A Large Permissive Regulatory Domain Exclusively Controls Tbx3 Expression in the Cardiac Conduction System

Jan Hendrik van Weerd1, Ileana Badi1, Malou van den Boogaard1, Sonia Stefanovic1, Harmen J G van de Werken2,3, Melisa Gomez-Velazquez4, Claudio Badia-Careaga4, Miguel Manzanares4, Wouter de Laat2, Phil Barnett1, Vincent M Christoffels1

1Dept. of Anatomy, Embryology & Physiology, Academic Medical Center, Amsterdam, The Netherlands; 2Hubrecht Institute-KNAW and University Medical Center Utrecht, Utrecht, The Netherlands; 3Dept. of Cell Biology, Erasmus MC Rotterdam, Rotterdam, The Netherlands, and; 4Cardiovascular Development and Repair Department, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain.

J.H.V.W. and I.B. contributed equally to this study.

Running title: Locus Confined Control of Cardiac Tbx3 Expression

Subject codes:
[6] Cardiac development
[139] Developmental biology
[141] Functional genomics
[143] Gene regulation

Address correspondence to:
Dr. Vincent M. Christoffels
Heart Failure Research Center
Academic Medical Center
Meibergdreef 15 L2-108
1105 AZ Amsterdam
The Netherlands
Tel: +31 20 5667821
v.m.christoffels@amc.uva.nl

In May 2014, the average time from submission to first decision for all original research papers submitted to Circulation Research was 14.87 days.

DOI: 10.1161/CIRCRESAHA.115.303591
ABSTRACT

**Rationale:** The evolutionary conserved Tbx3/Tbx5 gene cluster encodes T-box transcription factors that play crucial roles in the development and homeostasis of the cardiac conduction system (CCS) in human and mouse. Both genes are expressed in overlapping patterns and function in strictly tissue-specific and dose-dependent manners, yet, their regulation is poorly understood.

**Objective:** To analyse the mechanism underlying the complex regulation of the Tbx3/Tbx5 cluster.

**Methods and Results:** By probing the 3D architecture of the Tbx3/Tbx5 cluster using high-resolution 4C-sequencing in vivo, we found that its regulatory landscape is in a preformed conformation similar in embryonic heart, limbs and brain. Tbx3 and its flanking gene desert form a 1 Mbp loop between CTCF binding sites that is separated from the neighbouring Tbx5 loop. However, Ctcf inactivation did not result in transcriptional regulatory interaction between Tbx3 and Tbx5. Multiple sites within the Tbx3 locus contact the promoter, including sites corresponding to regions known to contain variations in the human genome influencing conduction. We identified an atrioventricular-specific enhancer and a pan-cardiac enhancer that contact the promoter and each other and synergize to activate transcription in the atrioventricular conduction system.

**Conclusions:** We provide a high-resolution model of the 3D structure and function of the Tbx3/Tbx5 locus and show that the locus is organized in a preformed, permissive structure. The Tbx3 locus forms a CTCF-independent autonomous regulatory domain with multiple combinatorial regulatory elements that control the precise pattern of Tbx3 in the CCS.

**Keywords:** Cardiac conduction system, chromatin topology, enhancer, heart development, Tbx3/Tbx5, morphology, gene expression/regulation, transcription factor genetics, cardiac development

**Nonstandard Abbreviations and Acronyms:**

- 4C-seq: Circular chromosome conformation capture and sequencing
- AV: Atrioventricular
- CCS: Cardiac conduction system
- CTCF: CCCTC-binding factor
- ChIP-seq: Chromatin immunoprecipitation and sequencing
- E: Embryonic day
- GWAS: Genome-wide association study
- Tbx: T-box transcription factor
INTRODUCTION

The heart comprises working muscle and an integrated cardiac conduction system (CCS) that initiates and propagates the electrical impulse required for the coordinated contraction of the chambers. The development of the CCS is regulated by transcription factors that act in strictly stage, location and dose-dependent manners. Disrupted function or expression of these factors may lead to disorders in development or function of the CCS, which are associated with morbidity, heart failure and sudden death. The evolutionary conserved \( Tbx3/Tbx5 \) gene cluster encodes T-box transcription factors that play key roles in the development and function of multiple organ systems including heart, limbs, lungs and liver. Haploinsufficiency of \( TBX5 \) causes Holt-Oram syndrome (OMIM: 142900), a developmental disorder characterized by hand-heart defects, whereas haploinsufficiency of \( TBX3 \) causes the ulnar-mammary syndrome (OMIM: 181450), a developmental disorder characterized by limb, mammary gland, tooth and genital abnormalities. Both genes play crucial roles in the development and patterning of the CCS. In the heart, \( Tbx5 \) is expressed in a pattern including the atria and CCS and acts as an activator of genes promoting conduction. \( Tbx3 \) is expressed specifically in the CCS components and functions by imposing a pacemaker phenotype on cells within its expression domain. Ectopic expression of Tbx3 in the myocardium induces a pacemaker phenotype. A decrease in Tbx3 below critical levels results in lethal arrhythmias. Therefore, the correct spatio-temporal pattern and dosage of expression of \( Tbx3 \) and \( Tbx5 \) are critical for development and functional integration of the CCS.

The regulatory architecture of loci harboring genes encoding transcription factors or other developmental factors often comprises long-range regulatory elements, each responsible for a subset of the temporal, spatial and/or quantitative aspects of expression of the gene. Genome-wide association studies (GWAS) suggest that a large number of non-coding variants underlie the increased risk to various common diseases, often by disrupting regulatory elements that control the expression of nearby genes. GWAS revealed that common variants in the non-coding regions around \( TBX3 \) and \( TBX5 \) are associated with cardiac impulse conduction and depolarization properties, whereas a congenital heart defect patient has been identified with a homozygous variant in an enhancer regulating \( TBX5 \). These findings further underscore the critical importance of understanding the regulatory system controlling the expression pattern and dosage of Tbx3 and Tbx5.

The three-dimensional (3D) organization of chromatin has been recognized as an important regulatory layer for gene regulation, setting the stage for physical contact of the control sequences, such as enhancers, and their target promoters. High-resolution 4C-seq (circular chromosome conformation capture combined with sequencing) is a recent technology examining the spatial organization of DNA and measuring the contact frequencies of a chosen genomic site, or ‘viewpoint’, with the entire genome. We assessed the 3D architecture of the \( Tbx3/Tbx5 \) cluster using 4C-seq to gain insight into its regulation and to find potential genomic targets that could underlie congenital defects. Our work provides a high-resolution model of the 3D structure and function of the \( Tbx3/Tbx5 \) locus, which reveals that the \( Tbx3 \) and \( Tbx5 \) loci form autonomous regulatory domains that do not interact, and that the gene desert upstream of \( Tbx3 \) harbours multiple regulatory elements that synergistically control the precise pattern of \( Tbx3 \) in the CCS.

METHODS

An expanded Methods section is provided in the Online Data Supplement.

