 and Mouse February 2006 genomes.

Analysis of TFBS within the Nkx2.5 Promoter

The genomic sequence for the 20 kb upstream of the murine and human NKX2.5 were obtained from ENSEMBL (http://www.ensembl.org) and aligned using the mVISTA tool (http://genome.lbl.gov/vista/index.shtml). TF binding sites within the 2.1 kb Nkx2.5 cardiac
enhancer were identified from five regions of greatest homology using the TESS tool as previously described (http://www.cbil.upenn.edu)\(^6\).

**Lentiviral production and infection into ES cells**

Viral infections were performed with replication defective doxycycline-inducible YY1 lentiviral expression vector and a lentiviral vector constitutively expressing the rtTA\(^1\). To produce infectious viral particles, 293T cells cultured on 10-cm dishes were transfected with the LV-tetO vectors (11 mg) together with the packaging plasmids VSV-G (5.5 mg) and D8.9 (8.25 mg) using FuGENE (Roche). Viral supernatants were harvested on three consecutive days starting 24 h after transfection, yielding a total of 30 ml of supernatant per viral vector. Viral supernatant was concentrated approximately 100-fold by ultracentrifugation at 50,000g for 1.5 h at 4°C, resuspended in 200 µl of medium, aliquoted, and stored at -80°C. Infections into NK and dTomato ESCs were carried out overnight in 0.5 ml ESC medium using 30 µl of each viral concentrate per well of 6-well plate. About 10 days post-transfection, colonies were individually expanded and scored by real-time PCR for the expression of YY1 following doxycycline treatment. One cell line, YY5-1, showing the most substantial level of YY1 over-expression was chosen for the subsequent studies.

**ESC in vitro differentiation**

All in vitro differentiation assays were performed according to Huang et al.\(^7\). In brief, ES cells were cultured on feeder-free gelatin-coated dishes in the presence of LIF for one day prior to differentiation. On the day of differentiation, ESCs were made into hanging droplets and cultured as suspension embryoid bodies (EBs) from day 2 onward. On the day of flow cytometry analysis, the EBs were harvested, washed once in PBS and dissociated with collagenase A (10 mg/mL) and B (10 mg/mL) (Roche) for 1h at 37°C. The dissociated EBs were then washed twice in PBS and single-cell resuspended in HANKS, 20% FCS, and 2 µg/mL propidium iodine. Flow cytometry data was acquired by CellQuest v3.3 (BD Biosciences, San Jose, CA) using FACscalibur (BD Biosciences) and processed using FlowJo v7.6 (Tree Star, Ashland, OR) software.

**Real-time PCR**

Total RNA from in vitro differentiated ES cells was isolated with the RNeasy Plus micro kit (Qiagen) according to the manufacturer suggested protocol. Quantitative PCR was performed on cDNA made from reverse transcribed RNA using the I-script cDNA synthesis kit (BioRad, Hercules, CA). Quantitative PCR was performed using the Realplex\(^4\) mastercycler (Eppendorf) with SYBR Green substrate (BioRad, Hercules, CA) for 40 cycles. Each sample was normalized against the B-actin value. Primers sequences used for quantitative PCR analysis are available upon request.

