

# Probing chromatin landscape reveals roles of endocardial TBX20 in septation

Cornelis J. Boogerd,<sup>1</sup> Ivy Aneas,<sup>2</sup> Noboru Sakabe,<sup>2</sup> Ralph J. Dirschinger,<sup>3</sup> Quen J. Cheng,<sup>3</sup> Bin Zhou,<sup>4</sup> Ju Chen,<sup>3</sup> Marcelo A. Nobrega,<sup>2</sup> and Sylvia M. Evans<sup>1,3,5</sup>

<sup>1</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, UCSD, La Jolla, California, USA. <sup>2</sup>Department of Human Genetics, University of Chicago, Chicago, Illinois, USA.

<sup>3</sup>Department of Medicine, UCSD, La Jolla, California, USA. <sup>4</sup>Departments of Genetics, Medicine and Pediatrics, Albert Einstein College of Medicine of Yeshiva University,

New York, New York, USA. <sup>5</sup>Department of Pharmacology, UCSD, La Jolla, California, USA.

Mutations in the T-box transcription factor TBX20 are associated with multiple forms of congenital heart defects, including cardiac septal abnormalities, but our understanding of the contributions of endocardial TBX20 to heart development remains incomplete. Here, we investigated how TBX20 interacts with endocardial gene networks to drive the mesenchymal and myocardial movements that are essential for outflow tract and atrioventricular septation. Selective ablation of *Tbx20* in murine endocardial lineages reduced the expression of extracellular matrix and cell migration genes that are critical for septation. Using the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq), we identified accessible chromatin within endocardial lineages and intersected these data with TBX20 ChIP-seq and chromatin loop maps to determine that TBX20 binds a conserved long-range enhancer to regulate versican (*Vcan*) expression. We also observed reduced *Vcan* expression in *Tbx20*-deficient mice, supporting a direct role for TBX20 in *Vcan* regulation. Further, we show that the *Vcan* enhancer drove reporter gene expression in endocardial lineages in a TBX20-binding site-dependent manner. This work illuminates gene networks that interact with TBX20 to orchestrate cardiac septation and provides insight into the chromatin landscape of endocardial lineages during septation.

## Introduction

Congenital heart disease (CHD) represents the most common form of birth defect (1), with a majority of CHDs involving anomalies in cardiac septation. Cardiac septation involves crosstalk between multiple cardiac cell lineages, but mechanisms by which this crosstalk occurs remain far from understood. Mutations in the T-box transcription factor (TF) *TBX20* are associated with a diverse array of CHDs, including dilated cardiomyopathy, double-outlet right ventricle (DORV), and defects in cardiac septation and valvulogenesis (2–6). The diversity of cardiac pathologies associated with *TBX20* mutations underscores the importance of *TBX20* for multiple aspects of heart development.

In mice, previous studies with a hypomorphic mutant of *Tbx20* demonstrated a requirement for *Tbx20* in outflow tract (OFT) remodeling and endocardial cushion formation (7). Mice that are globally null for *Tbx20* are embryonically lethal around E10 and show a severely hypomorphic heart that is poorly proliferative and lacks specialized chamber myocardium (7–10). Early lethality of the global mutants prevents the study of later roles of *Tbx20*.

To understand the mechanisms through which mutations in *TBX20* cause CHD, it is critical to elucidate the tissue-specific roles for *TBX20*. Recently, *Tbx20* was shown to be required in nuclear factor of activated T cells 1–*Cre* (*Nfatc1-Cre*) lineages for valve maturation

and elongation via regulation of Wnt signaling (11). Here, our analysis of *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants and *Nfatc1-Cre Tbx20<sup>fl/fl</sup>* mutants has revealed broader requirements for *TBX20* in endocardium for multiple aspects of cardiac septation, including normal formation of the OFT and the dorsal mesenchymal protrusion. To understand the molecular and cellular behaviors underlying OFT alignment and septation defects, we combined detailed histological analyses with the generation of several genome-wide data sets from whole heart or purified endocardial lineages, including microarray, *TBX20* ChIP-seq, and ATAC-seq. We intersected these genome-wide data sets with chromatin loop maps to identify long-range-acting, *TBX20*-targeted enhancers. Altogether, we believe our data provide new insights into the mechanisms by which mutations in *TBX20* result in cardiac septation defects, highlighting distinct roles for *TBX20* in endocardial lineages. Our work illustrates an integrative approach that links molecular mechanisms of TF-regulatory functions to whole-systems biology, using well-defined cells in relatively small numbers.

## Results

*Tbx20* is required in endothelial lineages for multiple aspects of cardiac septation. To better understand the roles of *TBX20* in endocardium and endocardium-derived cushion mesenchyme, *Tie2-Cre* was used to ablate *Tbx20*. *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants were never recovered after birth, indicating prenatal lethality of mutant embryos. Until E13.5, mutant embryos were recovered at expected Mendelian frequencies (Table 1). However, at E14.5 and later, no live mutants were recovered, indicating that *TBX20* is required in endothelial lineages before this stage for embryonic viability.

**Authorship note:** C.J. Boogerd, I. Aneas, and N. Sakabe contributed equally to this work.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Submitted:** November 2, 2015; **Accepted:** May 5, 2016.

**Reference information:** *J Clin Invest*. 2016;126(8):3023–3035. doi:10.1172/JCI85350.

**Table 1. Observed genotype distribution of live embryos recovered from matings between *Tie2-Cre Tbx20<sup>fl/null</sup>* males and *Tbx20<sup>fl/fl</sup>* females**

Stage	Without <i>Cre</i>		<i>Tie2-Cre</i>		$\chi^2$
	<i>Tbx20<sup>fl/fl</sup></i>	<i>Tbx20<sup>fl/null</sup></i>	<i>Tbx20<sup>fl/fl</sup></i>	<i>Tbx20<sup>fl/null</sup></i>	
E10.5	48 (48.75)	49 (48.75)	48 (48.75)	50 (48.75)	$P = 1.0$
E12.5	44 (52.5)	44 (52.5)	65 (52.5)	57 (52.5)	$P = 0.11$
E13.5	22 (25.5)	25 (25.5)	30 (25.5)	25 (25.5)	$P = 0.73$
E14.5	14 (10)	17 (10)	9 (10)	0 (10)	$P = 0.0009^A$
E15.5	10 (6.5)	7 (6.5)	9 (6.5)	0 (6.5)	$P = 0.024^A$

Expected number is in parentheses;  $\chi^2$  results were obtained;

<sup>A</sup> $P$  value of less than 0.05 was considered significant.

At E13.5, mutant embryos displayed overall normal size and morphology (Figure 1A). Histological analysis of mutant hearts at E13.5 revealed completely penetrant cardiac cushion abnormalities and septal defects (Figure 1B). Most notably, mutant hearts had OFT septation defects reminiscent of a DORV, marked by failure of the septal cushion to extend toward the interventricular septum (IVS). OFT and atrioventricular cushions were malformed. Mutant hearts had ventricular septal defects (VSDs) and ostium primum type atrial septal defects (ASDs) (Figure 1C), which are considered a type of atrioventricular septal defect (AVSD) (12). Additionally, the dorsal mesenchymal protrusion (DMP), a second heart field-derived structure that normally appears as a protrusion at the right pulmonary ridge (13), appeared to be lacking in the mutant hearts (Figure 1C). At E10.5 and E13.5, no defects in vascular patterning were observed (Supplemental Figure 1; supplemental material available online with this article; doi:10.1172/JCI85350DS1).

*Tbx20* is required for the proliferation of cushion endocardium. To address the question of how a lack of TBX20 might lead to malformed cushions, we analyzed proliferation in cardiac cushion endocardium and mesenchyme at E11.5 using the *Rosa26<sup>mTmGFP</sup>* lineage indicator. 5-Ethynyl-2'-deoxyuridine (EdU) incorporation rates in endocardium overlying cushions (*Tie2-Cre* lineage<sup>+</sup>, CD31<sup>+</sup>; Figure 2A) were decreased by 27% in OFT cushions. A more subtle decrease in atrioventricular cushion endocardium (14%) failed to reach statistical significance (Figure 2B,  $P < 0.08$ ). The proliferation rates of endocardium-derived mesenchymal cells were unchanged. Accordingly, the overall numbers of cushion mesenchymal cells in E11.5 mutant hearts were similar compared with those in controls (Supplemental Figure 2A). Similar rates of apoptosis were observed in mutant and control cushions (Supplemental Figure 2B).