Fore- and midbrain, limbs and hearts were dissected from E10.5 embryos and processed to obtain single-cell samples for circular chromosome conformation capture (4C), which was performed as described.
Bacterial artificial chromosomes were modified as described. Enhancer fragments were amplified from genomic DNA and cloned in the Hsp68-LacZ reporter vector. Embryos were isolated for F0 screening, fixed in 4% paraformaldehyde in PBS and processed for in-situ hybridisation and immunohistochemistry.

Animal work was performed in accordance with national and international guidelines and approved by the Animal Experiments Committee of the Academic Medical Center.

RESULTS

Physical separation of Tbx3 and Tbx5 regulatory domains.

To gain insight into the regulation of Tbx3, we assessed the 3D architecture of the Tbx3/Tbx5 cluster using high-resolution 4C-seq of embryonic day (E) 10.5 mouse hearts. The 4C-seq experiments show results typical for chromosome conformation capture experiments, in that most of the signal is close to the viewpoint, intrachromosomal contacts outnumber interchromosomal contacts, and contacts at distal sites tend to cluster. To assess the conformation of the Tbx3 locus and identify putative regulatory regions that contact Tbx3 and may regulate its expression, the Tbx3 promoter was used as viewpoint. The broad peak around the Tbx3 promoter reflects contact by linear proximity, whereas more distant signal clusters (peaks) represent discernable interactions through looping. The capture profile reveals that the Tbx3 promoter contacts multiple regions upstream of Tbx3, with interactions reaching up to neighbouring gene Med13l, located 903 kb upstream (Figure 1A). Analysis of these data reveals relatively strict boundaries of the contact profile, with most contacts being confined to the domain spanning approximately 900 kb upstream to 50 kb downstream of Tbx3 (Figure 1A). Interestingly, analysis of chromosomal contacts with TBX3 and TBX5 as assessed by Hi-C in human fibroblasts reveals, albeit in much lower resolution, a similar pattern with multiple regions in the TBX3 domain contacting the gene. Again, contacts with the TBX5 or MED13L domains are almost absent (Online Figure I).

The contact profile of the promoter of the adjacent Tbx5 gene reveals an independent, neighboring, regulatory domain with boundaries near Tbx3 and the neighbouring Rbm19 gene, located approximately 230 kb downstream of Tbx5 (Figure 1B). The two domains hardly overlap, indicating that sites in the Tbx3 domain do not contact the Tbx5 domain and vice versa. Med13l is isolated from these regulatory landscapes, as revealed by the contact profile of the Med13l promoter (Online Figure I). In contrast, the contact profile of Rbm19, downstream of Tbx5, partially overlaps with that of Tbx5 but not with the Tbx3 regulatory landscape (data not shown). This demonstrates that Tbx3 forms a large independent regulatory domain not utilized by the neighbouring genes Tbx5, Med13l or Rbm19. The strict boundaries of the contact profiles correspond with the location of CTCF-binding sites in heart and other tissues. We studied the interaction profile of one of the CTCF-binding sites in between Tbx3 and Tbx5. This site physically interacts with distal CTCF-binding sites-containing regions located 974 kb up- and 358 kb downstream, forming chromatin loops that demarcate the Tbx3 and Tbx5 regulatory landscapes, respectively (Figure 1C). These observations suggest that CTCF may mediate the physical separation of the regulatory domains of Tbx3 and Tbx5 (Figure 1E), preventing enhancer sharing between these genes. To test this possibility, we analysed the expression of Tbx3 and Tbx5 in embryos in which Ctf5 was inactivated specifically in the cardiomyocytes. However, neither the level nor the pattern of Tbx3 and Tbx5 were affected, indicating that CTCF is not required to prevent the separation of the regulatory sequences of Tbx3 and Tbx5 (Online Figure II B,C).

Active enhancers are marked by acetylation of histone H3 on lysine K27 (H3K27ac). ChIP-sequencing with an antibody against H3K27ac in both embryonic chamber and atrioventricular canal
(AVC) myocardial cells revealed that the region flanking \( Tbx3 \) is extensively acetylated on H3K27 in AVC cells, corresponding to the AVC-specific expression of \( Tbx3 \) (Online Figure III). In contrast, the region flanking \( Tbx5 \) is acetylated in both chamber myocardium and AVC, corresponding to its expression in both compartments. The boundary between the complementary H3K27ac signatures colocalizes roughly with the boundary of the contact profiles for \( Tbx3 \) and \( Tbx5 \). Taken together, our data indicate that the \( Tbx3 \) locus forms an independent regulatory domain which is physically separated from that of \( Tbx5 \), and that this domain contains multiple regions that contact the \( Tbx3 \) gene.

GFP-modified BACs recapitulate atrioventricular conduction system expression of \( Tbx3 \).

While 4C-seq physically wires gene promoters to potential enhancer sequences, physical contact does not necessarily reflect functional interaction. To identify regulatory elements within the \( Tbx3 \) locus, we generated transgenic mice with modified bacterial artificial chromosomes (BACs). BAC RP23-143N21 contains genomic sequences from -73 kb to +147 kb relative to the transcription start site of \( Tbx3 \), BAC RP23-366H17 contains sequences from -82 kb to +78 kb and BAC RP24-89K7 contains sequences from -149 to +66 kb. To investigate whether regulatory elements required for cardiac \( Tbx3 \) expression are located within the regions covered by these BACs, we modified them by inserting a GFP-encoding reporter gene at the translation start site of \( Tbx3 \) (143N21-GFP, 366H17-GFP, 89K7-GFP; Figure 2A), also resulting in the inactivation of \( Tbx3 \) within the BAC, negating any potential gain of \( Tbx3 \) function in transgenic mice. Analysis of the GFP expression pattern in F0 embryos of embryonic day (E) 11.5 revealed that all three modified BACs are able to recapitulate many aspects of the extracardiac \( Tbx3 \) expression pattern. GFP expression was found in the retina, limb buds, mammary glands and pharyngeal mesenchyme (described earlier for 366H17-GFP), corresponding to the endogenous \( Tbx3 \) expression pattern (Figure 2B). In situ hybridisation revealed that 143N21-GFP and 366H17-GFP recapitulate \( Tbx3 \) expression in the foregut endoderm and ventral body wall (Figure 2C). Cardiac GFP expression was detected in the AVC myocardium, but in a domain smaller than that of \( Tbx3 \) itself. The GFP expression domain included ventral, right-sided and dorsal aspects (primordial atrioventricular (AV) node) of the AVC. The \( Tbx3 \)-positive AV bundle primordium (interventricular ring) did not express GFP, indicating that the region from -82 kb to +147 kb of \( Tbx3 \) lacks DNA sequences for \( Tbx3 \) activation in the complete CCS. GFP expression was also observed in \( Tbx3 \)-positive cardiac neural crest cells distal from the outflow tract and the outflow tract cushions (neural crest-derived cells) (Online Figure IV). 89K7-GFP, spanning an additional 60 kb region upstream of the other BACs, drove GFP expression in the entire AVC, recapitulating the pattern of \( Tbx3 \). Furthermore, GFP expression was observed in the interventricular ring, indicating that the region -149 kb to -82 kb upstream of \( Tbx3 \) harbours regulatory elements required for \( Tbx3 \) expression in the prospective AV bundle. These data reveal that the genomic region of -82 kb to +147 kb of \( Tbx3 \) harbours regulatory elements that drive AV activity of \( Tbx3 \), whereas the region from -149 kb to -82 kb of \( Tbx3 \) harbours additional regulatory sequences driving the full AVC and bundle expression pattern of \( Tbx3 \). We observed no GFP expression in the sinoatrial node, indicating that the regulatory elements driving \( Tbx3 \) expression in this region are absent from the regions spanned by the tested BACs.