**Electromobility Shift Assays (EMSA)**

The 10 putative YY1 binding sites within the 2.1 kb fragment of the Nkx2.5 cardiac enhancer (-9435/-7353)\(^8\) were tested for their ability to bind YY1 in vitro. First, YY1 protein was translated in vitro with a TNT T7-coupled reticulocyte lysate system (Promega, Madison, WI), using pcDNA-YY1 as DNA template. Pairs of complementary oligonucleotides containing the consensus YY1 site were annealed in annealing buffer containing 10 mM Tris-HCl, 50 mM NaCl, and 0.1 mM EDTA. The double-strand annealed oligonucleotides were then labeled for 3h at 37°C with gamma \(^32\)P-dATP (Perkin Elmer) using the T4 Polynucleotide kinase (NEB). Gel shift assays were performed in binding buffer (25 mM Hepes pH 7.9, 100 mM KCl, 12.5 mg MgCl\(_2\), 20% glycerol, 0.1% NP-40, and 1 mM DTT). Binding occurred in a total volume of 20 µl containing 1 µg of Poly dl-dC (Roche, Indianapolis, IN) and 1 µg of salmon sperm DNA, 0.4
pmol of radioactively double-strand oligonucleotides and 5 µl (5 µg) of in vitro translated protein. The reactions were incubated for 15 minutes on ice. Unlabeled competitor and mutant competitor (YY1 sites mutated to GC rich sequence) in 10- and 100-fold excess were added to the mixture to test for specificity of DNA binding (which corresponds to 4 and 40 pM of unlabeled competitor DNA, respectively). The DNA-protein complex was separated for 3h by electrophoresis on a 6% non-denaturing polyacrylamide gel in 1x TBE. The gel was dried on a Whatman 3M paper for 2h and then exposed overnight on an X-ray film (Denville Scientific Inc. Metuchen, NJ). Three independent experiments were performed in triplicate.

Chromatin immunoprecipitation (ChIP) and ChIP-sequencing analysis

NK ESCs were in vitro differentiated and on day 6 of differentiation, embryoid bodies were dissociated with collagenase A (10 mg/mL) and B (10 mg/mL) (Roche, Nutley, NY) for 60 min at 37°C into single cells. FACS-purified eGFP+ or eGFP- cells or whole embryoid body (EB) cells were then isolated and fixed in formaldehyde and subjected to chromatin immunoprecipitation as previously described. DNA fragments were resuspended in 40 µl H2O, and 1 µl was used for 40 cycles of PCR with primers (5′-CCA GTC TGG GTC CTA ATG CGG GTG GCC TCT-3′ and 5′-ATC TAC TCG CCC GTC GCC TGT ACA AAC CCT-3′) spanning the first three putative YY1 sites in the Nkx2.5 cardiac enhancer. The expected PCR fragment was 0.27 kb.

For chromatin immunoprecipitation using embryonic hearts, timed-pregnant CD1 female mice from the Jackson laboratory were dissected on E9.5. Hearts from 78 embryos were collected and processed as described above.

For co-chromatin immunoprecipitation (Co-ChIP) of YY1 and Gata4, the chromatin was first pulled down with the anti-YY1 antibody overnight at 4°C. Subsequently, 45 µl of 50% slurry protein A/salmon sperm DNA agarose beads (Millipore) was added. After rotation for 2 h at 4°C, the beads were washed as described previously. The DNA fragments were not de-crosslinked but were purified using phenol/chloroform extraction. The purified DNA was diluted in 0.5 mL of ChIP dilution buffer. The DNA fragments were then pulled down with the anti-Gata4 antibody (Santa Cruz, C-20) overnight at 4°C.

The DNA fragments obtained from chromatin immunoprecipitation of embryonic hearts were amplified using the REPLI-g Mini Kit according to manufacturer’s suggested protocol and separated on a 0.8% agarose gel to be visualized by ethidium-bromide staining.

For ChIP-sequencing analysis, the de-crosslinked ChIP DNA fragments were converted into a DNA library according to manufacturer suggested protocol (Illumina® GAII, San Diego, California) with modifications. PicoGreen measurement of the library concentration was determined by Quant-it PicoGreen dsDNA reagent (Invitrogen, Carlsbad, California). This value was further confirmed using quantitative PCR against a standard reference sample. The size distribution of the library was analyzed using an Agilent 2100 Bioanalyzer prior to Next Generation Sequencing using an Illumina GAII sequencer by the Biopolymer Core Facility at Harvard Medical School. The resulting sequence files containing 60 million reads per lane in FASTQ format were analyzed on the Galaxy platform (http://main.g2.bx.psu.edu/). Alignment of sequencing reads to the genome was performed using Bowtie. Standard parameters were used. MM9 reference genome was used for the alignment. FATSQ files were used as the input. The output file format was SAM, which was then converted into BAM for peak calling. MACS platform was used for peak calling at a p value cutoff peak detection of 1e-05, tag size of 25, band width of 300, and regions with MFOLD high-confidence enrichment ratio against background to build model of 32. To identify the nearest genes to the peak binding sites, a custom-designed software made by the Molecular Biology Core at MGH was used. Analysis was limited to the strongest peaks and genes located within 20 kb from the binding peak. Gene ontology analysis was performed using the DAVID bioinformatics database.
Co-immunoprecipitation (Co-IP)