*Tbx20* is required for the normal migratory behavior of endocardial lineages after EMT. Endocardial cells contribute to OFT and atrioventricular cushions by undergoing endothelial-to-mesenchymal transition (EMT) and invading the cardiac jelly (14). Mutant atrioventricular and OFT cushions were colonized by *Tie2-Cre Rosa<sup>tdTom</sup>* lineage-traced mesenchymal cells (Figure 3A), demonstrating that mutant endocardial cells had undergone EMT. In control OFT cushions, endocardium-derived mesenchymal cells were loosely arranged and intermingled with mesenchymal cells of nonendothelial origin, such as neural crest cells (15). In

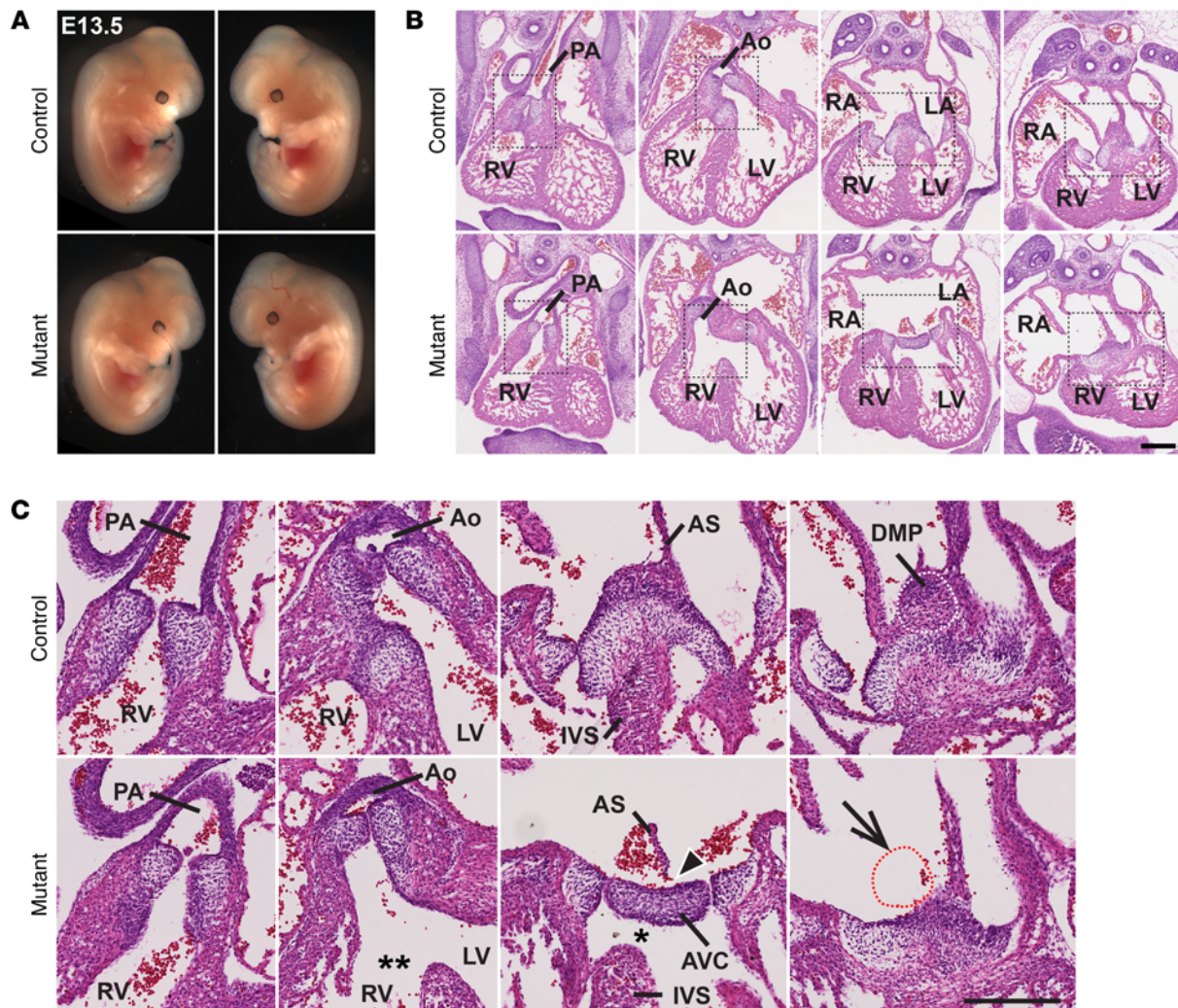
contrast, mutant cells appeared as dense clusters in the proximal cushions that did not mix with nonendocardium-derived mesenchymal cells and failed to disperse throughout the cushions (Figure 3B). Notably, we also observed these migratory defects by using a more endocardium-restricted *Cre* line, *Nfatc1-Cre* (16), to ablate *Tbx20* (Supplemental Figure 3, A and B).

To further characterize *Tbx20*-mutant endocardial cells, we analyzed their ability to undergo EMT in an ex vivo cushion explant assay (17). Mutant OFT cushion endocardial cells underwent EMT to the same extent as did their WT counterparts, as evidenced by abundant numbers of *Tie2-Cre* lineage-traced cells that migrated away from the explant over the surface of the gel (Figure 3C and Supplemental Figure 3C). In addition, the distances migrated by mutant endocardial cells atop the matrix were comparable to those of control cells (Supplemental Figure 3D). However, there was a striking reduction in the number of mutant endocardium-derived cells able to invade the gel relative to that seen in WT cells (Figure 3, C and D).

*Tbx20* in endocardial lineages is required for myocardialization of the OFT. After invasion by endocardial cells and neural crest cells, OFT cushions expand, fuse, and form a mesenchymal outlet septum. Subsequently, cardiomyocytes that flank cushion mesenchyme protrude into cushion mesenchyme. This process, called myocardialization, involves polarized cell movements of OFT cardiomyocytes (18, 19). In control embryos at E13.5, sarcomeric  $\alpha$ -actinin-expressing (*Actn2*-expressing) cardiomyocytes invaded the OFT septum, and cells protruding from the left and right sides met at the base of the proximal OFT septum (Figure 3B). In mutants, however, cardiomyocyte invasion was impaired, with a significant distance separating the protruding cardiomyocytes (Figure 3B). Notably, aberrantly clustered endocardium-derived cells resided in this region in mutants. These observations suggested that mutant endocardial cells were not able to provide the cues to myocytes required for cardiomyocyte invasion.

*Tbx20* in endothelial lineages is required for development of the DMP. To better understand the spatiotemporal requirements for *Tbx20* in cardiac septation, we used a more endocardium-restricted *Cre* line, *Nfatc1-Cre*, to ablate *Tbx20*. Notably, *Nfatc1-Cre Tbx20<sup>fl/fl</sup>* mutants also displayed OFT septation defects, although to a lesser extent than did *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants (Supplemental Figure 3A). However, in contrast to *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants, *Nfatc1-Cre Tbx20<sup>fl/fl</sup>* mutants did not show AVSDs at E13.5, and contribution of the DMP to the atrioventricular septal complex (AVD) was evident (Supplemental Figure 3B). Next, we performed genetic lineage tracing to compare *Nfatc1-Cre* and *Tie2-Cre* expression domains and found that at E9.5 and E10.5, *Nfatc1-Cre* did not label pulmonary venous endothelium and did not efficiently label the endocardium lining the sinus venosus and common atrium, whereas these cell populations were robustly labeled by *Tie2-Cre* (Supplemental Figure 4). Notably, these regions did express *Tbx20* (Supplemental Figure 5), suggesting that *Tbx20* expression in endothelium adjacent to the DMP, including endothelium of the pulmonary vein, sinus venosus, or atrium, is required for DMP development.