Two putative enhancers physically contact the \( Tbx3 \) promoter.

To identify potential enhancers within the \( Tbx3 \) BACs conferring tissue-specific expression, we analyzed the contact profile of the promoter by 4C in embryonic heart, brain and limb cells. Contact profiles of all three cell types revealed a highly similar pattern with multiple sites contacting the \( Tbx3 \) promoter (Figure 3A). Two regions showed strong interaction with the promoter, one approximately 25kbp upstream of the promoter and present in all three modified BACs, and one at approximately 90 kbp upstream of the promoter and present in only BAC 89K7-GFP. To identify putative enhancers within the contact regions, we evaluated ChIP-sequencing datasets of the genome-wide occupancy profiles of cardiac transcription factors Gata4, Nkx2-5, Tbx5, Tbx3, the enhancer-associated histone
acetyltransferase p300, RNA Polymerase II (Pol2) and histone modifications H3K27ac and monomethylation of histone H3 on lysine K4 (H3K4me1) (Figure 3B). Based on these occupancy profiles two candidate regions were identified upstream of Tbx3. Putative enhancer A (eA) is located 19 kb upstream of Tbx3, within the proximal contact region and is occupied by Tbx5, Gata4 and Nkx2-5 and marked by H3K4me1. Putative enhancer B (eB), which shows occupancy by all factors except Pol2, is located 95 kb upstream of Tbx3 within the distal contact region that is present only in BAC 89K7 (Figure 3B,C, Online Figure V). Tbx3, H3K27ac, p300 and Pol2 occupancy of eA and Pol2 occupancy of eB were not called by MACS. However, inspection of the raw data tracks revealed distinguishable peaks for all these factors on both sites (Online Figure VI A). The weak signal might be explained by the fact that the ChIP-seq datasets are derived from whole hearts, whereas the CCS in which Tbx3 is active comprises only a small fraction of the total cardiac cell population. The association of both enhancers with the enhancer-associated histone modifications H3K27ac and H3K4me1 was confirmed by ChIP-quantitative PCR (ChIP-qPCR). Both putative enhancers showed enrichment of H3K27ac and H3K4me1 occupancy over a negative control region, revealing their possible role as transcriptional enhancers (Online Figure VI B).

We used putative eA and eB as viewpoints in 4C-seq analysis, and found that eA strongly interacts with both eB and the Tbx3 promoter. The same holds for eB, albeit to a much lesser extent (Figure 4 and Online Figure VII). These findings reveal that both enhancers and the Tbx3 promoter are spatially clustered together. Comparison of contact profiles from heart, limb and brain indicated that the chromosomal conformation of the Tbx3 locus is not tissue-specific, since the contact profiles for all three cell types with different viewpoints reveal a striking similarity (Figure 4). Only minor differences in contact profiles are observed between heart and limb. In brain, slightly more appreciable differences are observed in the contact profiles. The region upstream of Tbx3, encompassing both eA and eB, shows more enhancer-promoter contacts in heart as compared to brain. Furthermore, the Tbx3 promoter contacts Tbx5 and several flanking intergenic regions in brain, exceeding the boundaries marked by CTCF binding sites, and vice versa (Figure 4B). The separation between both domains is therefore less strict in embryonic brain.

**eA and eB are required for CCS expression of Tbx3.**

To investigate whether putative enhancers eA and eB are involved in the regulation of Tbx3 expression, fragments containing the enhancers were deleted from 89K7-GFP to construct ΔeAeB-89K7-GFP. Inspection of GFP expression in E10.5 F0 ΔeAeB-89K7-GFP embryos revealed that extracardiac Tbx3 expression was largely unaffected by deletion of both fragments (Figure 5). GFP expression was detected in the limb buds, eye placode, pharyngeal arches and cardiac neural crest cells in a pattern that is similar to 89K7-GFP, indicating that Tbx3 expression in these regions is not mediated by eA and eB. However, GFP expression in the heart is completely abolished in ΔeAeB-89K7-GFP embryos. Immunohistochemistry revealed that GFP expression in the AVC, AV bundle, foregut endoderm and ventral body wall is absent in ΔeAeB-89K7-GFP embryos (Figure 5). These findings reveal putative enhancers eA and eB are required for CCS expression of Tbx3.

**Evolutionary conserved eA and eB are sufficient to drive atrioventricular Tbx3 expression.**

To elucidate whether both enhancer fragments on their own are sufficient to drive atrioventricular Tbx3 expression, we isolated fragments of approximately 2 kb spanning eA or eB, cloned them in a LacZ-reporter vector with a minimal promoter (eA-LacZ and eB-LacZ, respectively) and assessed their in vivo activity. F0-screening of transient transgenic eA-LacZ E10.5 embryos revealed that eA is sufficient to drive LacZ expression in the AVC, strong in the ventral, right and dorsal aspects, weak at the left side (Figure 6A). This cardiac expression pattern corresponds to the GFP expression as observed in BACs 143N21-GFP and 366H17-GFP, in which eA but not eB is located. Consistent LacZ expression was also...
observed in the pharyngeal arches, whereas ventral body wall expression was not detected in any of the embryos.

eB-LacZ embryos showed a variable expression pattern. Two embryos showed cardiac expression in only few spots in the heart not correlating with the pattern of endogenous \textit{Tbx3} expression, whereas two other embryos showed cardiac expression throughout the entire heart (Figure 6B). This indicates that on its own eB is cardiac-specific, is weak and influenced by repressive effects of the genomic region at the site of its integration, and, within the heart, lacks spatial information. Interestingly, when coupled to eA and tested in transient transgenic mouse embryos (eAeB-LacZ), robust expression was observed throughout the entire AVC and prospective AV bundle (interventricular ring; Figure 6C). Strong ventral body wall expression was also observed in eAeB-LacZ embryos, in contrast to either eA-LacZ or eB-LacZ. This expression pattern in eAeB-LacZ embryos is highly similar to the GFP-expression we observed in 89K7-GFP embryos, corresponding to the presence of both eA and eB in BAC 89K7. These data imply that eA and eB do not act as independent modules, but act in synergy, whereby eB enhances the activity of eA and complements the pattern of eA in the AV conduction system precursors (Figure 6G).

To investigate whether the mechanism of \textit{Tbx3} regulation in the AV conduction system is evolutionary conserved we tested the human and chicken orthologues of eA in vivo. Human eA-LacZ showed an expression pattern comparable to its murine counterpart, with expression in the major portion of the AVC, mainly in the ventral, right and dorsal side (Figure 6D). Weak activity was also observed in the prospective AV bundle. Chicken eA-LacZ shows similar AV-ring expression, however, expression in the prospective AV bundle was not detected in any of the embryos (Figure 6E). The chicken eA-LacZ construct also showed clear expression around the fore limbs, possibly reflecting \textit{Tbx3} expression in the zone of polarizing activity implicated in limb bud development, which was not detected in the human or murine orthologue of eA. The inter-species consistency in AVC activity of eA indicates that aspects of the \textit{Tbx3}-mediated formation of the CCS are evolutionary conserved.