To examine the direct interactions of YY1 with Gata4, EBs from day 6 of in vitro differentiation were collected, briefly washed once with PBS and dissociated with collagenase A and B for 1h at 37°C as described above. These cells were then washed twice with PBS and lysed in 500 µl of buffer containing 20 mM Tris-HCl [pH 8.0], 150 mM KCl, 10% glycerol, 5 mM MgCl₂, 0.1% NP-40, and protease inhibitors. Half of the extract was used for immunoprecipitation with 2 µg of anti-YY1 antibody (Cat# SC-1703, Santa Cru, Biotechnology) and the other half with 2 µg of the anti-GFP antibody (Cat#290, Abcam) Immunoprecipitation was performed overnight at 4°C. After rotation overnight at 4°C, 45 µl of 50% slurry protein A/salmon sperm DNA agarose beads was added. After rotation for 2 h at 4°C, the beads were washed four times with buffer B. Bound proteins were eluted in 30 µL of buffer B, boiled 10 minutes, separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and detected by immunoblotting with anti-YY1 and anti-Gata4 (SC25310, Santa Cruz Biotechnology) antibodies. PBS with 0.1% Tween 20 and 5% dry milk was used for membrane blocking and antibody incubation. Blots were developed with Supersignal chemiluminescent substrates (Pierce, Rockford, IL).

Luciferase assays

SuperFect transfection reagent (Qiagen, Valencia, CA) was used to transiently transfet a luciferase reporter plasmid (200 ng) and/or mammalian expression plasmids (100 to 200 ng) into H9C2 cells. pBluescript KSII(+) was used to normalize the total amount of plasmids used in each transfection, and pCMV-β-Gal (50 ng) was cotransfected for normalization of transfection efficiency. After 48 h, cells were lysed in situ, and luciferase reporter activity was determined by using D-(-)-luciferin (Boehringer Mannheim, Indianapolis, IN) as the substrate. Galactosidase activity was measured with Galacto-Light Plus (Tropix®) as the substrate. The chemiluminescence from activated luciferin or Galacto-Light Plus® was measured on a Luminometer plate reader (Berthold, Huntsville, Alabama). Each transfection was performed at least three times.

Immunohistochemistry

Embryos were fixed in 4% paraformaldehyde for 10 minutes and flash frozen in OCT and stored at -80°C until processing. The sections were then rinsed once in PBS for 10 minutes and blocked in 10% goat serum/PBS 0.1% saponin. Primary antibodies were incubated with the sections overnight at 4°C. Following three PBS washes (20 minutes/each), secondary antibodies conjugated with the appropriate fluorochrome were added to each section for 60 min at 37°C. The sections were then washed with PBS three times for 20 minutes each and mounted in Prolong® antifade mounting medium (Invitrogen, Carlsbad, California). Primary antibodies used were as follow: YY1 (Santa Cruz; Sc-1703), Gata4 (Santa Cruz; Sc-9053), MLC2A (Synaptic Systems; 311-011), Nkx2.5 (R&D Systems; MAB2444), Isl-1 (Developmental Studies Hybridoma Bank; 39.4D5).