*Tbx20* regulates extracellular matrix and cell guidance genes in endocardial lineages. To examine how ablation of *Tbx20* affected gene expression, we performed global gene expression analysis on FACS-purified endocardial lineages from E12.5 hearts (Sup-



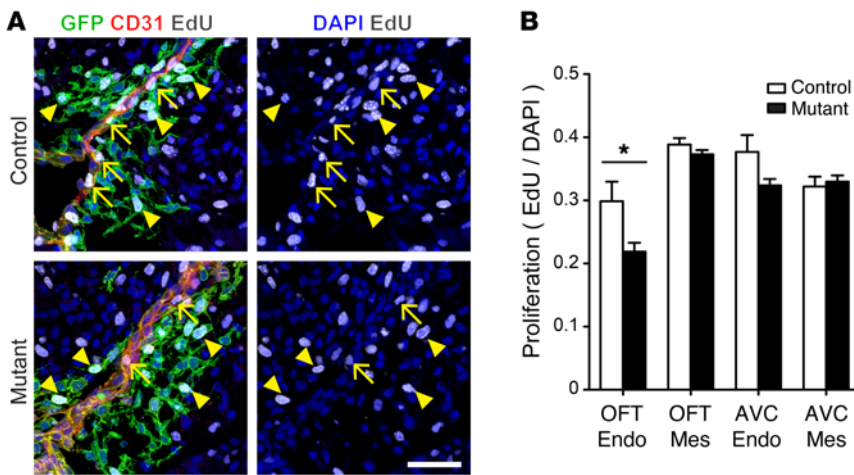
**Figure 1.** *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants show DORV and ASDs with reduced contribution from the DMP. (A) E13.5 control (*Tie2-Cre Tbx20<sup>WT/fl</sup>*) and mutant (*Tie2-Cre Tbx20<sup>fl/fl</sup>*) embryos. (B and C) Cardiac defects in E13.5 mutant hearts. The aorta (Ao) and pulmonary artery (PA) originate from the right ventricle (RV) in mutant embryos; the OFT septal cushion does not connect to the IVS in mutant hearts (double asterisks); VSD (single asterisk); AVSD (arrowhead); absence of a DMP (outlined by dashed lines, arrow). The defects shown are representative of 10 mutant hearts analyzed. RA, right atrium; LV, left ventricle; AS, primary atrial septum. Scale bars: 250  $\mu$ m (B and C).

plemental Figure 6). Pathway analysis of differentially expressed genes revealed a striking overrepresentation of extracellular matrix (ECM) and axon guidance genes among the downregulated genes (Figure 4A and Supplemental Table 1). Genes within the top category of downregulated ECM genes included versican (*Vcan*), transforming growth factor  $\beta$  2 (*Tgfb2*), and several genes encoding collagen proteins. We observed no changes in phosphorylated SMAD3 levels in *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutant cushions compared with levels in controls at E10.5 or E12.5 (not shown), indicating that decreased *Tgfb2* expression did not translate into a functional deficit in canonical TGF- $\beta$  signaling. Downregulation of *Vcan* was of particular interest, as *Vcan*-null mice die around E10.5 and have a complete lack of cushion EMT and a truncated OFT and right ventricle, findings that were recently shown to reflect a requirement for *Vcan* in *Tie2-Cre* lineages (17, 20, 21). Additionally, decreased VCAN levels lead to OFT, atrioventricular cushion, and DMP defects that are strikingly similar to the phenotypes we observed in *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants (22, 23).

Genes grouped by the axon guidance term may be of significance to the observed defects in endocardial cell migration as well as myocardialization defects. Axon guidance genes significantly ( $P < 0.05$ ) downregulated in mutants included roundabout 1 (*Robo1*) and slit homolog 2 (*Slit2*), components of the Slit/roundabout signaling pathway that regulate cardiac septation and valve development (24). *Myh10*, encoding a nonmuscle myosin required for cell motility (25), was also downregulated in *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants. Mutations in *Myh10* also cause cardiac septation defects (26).

Several additional genes, including the TFs *Twist1*, *Msx1*, and *Runx1t1*, that are associated with cushion development and cardiac septation were also downregulated in *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants (27–29). Expression of several genes of potential biological relevance was validated by quantitative real-time PCR (qRT-PCR) (Figure 4B).

*Genomic targets of Tbx20 in mid-gestation heart identified by ChIP-seq analysis.* TBX20 is a TF that binds DNA and can activate or repress the expression of target genes (30). To identify



**Figure 2. Proliferation in cushions of *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants and controls.** (A) EdU incorporation in E11.5 endocardial cells (arrows; CD31<sup>+</sup> and *Tie2-Cre Rosa<sup>mGFP+</sup>*) and endocardium-derived cells (arrowheads; CD31<sup>-</sup> and *Tie2-Cre Rosa<sup>mGFP+</sup>*). (B) EdU incorporation was decreased in mutant OFT cushion endocardial cells. \* $P < 0.05$ , by unpaired, 2-tailed  $t$  test; mean  $\pm$  SEM;  $n = 3$ . Endo, endocardium; Mes, mesenchyme. Scale bar: 50  $\mu$ m.

genes directly regulated by TBX20, we mapped TBX20-binding sites in E11.5 mouse heart by ChIP-sequencing (ChIP-seq), using a *Tbx20-GFP* transgene (31) that drives expression of a TBX20-GFP fusion protein under the control of the *Tbx20*-regulatory sequences in endocardium, myocardium, and cushion mesenchyme and that closely resembles the endogenous *Tbx20* expression pattern seen in embryonic heart (Supplemental Figure 5, C–E). Chromatin extracts prepared from *Tbx20-GFP*-transgenic hearts were subjected to ChIP to enrich for TBX20-bound chromatin and subjected to massive parallel sequencing. ChIP-seq analysis resulted in 2,988 peaks of TBX20 binding at a 0.1% FDR. Approximately 18% of these peaks were in promoter regions of RefSeq genes, indicating that, although TBX20 directly binds the promoter regions of several genes, the majority of its regulation relies on binding to longer-range DNA elements (Figure 5A).

TF motif discovery (32) within TBX20-bound regions revealed that the majority of regions (56%) had a TBX20 motif (Figure 5B and ref. 31). Canonical T-box motifs (TBX5, Eomes; AAGTGTC) were also enriched, though to a lesser extent (Supplemental Table 2). These data indicated that ChIP-seq fragments were highly enriched for genomic regions directly bound by TBX20. In addition, we found overrepresentation of motifs for GATA family TFs MAFA, a member of the bZIP family of TFs that includes AP1 factors MEF2C, homeodomain TFs NKX2-1 and NKX2-5, and TEAD TF-binding motifs (Figure 5B and Supplemental Table 2). NKX2-5, GATA, MEF2, and TEAD TFs have been previously identified as cofactors of TBX20 in cell lines and adult heart, indicating a high level of functional conservation (3, 30, 31, 33, 34).