\textit{Tbx3} regulation in the AVC by eA is mediated by BMP signalling.

We recently found that several AV canal-specific enhancers are GATA-binding site dependent and are synergistically activated by Gata4 and Smads (downstream BMP-signalling effectors), which recruit histone acetyltransferases such as p300. Since the enhancers here are occupied by Gata4 and p300, we hypothesized that they may be activated by Gata4 and BMP-signalling. In transient transfection assays, we found that eA and eB are not activated by Gata4 or BMP effectors alone. Co-transfection with Gata4 and BMP effectors, however, resulted in a synergistic induction of eA (Figure 6F). Transfection of two sub-fragments of eA revealed that the regulatory information of this enhancer is harboured within eA2. Consistently, this sub-fragment also drives AV canal expression in transgenic embryos (not shown). The eB enhancer fragment was not strongly activated. Together these data are consistent with the observed activity patterns of the two enhancer fragments \textit{in vivo}, in which eA provides AV canal specificity whereas eB is weak and does not show specificity within the heart.

DISCUSSION

The T-box transcription factors \textit{Tbx5} and \textit{Tbx3} are crucial for the development and function of the heart and its electrical components. In this study, we provided a high-resolution model of the 3D conformation of the evolutionary conserved \textit{Tbx3}/\textit{Tbx5} genomic locus using high-resolution 4C-seq \textit{in vivo}. We found that \textit{Tbx3} and \textit{Tbx5} are physically organized into separate loci and that the genomic organization is independent of tissue-type. We identified two enhancers within the \textit{Tbx3} domain that
contact each other and the \textit{Tbx3} promoter, and that synergize to drive \textit{Tbx3} expression in the AV conduction system.

\textit{Tbx3} and \textit{Tbx5} form an evolutionary conserved cluster of developmental transcription factor genes originating from a primitive T-box gene.\textsuperscript{37} Studies on the transcriptional regulation of clustered developmental genes, like the \textit{Irx} and \textit{Hox} clusters, revealed that regulatory DNA elements are not uniquely associated with single promoters.\textsuperscript{38,39} Instead, genes within these clusters extensively share enhancers, coordinating the regulation of their expression pattern. It has been proposed that enhancer sharing explains the conservation of the genomic organization of such clusters during evolution.\textsuperscript{38,39} Our data indicate that, contrary to the \textit{Irx} and \textit{Hox} gene clusters, the \textit{Tbx3/Tbx5} cluster is divided into two separate regulatory domains. Interdomain contacts appear almost absent in all tissue types we tested, indicating that \textit{Tbx3} and \textit{Tbx5} are regulated independently. This also implies that genomic variation affecting regulatory elements of either one of the genes does not directly influence the expression pattern and level of the other.

The strict boundaries between the \textit{Tbx3} and \textit{Tbx5} domains are marked by CTCF binding sites. CTCF-bound sites can function as boundary regions, interacting with each other to physically segregate the chromosome into topologically associated domains (TADs) within which enhancer-promoter contacts have been shown to be particularly frequent.\textsuperscript{26,40,41} CTCF binding sites have been suggested to isolate genes or gene clusters from neighbouring transcriptional interference. However, in cardiac-specific \textit{Ctcf} mutants we did not observe any ectopic expression of \textit{Tbx3} in the larger \textit{Tbx5} expression domain or any change in the level of expression of either gene. These data indicate that although CTCF binding marks regions involved in looping and locus separation, CTCF itself is not required to transcriptionally insulate the loci from neighbouring regulatory activity. Consistently, inactivation of \textit{Ctcf} in erythroid cells or limb buds did not lead to ectopic activation of genes adjacent to the \textit{\(\beta\)-globin} locus or \textit{Hoxd} locus, respectively.\textsuperscript{27,42}

Tissue-specific gene expression is mediated by specific interactions of the promoter with distal tissue-specific enhancers, and as such a tissue-specific 3D chromosomal conformation is expected for different cell types. Contrary to this, however, our data indicate that the chromosomal architecture of the \textit{Tbx3/Tbx5} locus is strikingly similar between different tissue types. Contact profiles in heart and limb cells reveal a fixed chromosomal architecture with multiple sites contacting the \textit{Tbx3} promoter, including eA and eB. In brain cells, the contact profile appears somewhat different, displaying slightly more \textit{Tbx3}-\textit{Tbx5} interdomain contacts than other tissues tested (Figure 4B,C). \textit{Tbx3} is expressed in the developing heart and limbs, whereas its expression in the developing brain is very limited or absent. The observed variation in the topology of the \textit{Tbx3} locus might thus correspond to its inactivity in this tissue. We propose that the genomic contacts of eA, eB and the \textit{Tbx3} promoter are mostly tissue-independent and organized in a preformed, permissive chromatin structure, in which functional enhancer-promoter interactions are facilitated by the rigid, cell-type invariant structural enhancer contacts.\textsuperscript{19} The enhancer regions and the promoter are in close proximity irrespective of tissue type, which thus facilitates CCS-specific transcription factor complexes in the activation of \textit{Tbx3} transcription in the CCS. In other cell types the CCS-specific enhancers are not activated but remain in close proximity to the promoter.

Analysis of the chromosomal interactions in human fibroblasts assessed by Hi-C\textsuperscript{24} reveals a similar pattern as observed in mouse. Multiple regions upstream of \textit{TBX3} contact the gene and these regions do not overlap with regions contacting \textit{TBX5} or \textit{MED13L} (Online Figure I). This indicates that the genomic organization of the \textit{Tbx3/Tbx5} locus is not only conserved between tissues, but also between mammalian species, which reveals a robust conservation of the topological \textit{Tbx3/Tbx5} domains. Although CTCF appeared to be not required for the insulation of the \textit{Tbx3} regulatory domain, it has been well-described as a factor frequently present at the boundaries of topologically associated domains. The highly
similar occupancy profiles of CTCF between different tissue types (Online Figure II A) therefore further indicates that the 3D organization of the Tbx3 locus is largely conserved and cell type-independent.

GWAS have identified common genetic variants in the human TBX3/TBX5 locus that influence CCS function.\textsuperscript{16,17} Tbx3 and Tbx5 are required for functional development, maturation and homeostasis of the CCS in a highly dosage-sensitive manner, indicating that minor changes in regulatory sequences could potentially have consequences for CCS function and homeostasis.\textsuperscript{12,43} Based on the contact profile of the Tbx3 and Tbx5 promoters and their separate regulatory domains, we propose that all the regulatory elements mediating Tbx3 or Tbx5 expression, respectively, are confined within their respective domains marked by CTCF binding sites. This indicates that the functional effect of genetic variation is limited to either the TBX3 or TBX5 domain, respectively, and within those domains to the regions contacting the respective promoters. Our data therefore facilitates the assignment of function to genomic variants identified in GWAS.