Mesoderm-specific loss of YY1 mouse embryos

All mouse-related studies were approved previously by the Subcommittee on Research Animal Care at Massachusetts General Hospital. To obtain mouse embryos that are deficient in YY1 expression specifically in cardiac mesoderm, we interbreed mice that harbor Mesp1-Cre+; YY1 Flox/+ alleles with mice carrying Mesp1-Cre-; YY1 Flox/+ alleles (a kind gift from Dr. Yang Shi at Children’s Hospital Boston) to generate embryos that carry homozygous YY1 Flox/Flox
alleles and Mesp1-Cre. At days 8.5, 9.0, 9.5, and 11.5 post coitum we euthanize the pregnant mother and dissected the embryos from the surrounding yolk sac and uterus. These embryos were subsequently imaged on a whole mount dissection microscope and their genotypes determined by PCR analysis of tail biopsy samples.

**Statistical Analysis**
Statistical analysis for all comparative studies were performed using the 2-tailed Student’s t-test. Statistical analysis for bioinformatics data are described within the section on such studies.

**Supplemental Figure Legends**

**Online Figure I.** (A) YY1 binding to chromatin sites near CNN1, SEMA3C, and FGFBP1 loci in FACS-purified eGFP⁺ cells. Note the absence of YY1 binding peaks in the same region in eGFP⁻ cells. (B) Confirmation of YY1 binding to promoter regions of CPC-enriched genes in CPCs by conventional PCR analysis. The genome region corresponding to the PCR amplicon is indicated by the red bar in (A). (C) Confirmation of YY1 binding to promoter regions of CPC-enriched genes in CPCs by quantitative PCR analysis. (D) Confirmation of YY1 binding to promoter regions of CPC-enriched genes in E9.5 embryonic hearts by conventional PCR analysis. (E) Confirmation of YY1 binding to promoter regions of CPC-enriched genes in E9.5 hearts by quantitative PCR analysis. (**) P < 0.05, (***) P < 0.001.

**Online Figure II.** Identification of YY1 targeted cardiac genes in CPCs by genome-wide ChIP-sequencing analysis. (A) Representative sarcomeric genes with YY1 binding in eGFP⁺ but not eGFP⁻ cells. (B) Predicted TFBS within genomic regions bound by YY1 according to bioinformatics analysis. The percentage of total YY1 reads that showed predicted binding to the indicated transcription factor is listed. (C) Cardiomyocyte specific genes among the 195 identified genes common to the YY1 and Gata4 ChIP-Seq. (**) P < 0.05, (***) P < 0.001.

**Online Figure III.** The proximal half (-8485/-7353) of the Nkx2.5 cardiac enhancer harbors an essential element for transcriptional activation of the full-length enhancer. (A) Schematic diagram of the YY1 and Gata4 binding sites in the full-length enhancer and the luciferase reporter constructs used. (B) Luciferase reporter activity from full-length (WT) or proximally truncated (1-950) reporter plasmids when co-transfected into H9C2 cells along with Gata4 (100 ng) and YY1 (100 and 200 ng) expression plasmids. (C) Overexpression of YY1 in H9C2 cells. The expression of YY1 and Gata4 was confirmed by Western blotting using anti-YY1 and anti-Gata4 antibodies, respectively. (D) Overexpression of Gata4 in H9C2 cells. The expression of YY1 and Gata4 was confirmed by Western blot using anti-YY1 and anti-Gata4 antibodies, respectively. (*** P < 0.001.

**Online Figure IV.** Immunohistological analysis of day 7.5 (E7.5) wild type and mutant YY1 embryos. (A) H&E staining of E7.5 embryos. (B) Immunohistochemical analysis of YY1 expression and co-expression of Nkx2.5, Isl-1, and Gata4 in wild type (+/+), heterozygous (+/F), and mutant (F/F) E7.5 mouse embryos. Note the loss of YY1 expression in mesodermal derivatives (yellow arrows). There are no identifiable mesodermal Nkx2.5+ or Isl-1+ cells in the region that represents the second heart fields (SHF) in the mutant (F/F) embryo (white arrows).