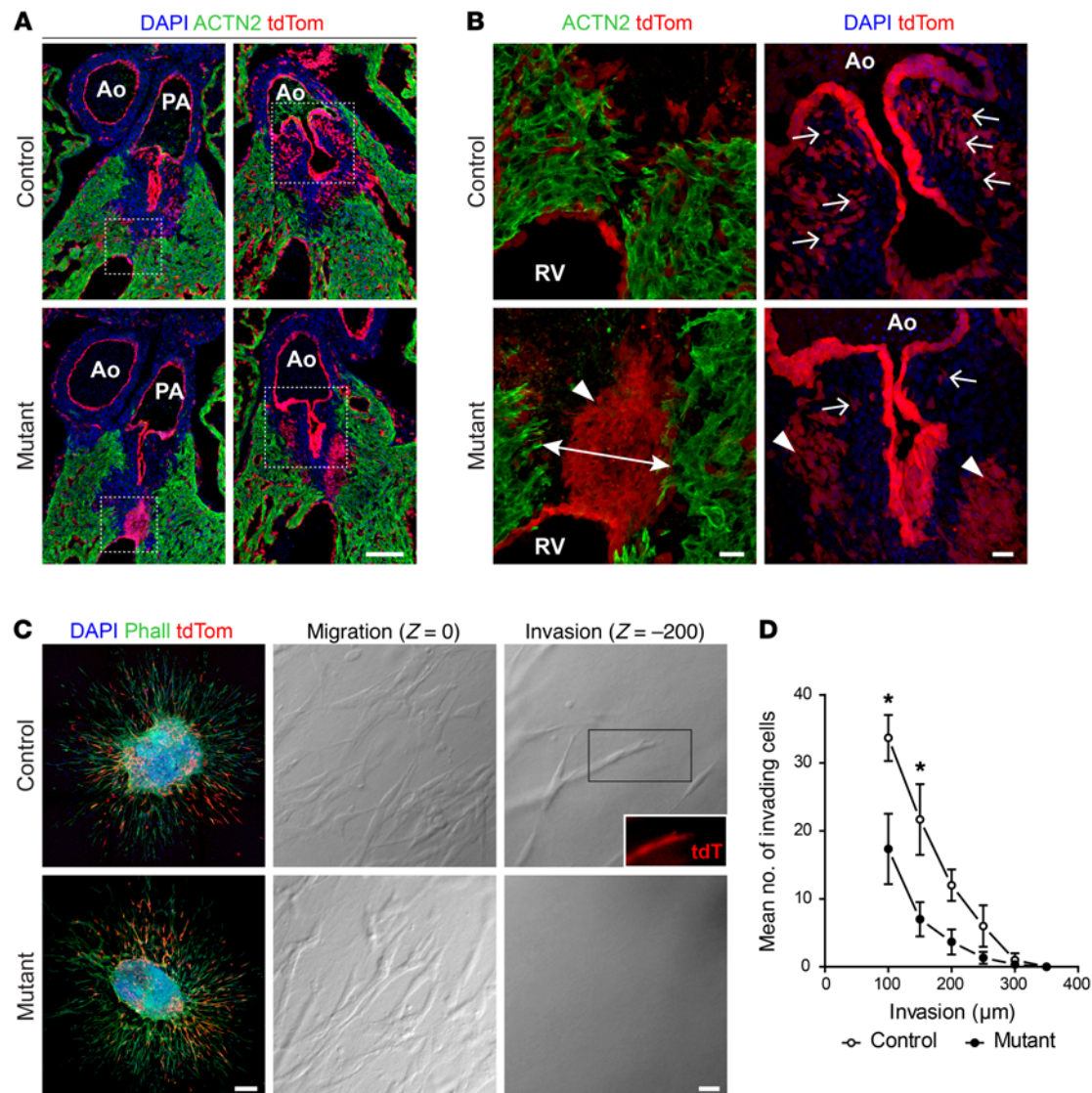
Genes proximal to TBX20-binding regions were subsequently analyzed for functional annotation using gene ontology (GO) enrichment analysis (35). In embryonic hearts, TBX20 preferentially bound near genes involved in cardiac chamber development, cardiac ventricle development, and cardiac septum morphogenesis (Figure 5C). It is likely that these terms reflect roles of TBX20 in cardiomyocytes, which make up approximately two-thirds of cells within the heart at this stage (data not shown). Genes involved in heart valve formation, EMT, OFT morphogenesis, mesenchyme development, and endocardial cushion formation were also markedly enriched and might reflect roles of TBX20 in endocardial lineages.

*Open chromatin in endocardial lineages identified by ATAC-seq.* We had used whole-heart chromatin extracts to map TBX20-binding sites, as it would have been technically challenging to isolate sufficient numbers of the endocardial lineage cells required for ChIP-seq experiments. To define subsets of TBX20-binding sites accessible for transcriptional regulation in endocardial lineages, we applied ATAC-seq (36). This method for mapping chromatin accessibility genome wide utilizes transposase, which will preferentially integrate sequencing adapters in regions of open chromatin such as active promoter regions, enhancers, and insulators and can be performed on relatively small amounts of cells. We performed ATAC-seq on FACS-sorted endocardial lineages from E12.5 hearts. Samples were subjected to massive parallel sequencing, and overlapping peaks from replicate samples were merged to identify high-confidence regions of open chromatin. Peaks were highly enriched at promoter regions (Figure 6A).

A first validation of ATAC-seq data revealed a high read count around promoter regions of genes, including *Snai1*, ET26 transformation-specific (*Ets*), and *Msx1*, expressed in endocardial lineage cells (Figure 6C and refs. 37–39). In contrast, troponin T (*Tnnt2*) and myosin heavy chain 7 (*Myh7*), 2 strictly myocardial genes, did not display significant read counts near their promoters. Several cushion-specific enhancers were also marked by open chromatin, confirming that ATAC-seq marks active promoters and enhancers (Figure 6C and refs. 40–42).

The most highly enriched motif in ATAC-seq regions of open chromatin was a CTCF motif (Figure 6B). CTCF is a TF that binds at chromatin domain boundaries, at enhancers and gene promoters, and inside gene bodies, the majority of which will be marked by open chromatin (43, 44). We also found enrichment for TF motifs of the ETS family, a large family of TFs marked by a highly conserved DNA-binding domain (45). ETS TFs are central regulators of endothelial gene expression, and at least 19 ETS factors are expressed in human endothelial cells (46, 47). Other highly overrepresented motifs are GATA and JUN-AP1 motifs (Supplemental Table 2).

*Identification of TBX20-binding sites within open chromatin of endocardial lineages.* To identify TBX20-binding sites that could be relevant to endocardial lineage development, we selected TBX20 ChIP-seq peaks within regions of open chromatin in endocardial



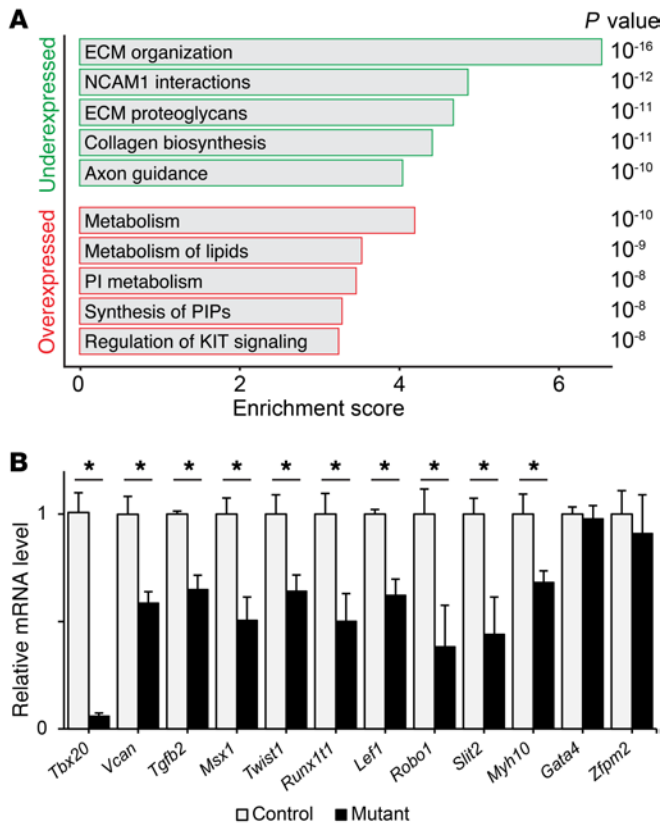
**Figure 3. TBX20 is required for endocardium-derived cell migration.** (A) *Tie2-Cre* lineage-traced mesenchymal cells (tdTom<sup>+</sup>) failed to migrate properly in mutants. (B) Mutant mesenchymal cells clustered (arrowheads) and failed to colonize the developing valves (arrows). Muscularization of OFT septum by cardiomyocytes (ACTN2<sup>+</sup>) was delayed (double arrow). (C) Ex vivo cushion explant showing EMT by mutant endocardium-derived cells (tdTom<sup>+</sup>, phalloidin<sup>+</sup>), cell migration over the surface of the collagen gel (middle; migration), and cell invasion into the matrix (right; invasion). Shown are representative images of 6 (*Tie2-Cre Tbx20*<sup>fl/null</sup> mutant) and 5 (*Tie2-Cre* control) explants. (D) Decreased matrix invasion in mutant endocardium-derived cells. \**P* < 0.05, by 2-way ANOVA with Bonferroni's post test; mean ± SEM; *n* = 3. Scale bars: 100 μm (A), 20 μm (B), 200 μm (C, left panel), and 20 μm (C, middle and left panels and inset). tdTom, tdTomato; Phall, phalloidin.

lineages. One-third of TBX20-binding sites in mid-gestation heart overlapped with regions of open chromatin, suggesting that these regions are potential active enhancer and promoter elements regulated by TBX20 in endocardial lineages (Figure 6D).