Identification of enhancers driving CCS-specific Tbx3 expression is crucial if one is to fully understand the mechanisms underlying CCS development and function. We report the identification of two enhancers within the Tbx3 domain that synergistically drive the AV conduction system expression of Tbx3. Major components of the Tbx3+ AV conduction system are the slow-conducting AV node and rings, and the fast-conducting AV bundle and bundle branches. These components derive from the embryonic AV canal and crest of the interventricular septum, respectively, which are interconnected throughout development.\textsuperscript{44} The AV canal component is evolutionarily conserved in vertebrates.\textsuperscript{45} Our data reveal that the regulatory sequences regulating Tbx3 in these respective components are separated in the genome, and suggest that eA is responsible for driving Tbx3 expression in the conserved AVC component through a BMP signalling-mediated mechanism. Although on its own weak, eB appears to synergize with eA to drive robust Tbx3 expression in the AVC and AV bundle precursor region in the interventricular septum. We speculate that eB may have been recruited during the evolution of the AV bundle, a structure, as yet, only found in birds and mammals. These elements may come together to form a single regulatory unit, leading to an expression pattern that is wider than the sum of each isolated element alone.\textsuperscript{46,47} The contribution of presumably weak elements such as eB could therefore be to provide robustness to the complex regulatory structure.

Taken together, our data show that the Tbx3 locus forms a conserved, independent regulatory unit physically separated from neighbouring loci. The chromosomal topology of this unit is in a preformed and permissive organization, which is tissue independent and evolutionary conserved. Multiple sites within this unit synergize to control the precise expression of Tbx3 in the CCS. This study provides insight into the tight relation between chromatin topology and the strict regulation of the evolutionary conserved Tbx3/Tbx5 gene cluster.
ACKNOWLEDGEMENTS
We thank Corrie de Gier-de Vries, Vincent Wakker and Karel van Duijvenboden for their technical assistance, and Niels Galjart (Erasmus MC, Rotterdam) for providing the Ctef-floxed mice.

SOURCES OF FUNDING
This work was supported by grants from the Rembrandt Institute of Cardiovascular Sciences, Academic Medical Center, Amsterdam (to V.M. Christoffels), the Netherlands Heart Foundation (NHS grant 2010B205, to V.M. Christoffels) and by the European Community’s Seventh Framework Programme contract (“CardioGeNet” 223463, to V.M. Christoffels). Work in the lab of Miguel Manzanares was funded by the Spanish Government (grant BFU2011-23083), the CNIC Translational Programme (grant CNIC-08-2009) and the ProCNIC Foundation.

DISCLOSURES
None

REFERENCES

3. Wolf CM, Berul CI. Inherited conduction system abnormalities—one group of diseases, many genes. J Cardiovasc Electrophysiol. 2006;17:446-455
32. 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;108:5632-5637

DOI: 10.1161/CIRCRESAHA.115.303591


FIGURE LEGENDS

Figure 1. Separation of the Tbx3 and Tbx5 regulatory domains. (A-D) Contact profiles of the Tbx3 and Tbx5 loci as determined by 4C reveal genomic regions that physically contact the Tbx3 promoter in embryonic heart cells. The upper track shows the UCSC genome browser traces of a running window approach for viewpoints of the Tbx3 promoter (A), Tbx5 promoter (B) and CTCF site (C,D). The second track shows the respective normalized contact intensities (gray dots) and their truncated mean trends (black line) for all viewpoints in and around Tbx3 and Tbx5. Means are computed for 5 kb windows and the grey band displays the 20–80% percentiles for these windows. The scaling-plot below depicts data points across differently scaled window sizes (from 2 kb (top row) to 50 kb (bottom)). The data points and the colors are scaled to the maximum mean value of a 5 kb window size. Local changes in colour codes indicate regions enriched for captured sequences, which correspond to the promoter-enhancer contacts described. The contact profile of Tbx5 (B) reveals a complementary pattern to that of Tbx3 (A), with hardly any overlap in contacting regions. (C) The contact profile of the CTCF-binding site in between Tbx3 and Tbx5 (marked by arrow in (D)) as determined by 4C reveals interaction with neighbouring CTCF-binding site islands (marked by asterisks). (D) Distribution of CTCF binding sites as determined by ChIP-seq in heart cells. 33 CTCF-binding sites are located in between Tbx3 and Tbx5 and near flanking genes Med13l and Rbm19, colocalizing with the boundaries in contact profiles of Tbx3 and Tbx5 promoters. Arrow marks CTCF site used for 4C. (E) Schematic overview of the Tbx3 and Tbx5 loci, depicting the physical separation of the Tbx3 and Tbx5 regulatory domains.

Figure 2. GFP-modified BACs recapitulate AV conduction system expression of Tbx3. (A) Tbx3 locus with regions spanned by BACs. A GFP cassette was inserted in the translational start site of Tbx3 to obtain BACs 143N21-GFP, 366H17-GFP and 89K7-GFP. (B) Whole-mount in-situ hybridisation on E11.5 embryo showing endogenous Tbx3 expression pattern. GFP expression patterns in F0 embryos of 143N21-GFP, 366H17-GFP and 89K7-GFP reveals recapitulation of many aspects of extracardiac Tbx3 expression pattern. Numbers depict number of embryos with reproducible GFP expression patterns among the total number of independent transgenic embryos. (C) In-situ hybridisation on sections through 366H17-GFP and 89K7-GFP reveals Tbx3 and GFP expression patterns. Arrows and arrowheads depict expression in AVC and interventricular region, respectively. (D) Table depicting GFP-positive regions in 89K7-GFP, 366H17-GFP and 143N21-GFP embryos. av; atrioventricular, avc; atrioventricular canal, fg; foregut, fl; forelimb, la; left atrium, lv; left ventricle, ra; right atrium, rv; right ventricle, san; sinoatrial node, oft; outflow tract, cnccs; cardiac neural crest cells.

Figure 3. Identification of two putative enhancers contacting Tbx3. (A) Contact profile in heart with Tbx3 promoter as viewpoint as assessed by 4C reveals multiple sites contacting Tbx3. Y-axis depicts contact tag count. (B) Occupancy profiles of the Tbx3 locus of cardiac transcription factors Tbx5, Tbx3, Gata4 and Nkx2, histone modifications H3K4me1 and H3K27ac and enhancer-associated proteins p300 and Pol2 and as assessed by ChIP-seq. Peaks were called using the MACS peak calling algorithm. (C) Two potential enhancers are identified which are present in all three modified BACs (eA) or exclusively in 89K7-GFP (eB).