**Online Figure V.** Overexpression of YY1 enhances cardiac differentiation of NK ES cells (A) Diagram of the YY1 overexpression lentiviral plasmid. (B) YY1 expression by quantitative PCR analysis at days 6 or 10 of in vitro differentiation in the presence or absence of doxycycline treatment. (C) Quantitative PCR analysis of cardiac transcription factor expression at day 10 of
*in vitro* differentiation in the presence or absence of doxycycline treatment. (**) $P < 0.05$, (***) $P < 0.001$.

**Online Figure VI.** Absence of enhanced endodermal or ectodermal differentiation by YY1 overexpression in ES cells. NK−YY1 ESCs were *in vitro* differentiated in the presence or absence (Ctrl) of doxycycline treatment from days 0 to 3 (0-3) or 3 to 6 (3-6). The expression of markers for endodermal (FGF5) (A), ectodermal (Nestin) (B), and mesodermal sublineages (C) such as skeletal muscle (Myogenin, MyoD) and blood (Gata1) were assessed by quantitative PCR analysis. (**) $P < 0.05$, (***) $P < 0.001$. 
References

Online Figure II

A

B

<table>
<thead>
<tr>
<th>Genes</th>
<th>% of reads</th>
<th>Consensus Binding sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATA</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>AP2</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>CEBPα</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>NCX</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Pax2</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>En1</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>FOXC1</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>ZIC3</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>TST1</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>BRCA1</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>YY1</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

C

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>actn4</td>
<td>Alpha-actinin 4</td>
</tr>
<tr>
<td>casz1</td>
<td>Zinc finger protein castor homolog 1</td>
</tr>
<tr>
<td>mylk</td>
<td>Smooth muscle myosin light chain kinase</td>
</tr>
<tr>
<td>myo5a</td>
<td>Myosin Va</td>
</tr>
<tr>
<td>tbx3</td>
<td>T-box 3</td>
</tr>
<tr>
<td>tbx18</td>
<td>T-box 18</td>
</tr>
<tr>
<td>tpm2</td>
<td>Tropomyosin beta chain</td>
</tr>
</tbody>
</table>
Online Figure III

A. Nkx2.5 cardiac enhancer

B. Luciferase activity (Fold over control)

C. YY1 overexpression

D. Gata4 overexpression
Online Figure V

A

B

Day 6

YY1 Expression
(Fold over Dox -)

Dox
- +

0  20  40  60

**

**

Day 10

Dox
- +

0  20  40  60

**

**

C

Day 10

Relative Expression
(Fold over Dox -)

Dox
- +

0  10  20  30  40  50  60

***

**

**

***

**

**

***

**

**

**

Nkx2.5
IsI1+
Gata4
Tbx5
Hand1
Hand2
**Online Table I.** Identification of over-represented transcription factor binding sites within 5 kb upstream of transcription start sites of CPC-enriched genes.