To identify potential cofactors that might interact with TBX20 in the regulation of gene expression in endocardial lineages, we analyzed TF motifs for this set of overlapping sequences. As anticipated, the most highly enriched motif was a TBX20 motif, present in 45% of sequences (Figure 6E). Additionally, canonical T-box motifs (TBX5, Eomes) were overrepresented at lower ranks. Interestingly, several factors known to interact with TBX20 in adult myocardium were also enriched in endocardial lineages, including GATA, TEAD, and MEF2 families. TBX20 may also rely on these cofactors for regulation of endocardial and mesenchymal gene expression. These

enhancers might also be shared between endocardium and myocardium. Additionally, we found significant overrepresentation of ETS motifs and AP1 motifs. These have not previously been identified as cofactors of TBX20 and were specifically enriched in the subset of TBX20 peaks that overlapped with endocardial lineage open chromatin (Figure 6E and Supplemental Table 2).

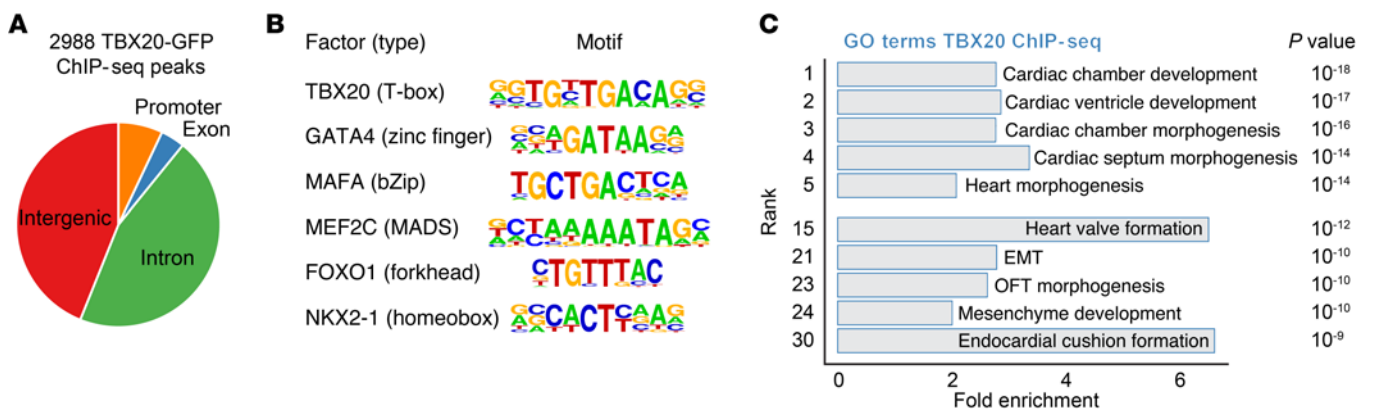
To identify putative direct targets of TBX20 in endocardial lineages, we selected differentially expressed genes that had TBX20-binding sites in a nearby region of open chromatin. We found 44 putative direct endocardial targets downregulated and 72 putative direct endocardial targets overexpressed in *Tie2-Cre Tbx20*<sup>fl/null</sup> mutants (Figure 6F). Among the downregulated putative direct target genes that might impact observed mutant phenotypes were *Robo1* and *Tbx20* itself (Supplemental Table 3).



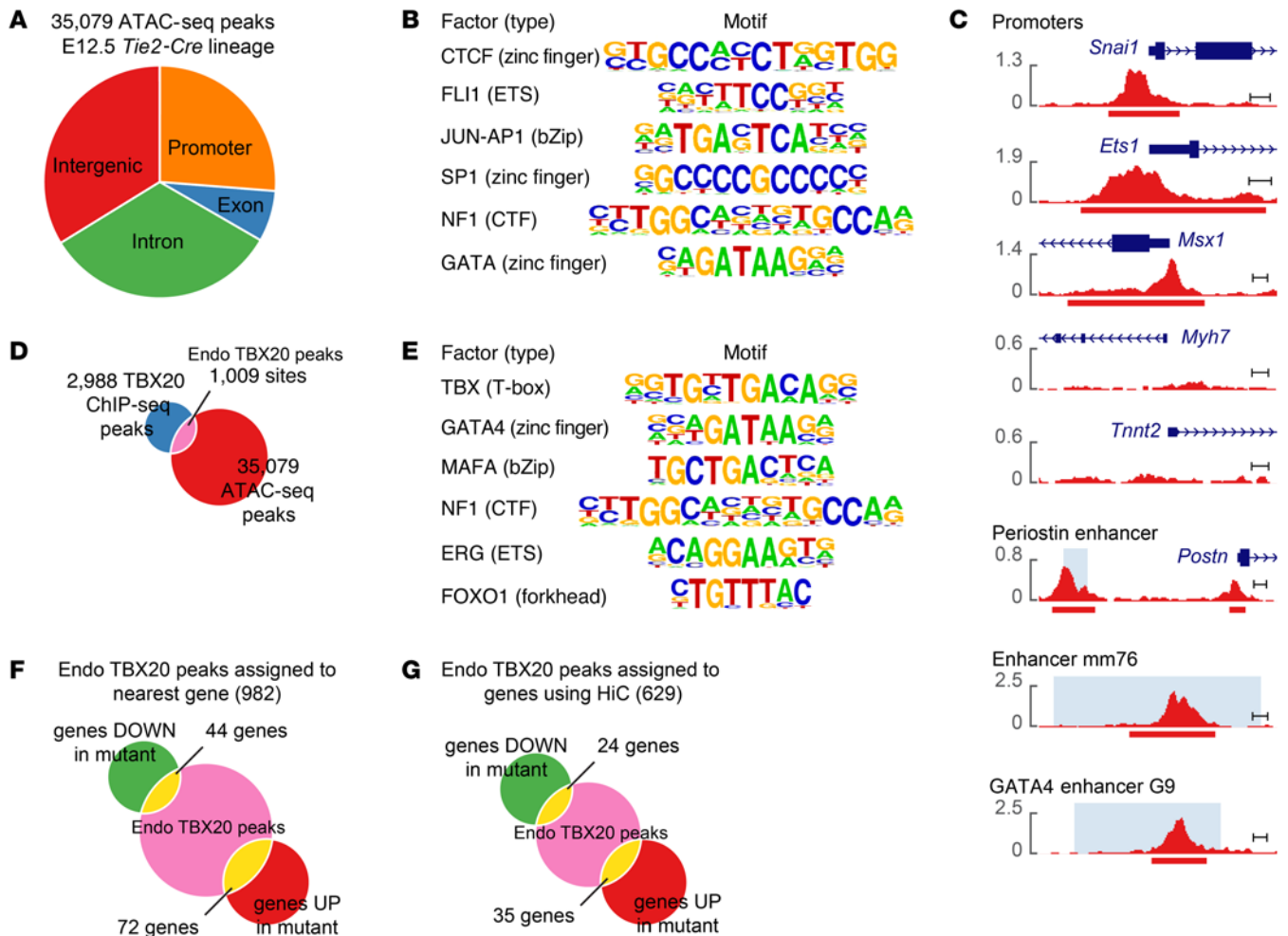
**Figure 4. Expression analysis of FACS-sorted *Tie2-Cre* lineages in E12.5 *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutant versus control hearts. (A) Pathway analysis of differentially expressed genes (top-5 terms). (B) qRT-PCR validation of relative gene expression for selected genes. \**P* < 0.05, by unpaired, 2-tailed *t* test; mean ± SEM; *n* = 3. KIT, proto-oncogene receptor tyrosine kinase; NCAM1, neural cell adhesion molecule 1; PI, phosphatidylinositol; PIPs, phosphatidylinositol phosphates.**