Figure 4. eA and eB physically interact with Tbx3 in a tissue-independent manner. (A) Contact profiles for Tbx3 promoter, eA and eB as assessed by 4C in different cell types. Arrows indicate points of view for the depicted tracks. Both eA and eB physically contact the Tbx3 promoter and vice versa. Asterisks indicate contacts close to Tbx5 with eA and the Tbx3 promoter. (B,C) Overlap of heart and brain contact profiles for Tbx3 promoter (B) and eA (C) reveal minor differences. In heart, more contacts are observed in the region upstream of Tbx3 for both viewpoints (asterisks), whereas in brain more contacts with the Tbx5 promoter are observed (double asterisks).
**Figure 5. eA and eB are required for AVCS expression of Tbx3**
89K7-GFP recapitulates (extra)cardiac expression pattern of *Tbx3*. Deletion of both eA and eB from 89K7-GFP (∆eAeB-89K7-GFP) leads to complete loss of cardiac GFP expression, whereas extracardiac expression is unaffected. Rectangles depict regions shown in greater magnification below. Circle depicts loss of GFP expression in the heart. Arrows depict atrioventricular canal, arrowheads depicts interventricular region.

**Figure 6. eA and eB are sufficient to drive AVC Tbx3 expression and are evolutionary conserved**
(A) Transient transgenic eA-LacZ embryos (E10.5) reveals that eA is sufficient to drive Tbx3 expression in the AVC and part of the interventricular septum. (B).eB-LacZ expression is observed in only few spots in the heart or throughout the entire heart (not shown). (C) eA and eB synergize to drive robust Tbx3 expression in the AVC and interventricular septum. (D,E) Human (D) and chicken (E) orthologues of eA drive expression in the AVC indicating evolutionary conservation of eA. (F) Luciferase reporter assay on eA and its two subfragments, eB and a previously described tandemly repeated Tbx2 enhancer. Cos-7 cells were co-transfected with Gata4 and Smad expressing vectors. Bars represent average ± s.d. (n=3). (G) Model depicting the synergy between eA and eB to drive AVC expression of *Tbx3*. 
Novelty and Significance

What Is Known?

- The T-box transcription factors Tbx3 and Tbx5 play crucial roles in the development and function of the cardiac conduction system.
- Tissue-specific gene expression is mediated by genomic regulatory sequences that act in a time, tissue and dose-dependent manner.
- Common genetic variants in non-coding sequences close to Tbx3 and Tbx5 are associated with cardiac conduction system parameters.

What New Information Does This Article Contribute?

- The Tbx3/5 genomic cluster is organized in a pre-formed three-dimensional structure, independent of tissue-type.
- The Tbx3 genomic locus is physically separated from that of Tbx5, indicating both genes do not share regulatory sequences to drive their transcription.
- We identified two genomic enhancers that are required and sufficient to drive Tbx3 expression in the atrioventricular conduction system.

Tbx3 and Tbx5 are crucial for the development and function of the cardiac conduction system. Recent studies have identified common variations in the human genome close to Tbx3 and Tbx5 to be associated with cardiac impulse conduction. We show here that the Tbx3/5 genomic locus is organized in a pre-formed three-dimensional conformation, independent of tissue-type. The Tbx3 domain is physically separated from the neighboring Tbx5 domain, indicating both genes do not share regulatory sequences to drive their transcription. This suggests that the functional effect of genetic variation in either one of the domains can be exclusively attributed to the respective gene, which facilitates the assignment of function to genomic variations associated with cardiac conduction. Furthermore, we identified two transcriptional enhancers that are required and sufficient to drive Tbx3 expression in the atrioventricular conduction system. Together, these findings provide insight into the close relationship between chromatin conformation and the complex regulation of the Tbx3/5 gene cluster.
Figure 2

A Mammalian Conservation

RP23-143N21
RP23-368H17
RP24-89K7

-73 kb
-82 kb
-149 kb

78 kb
66 kb

100 kb

B

<table>
<thead>
<tr>
<th>Tbx3</th>
<th>143N21-GFP</th>
<th>366H17-GFP</th>
<th>89K7-GFP</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>eye</td>
<td>heart</td>
<td>fl</td>
<td>3/3</td>
</tr>
</tbody>
</table>

C

<table>
<thead>
<tr>
<th>Tbx3</th>
<th>GFP</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>ra</td>
<td>la</td>
</tr>
<tr>
<td>rv</td>
<td>lv</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>366H17-GFP</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image7.png" alt="Image" /></td>
</tr>
<tr>
<td>ra</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>89K7-GFP</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image9.png" alt="Image" /></td>
</tr>
<tr>
<td>ra</td>
</tr>
</tbody>
</table>

D Tbx3 endogenous expression

<table>
<thead>
<tr>
<th>89K7-GFP</th>
<th>366H17-GFP</th>
<th>143N21-GFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAN</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AV canal</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>AV cushions</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>AV bundle</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>OFT cushions/cNCCs</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pericardium</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>foregut endoderm</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>tongue muscle precursors</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>limb buds</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pharyngeal mesenchyme</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>retina</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

* GFP expression observed in AVN primordium and ventral/cranial and right side of AVC
Figure 3
Figure 5
Figure 6
A Large Permissive Regulatory Domain Exclusively Controls Tbx3 Expression in the Cardiac Conduction System
Jan H van Weerd, Ileana Badi, Malou van den Boogaard, Sonia Stefanovic, Harmen J van de Werken, Melisa Gomez-Velazquez, Claudio Badia-Careaga, Miguel Manzanares, Wouter de Laat, Phil Barnett and Vincent M Christoffels

Circ Res. published online June 24, 2014;
Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 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/2014/06/24/CIRCRESAHA.115.303591
Data Supplement (unedited) at:
http://circres.ahajournals.org/content/suppl/2014/06/24/CIRCRESAHA.115.303591.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 MATERIAL

DETAILED METHODS

Circular chromosome conformation capture

4C was performed as previously described. In short, fore- and midbrain, limbs and heart were dissected from E10.5 embryos and processed as described to obtain single-cell samples. Approximately 10^7 cells were used for fixation and lysis, and cross-linked chromatin was digested with DpnII (New England Biolabs, cat.no. R0543M). Samples were treated with T4 DNA ligase (Roche, cat.no. 10799009001) to obtain large DNA circular molecules, encompassing multiple restriction fragments held together by cross-links. After cross-link removal by heating, DNA was digested with Csp6I (Fermentas, cat.no. ER0211). Final ligation of restriction fragments yields circular DNA templates for 4C PCR amplification. For all experiments, 200 ng of the resulting 4C template was used for the subsequent PCR reaction, of which 16 (total: 3.2 μg of 4C template) were pooled and purified for next-generation sequencing. The PCR products were purified using two columns per sample of the High Pure PCR Product Purification Kit (Roche cat. no. 11732676001). Inverse PCR primers specific for fragment of interest (viewpoint) are listed in Table S1. Primers were designed and data analysis for mapping and normalizing 4C-seq data was performed as described in ref. In short, the rich fragment pool (around 6-8 fragment ends per 1kb) was used to generate statistically robust semiquantitative contact maps in the 10kb to 1Mb region surrounding the viewpoint. Moreover, the high-resolution restriction site grid was used to quantify contact intensities in genomic windows varying in size from a few kilobases to tens of kilobases. Chromosomal contacts in human fibroblasts were assessed using Hi-C data generated by Jin et al.