<table>
<thead>
<tr>
<th>TF</th>
<th>Total #Hits</th>
<th># Hits in genome</th>
<th>-log (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MINI20</td>
<td>51</td>
<td>30053</td>
<td>9.4684</td>
</tr>
<tr>
<td>E47</td>
<td>28</td>
<td>11753</td>
<td>8.4787</td>
</tr>
<tr>
<td>GLI</td>
<td>53</td>
<td>35007</td>
<td>8.1544</td>
</tr>
<tr>
<td>MYOGENIN</td>
<td>30</td>
<td>14571</td>
<td>7.6286</td>
</tr>
<tr>
<td>AREB6</td>
<td>35</td>
<td>19072</td>
<td>7.5474</td>
</tr>
<tr>
<td>SP3</td>
<td>38</td>
<td>23151</td>
<td>6.9093</td>
</tr>
<tr>
<td>AP4</td>
<td>48</td>
<td>33239</td>
<td>6.8725</td>
</tr>
<tr>
<td>NFKAPPAB</td>
<td>33</td>
<td>18887</td>
<td>6.688</td>
</tr>
<tr>
<td>LBP1</td>
<td>29</td>
<td>15762</td>
<td>6.4087</td>
</tr>
<tr>
<td>SMAD3</td>
<td>38</td>
<td>24813</td>
<td>6.1897</td>
</tr>
<tr>
<td>CMYB</td>
<td>32</td>
<td>20915</td>
<td>5.3236</td>
</tr>
<tr>
<td>ERR1</td>
<td>45</td>
<td>34637</td>
<td>5.3172</td>
</tr>
<tr>
<td>YY1</td>
<td>15</td>
<td>6013</td>
<td>5.1082</td>
</tr>
<tr>
<td>STAF</td>
<td>31</td>
<td>20894</td>
<td>4.9264</td>
</tr>
<tr>
<td>LMO2COM</td>
<td>28</td>
<td>18443</td>
<td>4.6906</td>
</tr>
<tr>
<td>NFE2</td>
<td>17</td>
<td>9153</td>
<td>4.0957</td>
</tr>
<tr>
<td>TFIIA</td>
<td>40</td>
<td>33913</td>
<td>3.9255</td>
</tr>
<tr>
<td>ZID</td>
<td>31</td>
<td>23763</td>
<td>3.9185</td>
</tr>
<tr>
<td>AML</td>
<td>22</td>
<td>14678</td>
<td>3.7569</td>
</tr>
<tr>
<td>AP2ALPHA</td>
<td>18</td>
<td>10896</td>
<td>3.6783</td>
</tr>
<tr>
<td>LEF1TCF1</td>
<td>40</td>
<td>34906</td>
<td>3.6763</td>
</tr>
<tr>
<td>NRSF</td>
<td>10</td>
<td>3982</td>
<td>3.6384</td>
</tr>
<tr>
<td>TAL1ALPHAE47</td>
<td>11</td>
<td>4791</td>
<td>3.6113</td>
</tr>
<tr>
<td>HNF4ALPHA</td>
<td>41</td>
<td>37460</td>
<td>3.3599</td>
</tr>
<tr>
<td>ATF4</td>
<td>27</td>
<td>21283</td>
<td>3.3195</td>
</tr>
<tr>
<td>MEIS1</td>
<td>28</td>
<td>22446</td>
<td>3.307</td>
</tr>
<tr>
<td>FOXO1</td>
<td>39</td>
<td>35798</td>
<td>3.1924</td>
</tr>
<tr>
<td>ETS1</td>
<td>24</td>
<td>18391</td>
<td>3.1862</td>
</tr>
<tr>
<td>NRF2</td>
<td>31</td>
<td>26590</td>
<td>3.1161</td>
</tr>
<tr>
<td>DR3</td>
<td>35</td>
<td>31451</td>
<td>3.0855</td>
</tr>
<tr>
<td>TAL1BETA47</td>
<td>7</td>
<td>2583</td>
<td>2.9002</td>
</tr>
<tr>
<td>R</td>
<td>12</td>
<td>6857</td>
<td>2.8585</td>
</tr>
<tr>
<td>GCM</td>
<td>8</td>
<td>3389</td>
<td>2.8571</td>
</tr>
<tr>
<td>ARP1</td>
<td>19</td>
<td>14041</td>
<td>2.8246</td>
</tr>
<tr>
<td>CREBATF</td>
<td>19</td>
<td>14222</td>
<td>2.7632</td>
</tr>
<tr>
<td>RSRFC4</td>
<td>7</td>
<td>2736</td>
<td>2.7598</td>
</tr>
<tr>
<td>FOXO4</td>
<td>31</td>
<td>28117</td>
<td>2.7469</td>
</tr>
<tr>
<td>TEF1</td>
<td>15</td>
<td>10105</td>
<td>2.7214</td>
</tr>
<tr>
<td>HES1</td>
<td>24</td>
<td>20683</td>
<td>2.53</td>
</tr>
<tr>
<td>ETS2</td>
<td>14</td>
<td>9744</td>
<td>2.4558</td>
</tr>
<tr>
<td>SREBP</td>
<td>8</td>
<td>4006</td>
<td>2.4172</td>
</tr>
<tr>
<td>E12</td>
<td>11</td>
<td>6826</td>
<td>2.4015</td>
</tr>
<tr>
<td>LFA1</td>
<td>25</td>
<td>22988</td>
<td>2.2628</td>
</tr>
<tr>
<td>PEBP</td>
<td>17</td>
<td>13681</td>
<td>2.2324</td>
</tr>
</tbody>
</table>