We next investigated whether the endocardial genetic programs regulated by TBX20 that we uncovered in mouse hearts are also conserved in humans. 3D-ordered chromatin structures connect long-range enhancers to cognate gene promoters. We postulated that chromatin conformation could be used as a guide in assigning TBX20 ChIP-seq peaks within long-range enhancers to their target genes. We used genome-wide interaction maps generated by high-resolution chromatin conformation capture (HiC) of a human fibroblast cell line that has many characteristics of endocardium-derived mesenchyme (51). By projection of chromatin interactions in the human genome onto orthologous conserved regions of the mouse genome and intersection with our TBX20 ChIP-seq peaks within open chromatin of endocardial lineages, as determined by ATAC-seq, we identified 906 transcription start site-bound (TSS-bound) DNA regions with TBX20 peaks and assigned these peaks to 629 orthologous mouse genes. Notably, TBX20 ChIP-seq peaks and ATAC-seq peaks overlap HiC interactions in excess of approximately 2-fold when compared with heterochromatin regions in human fetal heart identified by the Epigenome Roadmap (52). This observation suggests that the human fibroblast HiC data are useful for uncovering functional regions in mouse endocardium (Supplemental Figure 9D). Interestingly, 25% of these peaks were assigned to the most proximal TSS, whereas others were assigned to more distant genes, suggesting that linear proximity is a poor metric for the functional assignment of target genes for distal enhancers (50). With this approach, we found that 27 TBX20 peaks were associated with downregulated genes and 40 TBX20 peaks with upregulated genes (Figure 6G). Among downregulated putative direct targets of TBX20 with potential long-range-acting enhancers in endocardial lineages were *Vcan*, *Myh10*, and *Runx1t1* (Supplemental Table 4). These data suggest that at least some of the most relevant downstream

*Identification of long-range enhancers using publicly available chromatin loop maps.* Our understanding of how alterations in the expression of putative direct targets of TBX20 in endocardial lineages might contribute to the observed phenotypes in *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants was limited by the fact that their function has not often been explored in endocardium. Additionally, these direct targets were defined by assigning TBX20 ChIP-seq peaks to the nearest genes. While linear proximity to regulatory sites is often used to predict target genes, many examples exist in which enhancers ignore the nearest genes and act on genes further away (48–50).



**Figure 5. TBX20-GFP ChIP-seq in E11.5 mouse hearts. (A) Genomic distribution of TBX20-GFP ChIP-seq peaks. (B) Select TF motifs enriched in TBX20-GFP peaks (the most enriched factor for each family is shown; see the complete list in Supplemental Table 2). (C) GO analysis of the predicted TBX20 target genes.**



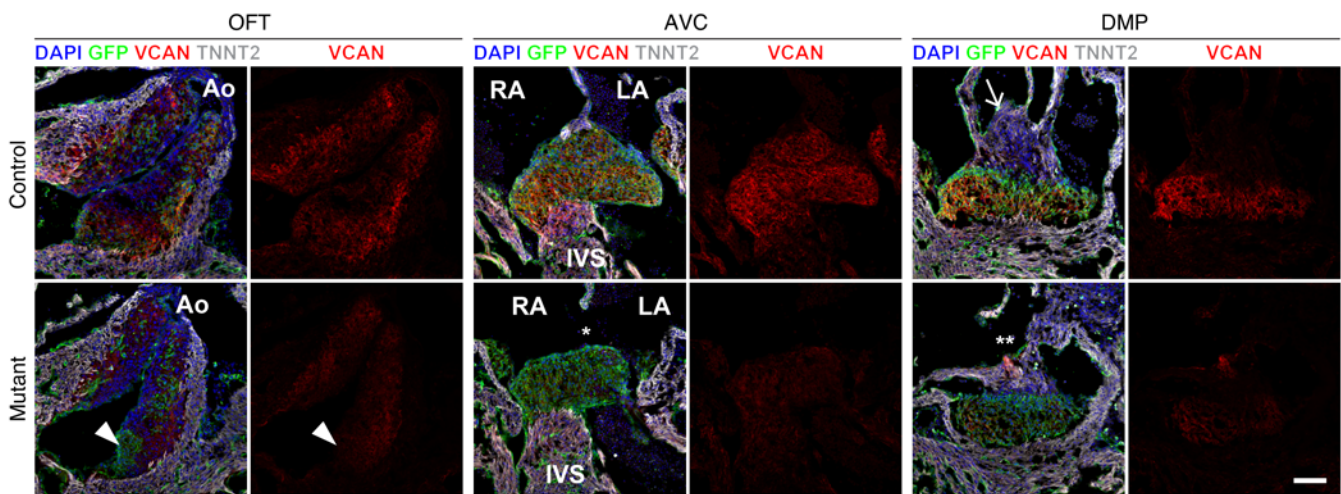
**Figure 6. E12.5 endocardial lineage ATAC-seq and intersection with TBX20 ChIP-seq.** (A) Genomic distribution of ATAC-seq peaks. (B) Select TF motifs overrepresented in ATAC-seq peaks (the most enriched factor for each family is shown; see the full list in Supplemental Table 3). (C) ATAC-seq read counts (red) and peaks (red bars) near known promoters and enhancers (light blue regions). (D) TBX20 ChIP-seq peaks in endocardial lineage ATAC-seq regions identified putative endocardial lineage-relevant regulatory elements bound by TBX20. (E) Select TF motifs enriched in TBX20 ChIP-seq peaks within endocardial open chromatin. (F and G) Intersection of differentially expressed genes and endocardial lineage-relevant regulatory elements bound by TBX20 using (F) the nearest-gene approach or (G) HiC-guided peak assignment. UP, upregulated; DOWN, downregulated; FLI1, Friend leukemia integration 1 transcription factor; SP1, transcription factor Sp1; NF1, neurofibromin; ERG, transcriptional regulator ERG.

targets of TBX20 in mouse endocardium are regulated through enhancers conserved between mouse and humans, giving mechanistic insights into the molecular basis for CHD in patients harboring *TBX20* mutations.

*Versican expression is diminished in mutant cushions.* Because mutants with decreased VCAN levels have phenotypes strikingly similar to those of *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants (22, 23), we further investigated *Vcan* downregulation as a potential direct target that contributes to the observed *Tie2-Cre Tbx20<sup>fl/null</sup>* phenotypes. qRT-PCR on RNA from purified endothelial lineages with primers specific for transcript variants of *Vcan* indicated that expression of all 4 variants examined was decreased in mutants (Supplemental Figure 7A). Furthermore, we found that downregulation of *Vcan* transcripts had occurred by E10.5 in mutant endocardial cells (Supplemental Figure 7B). To investigate whether decreased mRNA levels resulted in decreased protein levels, we performed anti-VCAN immunostaining. In control hearts, VCAN was highly expressed in OFT and

atrioventricular cushions, suggesting its secretion by endocardial lineages and, potentially, by mesenchymal cells of nonendothelial origin (Figure 7). In contrast, in *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants, VCAN was markedly reduced in OFT and atrioventricular cushions. Mutant cushions expressed normal levels of the mesenchymal cell marker platelet-derived growth factor receptor  $\alpha$  (PDGFR $\alpha$ ) and showed normal levels of ECM proteoglycans (Supplemental Figure 7, D and E). Moreover, VCAN levels surrounding the main bronchi were similar in controls and mutants (Supplemental Figure 7C), confirming that the reduction of VCAN levels in cushions was specific to a loss of TBX20 in endocardial lineages.