BAC modification and generation of transgenic mice

BACs RP24-89K7 and RP23-143N21 were obtained from C57BL/6J mouse BAC libraries (CHORI, BACPAC Resources) and modified to generate 89K7-GFP and 143N21-GFP. A GFP reporter gene cassette was inserted in the translational start site of Tbx3 according to the two-step BAC modification protocol as described in ref. 3. Generation of the 366H17-GFP mouse line was described previously. To obtain ΔeAeB 89K7-GFP, eA and eB enhancer sequences were sequentially deleted from GFP-89K7 using the abovementioned protocol with shuttle vectors in which both fragments were absent. Modified BAC DNA was purified with the NucleoBond PC 20 kit (Macherey-Nagel) according to manufacturer's instructions and injected into mouse pronuclei of FVB/N mice. Microinjected zygotes were implanted into pseudo-pregnant females (E0.5) and transgenic embryos were isolated in PBS at the indicated day, visualized by fluorescence microscopy, then fixed in 4% paraformaldehyde in PBS for 4 h and processed for in situ hybridization or immunohistochemistry. At least two independent transgenic lines per BAC construct were analysed to define the GFP expression pattern and to account
for variation in expression pattern due to position effect. The sequences of primers used for cloning and genotyping are available upon request. All animal work was approved by the Animal Experimental Committee of the Academic Medical Center, University of Amsterdam, and carried out in compliance with Dutch government guidelines.

**Ctcf conditional knockout**
The Ctcf floxed allele and Nkx2.5-Cre line were previously described. Mice were bred in the core animal facility in the Centro Nacional de Investigaciones Cardiovasculares in accordance with national and European legislation. Three pools of 2 to 4 wildtype and Ctcf KO hearts were used for qRT-PCR. 500 ng of total RNA was reverse transcribed using random primers (Applied Biosystems). Expression changes were normalized to actin. PCR was performed in an AB7900_384 (Applied Biosystems) using Sybrgreen (Applied Biosystems) as reporter. qRT-PCR primers are available upon request. Homo- or heterozygous mice for the Ctcf-floxed allele that did not carry the Nkx2.5-Cre were used as controls.

**In-situ hybridisation**
After fixation in 4% PFA in PBS, embryos were embedded in paraplast and sectioned at 10 μm. Non-radioactive in situ hybridization on sections was performed as previously described. Probes for Tbx3 and GFP have been described in ref. Sections were observed with a Zeiss Axiophot microscope and photographed with a Leica DFC320 Digital Camera.

**Immunohistochemistry**
After fixation, embryos were embedded in paraplast and sectioned at 8-10 μm for immunohistochemistry. Anti-Tbx3 (E-20, sc-31656) goat polyclonal (1:200; Santa Cruz Biotechnology Inc.), anti-GFP (ab5450) goat polyclonal (1:200; Abcam) primary antibodies and Alexa Fluor 488 donkey anti-goat IgG, Fluor 568 donkey anti-goat IgG or Fluor 680 donkey anti-goat IgG (1:250; Invitrogen) secondary antibodies were used. Stained sections were photographed with a fluorescence microscope Leica DM6000.

**ChIP-sequencing**
ChIP-sequencing data sets for Nkx2-5, Gata4, Tbx5 and Tbx3 were described in refs. ChIP-seq datasets for p300, Pol2, CTCF and DNase hypersensitivity sites were generated by the ENCODE consortium. ChIP-sequencing datasets for H3K27ac are described in ref. Peaks were called using the peak caller MACS, using corresponding input control tissue, and by an in-house peak calling algorithm dubbed Occupeak (de Boer, van Duijvenboden et al., PLoS-One, in press).
Identification of consensus binding sites

Consensus binding sites were identified using JASPAR\textsuperscript{13} and TRANSFAC\textsuperscript{14} software. Sequences were tested for enrichment of binding sites for cardiac transcription factors Nkx2-5, Gata4, T-box factors and Bmp signaling effector Smads. Relative profile score threshold for JASPAR was set to 85%. TRANSFAC parameters were set at the default setting for vertebrate sequences and the minFP setting to minimize false positive scores.

Chromatin immunoprecipitation

E14.5 embryos were dissected, and the hearts were isolated and collected in cold PBS. ChIP was performed using the True MicroChIP kit (Diagenode) according to the manufacturer’s protocol. The antibodies used for ChIP were anti-H3K27ac (0.5μg, Abcam ab4729) and anti H3K4me1 (0.5μg, Abcam ab8895). Quantitative PCR was performed on a Roche LightCycler 480 System using Sybr Green detection. Fold enrichment indicates the ratio of ChIPed DNA to a negative control region, normalized for input DNA. Primer sequences are provided in Table S2.

In vivo enhancer screening assay

Mouse eB (chr5:120,024,484-120,025,600; NCBI37/mm9), 3.4kb-eB (chr5:120,024,225-120,027,612), eA (chr5:120,100,494-120,102,560), eAb (chr5:120,101,525-120,102,566), human eA (chr12:115,140,877-115,142,783; NCBI36/hg18) and chicken eA (chr15:12,509,382-12,511,893; WUGSC2.1/galGal3) enhancer sequences were PCR-amplified and cloned in the HindIII restriction site of the Hsp68-LacZ reporter vector.\textsuperscript{15} eA was subcloned into the SmaI site of eB-Hsp68-LacZ vector to obtain the eAeB-Hsp68-LacZ construct. After vector backbone excision, linearized fragments were injected into the pronucleus of 0.5-day-old fertilized FVB/N eggs, which were transferred into the oviducts of CD-1 pseudopregnant foster females (Cyagen Biosciences). Embryos were harvested and stained with X-gal to detect LacZ activity.

In vitro reporter assay

eA (chr5:120,100,494-120,102,560), eA1 (chr5:120,100,494-120,101,620), eA2 (chr5:120,101,526-120,102,560), eB (chr5:120,024,484-120,025,600) and eAeB were cloned in the pGL2-SV40 luciferase construct containing a minimal promoter (Promega). (Tbx2)4 construct was previously described.\textsuperscript{12} The constructs were transfected in Cos-7 cells using PEI as transfection reagent (DNA:PEI ratio 1:4). Constructs were co-transfected with expression constructs for Gata4, Smad1, Smad4 (all cloned in pcDNA3.1) and constitutively active Bmp-receptor Alk3 (Alk3CA; pCS2.BmpR1a.CA). Cell extracts and luciferase assays were performed according to the manufacturer’s protocol (Promega). Mean luciferase activities and standard deviations (s.d.) were
plotted as fold activation compared to the empty pGL2-SV40 expression plasmid. Transfections were carried out in triplo.
Online Figure V

>Tbx3_eA (2067bp; chr5:120,100,494-120,102,560)

> Tbx3_eB (1117bp; chr5:120,024,484-120,025,600)

Consensus binding sites for Gata4, Nkx2-5, Tbx3/5 and Smad identified by JASPAR\textsuperscript{13} and TRANSFAC\textsuperscript{14}.
SUPPLEMENTAL FIGURE LEGENDS

Online Figure I. Conservation of topological Tbx3/Tbx5 domains
Contact profiles with Med13l, Tbx3 and Tbx5 promoters as viewpoints in mouse heart (upper tracks, 4C data) and human fibroblasts (lower tracks, Hi-C data generated by Jin et al.). Arrows depict point of view for each track. Similar to the mouse genome, contacts with TBX3 in human are restricted to the upstream domain and do not overlap with regions contacting MED13L (upstream) and TBX5 (downstream).