The frequency of appearance of these sites within the 5 kb promoter region (Total # Hits), the frequency of their appearance within the entire genome (# Hits in genome), and the p value for the over-representation are listed.
Online Table II. Identification of overrepresented transcription factor binding sites (TFBS) within five highly conserved genomic regions upstream from the transcript start site of *NKX2.5*.

<table>
<thead>
<tr>
<th>Genomic Region</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transcription factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YY1</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Pou1f1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GATA</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>AP-2?</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Isl-1</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>SRF</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Myogenin</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETS-1</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Cutl-1</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>ATF-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>CREB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>NF-E2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Oct-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The five genomic regions (A, B, C, D, E) are diagrammed in Figure 1D. The numbers listed are the start and end positions of each region relative to the transcription start site of murine *NKX2.5*. Within these conserved regions, the presence of over-represented TFBS are indicated.
Online Table III

ChIP sequencing analysis confirms the preferential binding of YY1 to promoter regions of CPC-enriched genes.

<table>
<thead>
<tr>
<th>CPC-enriched genes</th>
<th>Non-CPC-enriched genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genes* Peaks** Confirmed^</td>
<td>Genes Peaks Confirmed</td>
</tr>
<tr>
<td>Col5a1 Acta1 Tnnc1 SM22? Acta2 Actc1 Fgfbp1 Myl9 Cnn1 Sema3c Samd4 Sfn Fam162b Itga3</td>
<td>Pmp22 Edrg1 Ccng2 Lfitm1 x x Ephx2</td>
</tr>
</tbody>
</table>

* List of genes from Figure 1C

** Peaks are identified by YY1 occupancy in the 5 kb upstream or intronic regions of each gene on ChIP sequencing analysis

^ Confirmed represents peaks that show enrichment by conventional and quantitative PCR analysis.
**Online Table IV.** Results from the interbreeding of Mesp1 Cre;YY1 Flox/+ male with YY1 Flox/+ females.

<table>
<thead>
<tr>
<th></th>
<th>Mesp1-cre</th>
<th>YY1 genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+/+</td>
<td>+/F</td>
</tr>
<tr>
<td><strong>e8.5</strong></td>
<td>8/5 (18.6/12.5)</td>
<td>7/11 (16.3/25)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>8/5 (18.6/12.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mesp1-cre</th>
<th>YY1 genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+/+</td>
<td>+/F</td>
</tr>
<tr>
<td><strong>e9.5</strong></td>
<td>8/8 (12.5/12.5)</td>
<td>19/16 (30/25)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>9/8 (14/12.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mesp1-cre</th>
<th>YY1 genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+/+</td>
<td>+/F</td>
</tr>
<tr>
<td><strong>e11.5</strong></td>
<td>7/11 (7.7/12.5)</td>
<td>35/23 (38/25)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>12/11 (13.2/12.5)</td>
</tr>
</tbody>
</table>

The data shown represents actual/expected number of embryos at embryonic days 8.5 (e8.5), 9.5 (e9.5), and 11.5 (e11.5) for each genotype. The actual/expected percentages of each genotype are displayed in parentheses next to the corresponding embryo number. + = wild type allele. F = Floxed allele.