*A Vcan enhancer bound by TBX20 in endocardial lineages drives expression in endocardial lineages in vivo.* Because *Vcan* was significantly downregulated in *Tbx20*-mutant endocardial lineages, and since this reduction may contribute to cardiac defects in our mutant, we further investigated the direct regulation of *Vcan* by TBX20. Remarkably, the HiC data suggested



**Figure 7. Expression of VCAN in E12.5 hearts is reduced in mutant cushions.** GFP marks *Tie2-Cre*-labeled lineages. TNNT2 marks cardiomyocytes. Clustered mutant cells in the OFT (arrowhead); AVSD (single asterisk); and absence of a DMP (double asterisks; arrow in WT) in mutant heart. Scale bar: 100  $\mu$ m. LA, left atrium.

that in human cells, VCAN participates in long-range genomic interactions with sequences spread over 1.1 million bp around VCAN TSSs and spanning at least 5 RefSeq genes. Within the mouse *Vcan*-regulatory landscape, we identified 4 TBX20-binding sites, 2 of which mapped to regions of open chromatin in endocardial lineages (Figure 8A and Supplemental Figure 8), the human counterparts of which were linked to VCAN TSSs by HiC (51). By inspection of the Human Epigenome Roadmap data, we noticed that the orthologous human regions corresponding to these candidate enhancers also harbor epigenetic marks that are hallmarks of enhancers in human heart (Supplemental Figure 9). Remarkably, these putative cardiac enhancers are in the vicinity of *ATP6AP1L*, a gene separated from VCAN by 4 genes and a gene desert, highlighting the power of overlaying multiple genomic databases as a tool kit to uncover possible functional interactions in human and mouse genomes. The functionality of these candidate cardiac enhancers was directly tested in vivo using a mouse transgenic reporter assay. Enhancer 1 was expressed in multiple embryonic regions, with some variation between embryos (Figure 8, B and C), while enhancer 2 was highly cardiac specific, with only minor expression outside the heart in some embryos (Supplemental Figure 8, C and D, and data not shown). Section analysis revealed that enhancer 1 consistently drove reporter gene expression in cushion endocardium and/or mesenchyme (Figure 8D). In contrast, enhancer 2 drove robust expression in cardiomyocytes, most notably in IVS, ventricles, atrial septum, and venous valves (Supplemental Figure 8, C and D). Enhancer 1 contained 2 perfect consensus TBX20-binding motifs. Mutation of both TBX20 sites almost completely abolished enhancer activity in cardiac cushions in transgenic embryos in vivo. (Figure 8E). Taken together, these results link TBX20 binding to a distant enhancer that regulates *Vcan* in endocardial lineages, uncovering a mechanism by which *Vcan* expression is TBX20 dependent and providing a framework for how downregulation of *Vcan* in TBX20 endocardial mutants may contribute to congenital heart defects.

## Discussion

Using a conditional mouse model and genome-wide analyses, we investigated the functional relevance of TBX20 in endocardial lineages for cardiac development. Analysis of *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants revealed previously unrecognized requirements for TBX20 in endocardial lineages for multiple aspects of cardiac septation. Although we did not observe defects in vascular patterning, we cannot formally rule out the possibility that vascular defects may have contributed to the observed phenotypes or embryonic lethality.

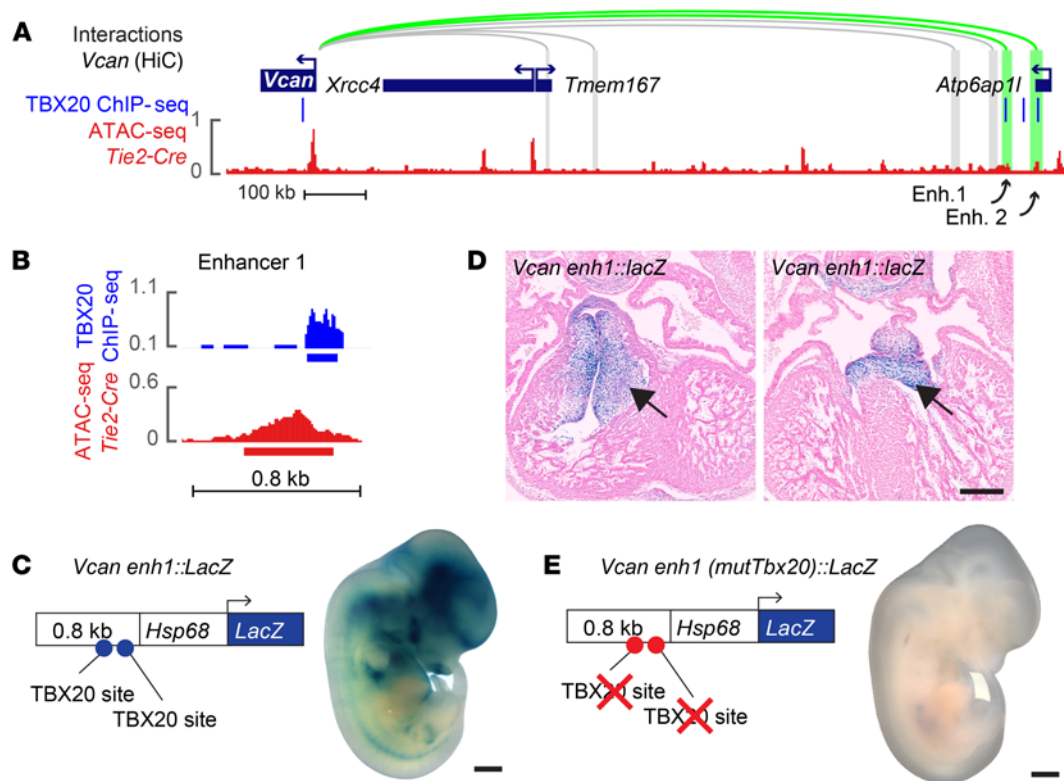
*Tbx20 regulates the endocardial proliferation and migratory behavior of endocardium-derived mesenchymal cells.* We observed perturbations in several aspects of cushion development in mutant embryos. Cushion malformations and decreased proliferation of endocardium overlaying the cushions were consistent with the results obtained with *Nfatc1-Cre Tbx20<sup>fl/fl</sup>* mutants (11), in which decreased endocardial proliferation was attributed to decreased  $\beta$ -catenin signaling via downregulation of lymphoid enhancer-binding factor 1 (*Lef1*), a phenomenon we also observed in *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants.

We also found that TBX20 was required for normal dispersion of endocardium-derived cushion cells. Ex vivo endocardial explant assays demonstrated that TBX20 is not required for EMT, but that mesenchymal cells require TBX20 for subsequent matrix invasion. In vivo, mutant endocardium-derived mesenchymal cells were abnormally clustered and failed to disperse and populate the forming valves in the OFT.

*Tbx20 in endocardium regulates cardiomyocyte migration.* We found that loss of TBX20 in endocardial lineages resulted in decreased myocardialization of the OFT, indicating a disturbed crosstalk between endocardial lineages and cardiomyocytes. Myocardialization involves polarized cell movements of OFT cardiomyocytes and is required for normal OFT alignment and septation (18, 19). Decreased myocardialization in *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants probably contributes to the observed DORV phenotype.

We describe another endocardial requirement for TBX20 in DMP formation. *Tie2-Cre Tbx20<sup>fl/fl</sup>* mutants displayed a striking absence of a DMP, which is a second heart field-derived struc-





**Figure 8. A distant *Vcan* enhancer drives expression in endocardium and cushion in E12.5 hearts.** (A) *Vcan* locus, with projected chromatin interactions from IMR90 fibroblasts (HiC; gray and green [overlap with ATAC-seq] lines and bars); RefSeq genes (dark blue); TBX20-GFP ChIP-seq peaks (blue); and ATAC-seq read counts (red). (B) Alignment of TBX20 ChIP-seq and ATAC-seq peaks in *Vcan* enhancer 1. (C) *Vcan* enh1::LacZ transgenic reporter embryos showed a strong pattern for X-gal staining. (D) Section analysis revealed that 6 of 8 *Vcan* enh1::LacZ had LacZ expression in the OFT and atrioventricular cushions (arrows). (E) Mutation of TBX20 sites in *Vcan* enh1 largely disrupted the pattern of X-gal staining, with 1 of 8 embryos showing expression in cushions. Scale bars: 1 mm (C and E), 200 μm (D).

ture required for proper atrioventricular septation (13, 53). Notably, this defect was not observed in our analysis of *Nfatc1-Cre Tbx20<sup>fl/fl</sup>* mutants. Unlike the atrioventricular cushions and the mesenchymal cap of the atrial septum, the DMP is not derived from *Tie2-Cre* lineages (13). Therefore, defects in DMP development in *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants suggest a crosstalk between *Tbx20*-expressing cells of endocardial lineage and myocardial progenitors of the DMP. Given our comparison of *Nfatc1-Cre* and *Tie2-Cre* lineages and their proximity to the DMP, we suggest that candidate regions for such crosstalk are the atrial endocardium, sinus venosus endothelium, and pulmonary vein endothelium (Supplemental Figure 10). To the best of our knowledge, our study is the first to demonstrate that endothelial-to-DMP signaling is required for atrioventricular septation.