Online Figure II. cKO of CTCF in cardiomyocytes does not result in altered gene expression
(A) UCSC browser views of ChIP-seq data (ENCODE consortium) depicting occupancy of CTCF in heart, cerebellum and limb (E14.5). Lower traces depict peaks called by MACS, indicating the high similarity of occupancy profiles between the three tissue types. (B) In-situ hybridisation on sections through WT and Ctf/- embryos at E10.5 reveal no misexpression of Tbx3 and Tbx5 in the heart. (C) qPCR validation of CTCF, Tbx3 and Tbx5 mRNA levels. CTCF mRNA levels are significantly reduced in Ctcf/- embryos, whereas Tbx3 and Tbx5 levels were unaffected.

Online Figure III. Cell-type specific H3K27 acetylation associates with gene activity
Occupancy profiles of H3K27ac in the Tbx3 and Tbx5 locus as assessed by ChIP-seq in AVC and chamber myocardium. Y-axis in upper tracks depicts normalized tag count. Lower tracks depict peaks called by our in-house peakcalling algorithm called OccuPeak (de Boer, van Duijvenboden et al., PLoS-One, in press).

Online Figure IV. GFP-modified BACs recapitulate Tbx3 expression in neural crest-derived cell
In-situ hybridisation and immunohistochemistry on sections through 89K7-GFP, 143N21-GFP and 366H17-GFP (not shown) embryos at E11.5 reveal GFP expression in cardiac neural crest cells flanking the outflow tract (arrows) and the outflow tract cushions (neural crest-derived cells; asterisks).

Online Figure V. Overview of transcription factor binding sites in eA and eB
Enrichment of binding sites for cardiac transcription factors Nkx2-5, Gata4, T-box factors and Bmp signalling effector Smads as revealed by JASPAR and TRANSFAC software.

Online Figure VI. Transcription factor occupancy and histone modifications of eA and eB
(A) UCSC browser view of ChIP-sequencing datasets, revealing the occupancy of eA and eB by Tbx3, p300, Pol2 and H3K27ac. Y-axis depict normalized tag count. (B) ChIP-qPCR confirming eA and eB...
are marked by H3K4me1 and H3K27ac. Y-axis depicts fold activation over a negative control region. Bars represent mean±s.d. Experiments are performed in duplo.

**Online Figure VII. Physical interactions of Tbx3, eA and eB**

Contact profiles of with Tbx3 promoter, eA and eB as viewpoints (arrows) reveal both enhancers and the Tbx3 promoter physically contact each other (asterisks). The legend for the color-coded scaling plots can be found in the legend of Figure 1.
### Online Table I. 4C-seq primer sequences

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tbx3_promoter_F</td>
<td>AATGATACGGCGACCACCAGCAACTCTTTCCTACACGACGCTCTTCCGATCTGAAATGTCTGGGATTTGATC</td>
</tr>
<tr>
<td>Tbx3_promoter_R</td>
<td>CAAGCAGAAGAAGCAAGGCTACGAGTCGATGTAGGTGTAAGATC</td>
</tr>
<tr>
<td>eA_F</td>
<td>AATGATACGGCGACCACCAGCAACTCTTTCCTACACGACGCTCTTCCGATCTGACTGGATGTAGGTGTAAGATC</td>
</tr>
<tr>
<td>eA_R</td>
<td>CAAGCAGAAGAAGCAAGGCTACGAGTCGATGTAGGTGTAAGATC</td>
</tr>
<tr>
<td>eB_F</td>
<td>AATGATACGGCGACCACCAGCAACTCTTTCCTACACGACGCTCTTCCGATCTGACTGAGACCACCACCCTGATC</td>
</tr>
<tr>
<td>eB_R</td>
<td>CAAGCAGAAGAAGCAAGGCTACGAGTCGATGTAGGTGTAAGATC</td>
</tr>
<tr>
<td>Tbx5_promoter_F</td>
<td>AATGATACGGCGACCACCAGCAACTCTTTCCTACACGACGCTCTTCCGATCTGATCCCTGTGGGATGATC</td>
</tr>
<tr>
<td>Tbx5_promoter_R</td>
<td>CAAGCAGAAGAAGCAAGGCTACGAGTCGATGTAGGTGTAAGATC</td>
</tr>
<tr>
<td>CTCF_F</td>
<td>AATGATACGGCGACCACCAGCAACTCTTTCCTACACGACGCTCTTCCGATCTGAAATGTCTGGGATGATC</td>
</tr>
<tr>
<td>CTCF_R</td>
<td>CAAGCAGAAGAAGCAAGGCTACGAGTCGATGTAGGTGTAAGATC</td>
</tr>
<tr>
<td>Rbm19_promoter_F</td>
<td>AATGATACGGCGACCACCAGCAACTCTTTCCTACACGACGCTCTTCCGATCTGAGGAGAAGTCTGGGATGATC</td>
</tr>
<tr>
<td>Rbm19_promoter_R</td>
<td>CAAGCAGAAGAAGCAAGGCTACGAGTCGATGTAGGTGTAAGATC</td>
</tr>
<tr>
<td>Med13l_promoter_F</td>
<td>AATGATACGGCGACCACCAGCAACTCTTTCCTACACGACGCTCTTCCGATCTGAGGAGAAGTCTGGGATGATC</td>
</tr>
<tr>
<td>Med13l_promoter_R</td>
<td>CAAGCAGAAGAAGCAAGGCTACGAGTCGATGTAGGTGTAAGATC</td>
</tr>
</tbody>
</table>
Online Table II. qPCR primer sequences

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tbx3_Promoter_F</td>
<td>CAGCACTCGACCTGTGAAAA</td>
</tr>
<tr>
<td>Tbx3_Promoter_R</td>
<td>ATTGGCTCTTTGACGCTTTTC</td>
</tr>
<tr>
<td>Tbx3_eA_F</td>
<td>CAGCTGAGCCCTTCAGGAT</td>
</tr>
<tr>
<td>Tbx3_eA_R</td>
<td>GGTGTCAAACATCCCTTCTGA</td>
</tr>
<tr>
<td>Tbx3_eB_F</td>
<td>TCCCCGAACCTGACCTTCTGA</td>
</tr>
<tr>
<td>Tbx3_eB_R</td>
<td>CGTGTGTCGCTTCTTTCTTG</td>
</tr>
<tr>
<td>negative control region 1'F</td>
<td>TCCAGACACACTGAAGCAC</td>
</tr>
<tr>
<td>negative control region 1'R</td>
<td>TCCCCGAGATCAAGTGCTAC</td>
</tr>
<tr>
<td>negative control region 2''F</td>
<td>CCAGACACACTGAAGCAC</td>
</tr>
<tr>
<td>negative control region 2'R</td>
<td>TGCCAGTCTCTGGAAATC</td>
</tr>
</tbody>
</table>

* (chr4: 147397477 – 147397580)
** (chr4: 147397478 – 147397598)
References