*Tbx20* directly regulates the expression of ECM and migration genes. Consistent with the kinds of abnormal cellular behaviors we observed in *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants, microarray analysis of purified endocardial lineages showed downregulation of ECM and axon guidance genes. The intersection of this data set with TBX20 ChIP-seq data from whole embryonic heart revealed that, in mutants, several direct targets of TBX20 that have essential roles in cardiac septation were downregulated, including *Robo1*, *Myh10*, and *Vcan*. *Robo1* mutants had VSDs similar to those seen in *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants, and *Robo1 Robo2* compound mutants

had defects in valve development, including immature atrioventricular valves and bicuspid aortic valves (24). *Myh10*-null mice and mice that carry a point mutation in *Myh10* have defects strikingly similar to those seen in *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants, including DORV, decreased OFT myocardialization, and defects in the fusion of atrioventricular cushions (25, 26). While the tissue-specific roles of these factors remain to be addressed, we postulate that their downregulation in *Tie2-Cre Tbx20<sup>fl/null</sup>* mutants probably contributes to the observed phenotypes.

*Cofactors of TBX20 in endocardial lineages.* Our analysis of genomic regions that were available for TF interaction through ATAC-seq in endocardial lineages and bound by TBX20 in embryonic heart revealed overrepresentation of specific TF motifs. Several of these TFs, including NKX2-5, GATA4, MEF2, and TEAD, have been previously associated with TBX20 in adult cardiomyocytes and cardiac fibroblasts (31, 34). Interestingly, motifs highly enriched near TBX20-binding sites within endocardial lineages included those recognized by the AP1 family, which contains JUN, FOS, ATF, and MAF proteins. AP1 proteins regulate a variety of processes including cell proliferation and survival, differentiation, growth, migration, and transformation. Intriguingly, conditional deletion of the AP1 family member *Jun* using *Tie2-Cre* leads to DORV, VSDs, and valve defects (54), suggesting that perhaps JUN cooperates with TBX20 in endocardial lineages to effect

septation. Both ETS and FOX factor motifs were also overrepresented in our ATAC-seq peaks, in keeping with the important role of ETS-FOX factors in endothelium-specific gene expression (55). A recent analysis of NKX2-5 mutants associated with CHD found that ETS factors are embedded in the cardiac gene-regulatory network (56). Moreover, the bioinformatics analysis of multiple DNA adenine methyltransferase identification (DAM-ID) and ChIP-seq experiments predicts a functional interaction between ETS factors and T-box factors in the regulation of cardiac target gene expression (56). Our finding that ETS motifs were highly enriched near endocardial TBX20-binding sites provides additional evidence for such interactions and further underlines the importance of TF interactions in cardiac development.

Overrepresentation of GATA TF motifs is of particular interest, as cardiac septation is especially sensitive to the dosage of GATA4, GATA5, and GATA6. Ablation of *Gata4* from endocardium is embryonically lethal before E12.5 due to EMT defects in cushions (57). Additionally, a *Gata4* hypomorphic allele is associated with DORV and AVSDs (58), and, interestingly, compound heterozygotes for *Gata4* and *Gata5* or for *Gata5* and *Gata6* also show DORV and VSDs (59). The endocardium-specific requirements for GATA factors remain to be addressed, but the foregoing findings suggest that TBX20 may interact with GATA factors within endocardial lineages to effect septation.

*Identification of long-range enhancer targets with limited numbers of cells.* Mapping enhancers to their true target genes remains challenging, especially when it concerns developmental processes that occur in small numbers of cells. In this study, we show that the intersection of whole-heart TBX20 ChIP-seq and endocardial lineage-specific assessment of open chromatin is a powerful tool for identifying putative endocardial lineage enhancers. In addition, we found that the intersection of these data sets with existing HiC data from a similar cell type (51) proved instrumental in mapping long-range-acting, cell-type-specific enhancers to their putative target gene. With the observation that chromatin architecture is remarkably well conserved between cell types and between species (60, 61), we predict that our approach can be broadly applied to other systems, working with limited numbers of cells.

*TBX20 regulates Vcan expression in endocardial and mesenchymal cells via a distant enhancer.* The intersection of our data sets with HiC data suggested that *Vcan*, downregulated in *Tie2-Cre Tbx20<sup>f/null</sup>* mutant endocardial lineages, might have long-range enhancers that were directly targeted by TBX20. This was of particular interest, as mice homozygously mutant for hypomorphic alleles of *Vcan* have phenotypes, including hypomorphic atrioventricular cushions, reduced DMP, and lack of myocardialization of the OFT, that are strikingly similar to those of *Tie2-Cre Tbx20<sup>f/null</sup>* mutants (22). Ex vivo, cushion cells hypomorphic for *Vcan* have a reduced capacity to invade a collagen matrix, demonstrating that *Vcan* is a dosage-sensitive regulator of cell migration and invasion in cardiac septation (22, 62). Additionally, recent studies have demonstrated that *Vcan* is required in *Tie2-Cre* lineages for multiple aspects of cardiac septation (17, 20, 21).

Our analysis of distant enhancers in the *Vcan* locus identified lineage-specific enhancers that drive expression in cells that normally express *Vcan*, strongly suggesting that we have successfully identified cardiac enhancers for *Vcan*. These enhancers are

likely to be functionally conserved between species, as chromatin interactions between the *Vcan* enhancers we identified and the *Vcan* TSSs were found in human and mouse HiC experiments (51, 60). Furthermore, the enhancers have epigenetic marks of active enhancers in human hearts. Taken together, our results provide evidence that the regulation of *Vcan* by TBX20 in endocardial lineages is conserved between humans and mice and that *VCAN* misregulation in patients with TBX20 mutations may contribute to CHDs such as DORV or AVSD (3, 6).

High expression levels of *VCAN* have been found in most malignancies and are associated with cancer relapse and poor patient outcomes in numerous cancer types, including breast and prostate cancers (63). *VCAN* is predominantly secreted by activated peritumoral fibroblasts and vasculature and has also been reported to be secreted by some cancer cells (63, 64). Inhibition of *VCAN* expression or function in glioma cells, ovarian cancer cells, or lung cancer cells has been shown to inhibit cell invasion (64–66). Targeting of *VCAN* thus provides a potential strategy for suppressing tumor invasion (65, 67, 68). Our identification of a transcriptional enhancer that is active in fibroblasts may prove useful in understanding *Vcan* regulation in the context of tumorigenesis and could provide a therapeutic target.

## Methods

*Mouse strains.* Mice were maintained on a Black Swiss (NIHBL(S); Taconic Biosciences) outbred background. *Tbx20*-conditional and -null alleles and the *Tbx20-GFP* and *Nfatc1-Cre* mouse lines were generated in the authors' laboratories and have been previously described (8, 16, 31). *Tie2-Cre* (69), *Rosa26 flox-stop-flox tdTomato (Rosa tdTom)* (70), and *Rosa26* membrane-targeted (*m*) *tdTom flox-stop-flox mGFP (Rosa mT/mGFP)* (71) mice were obtained from The Jackson Laboratory.

*Histology.* Embryos were isolated in PBS and fixed in 4% paraformaldehyde in PBS at 4°C for 6 to 12 hours. Tissue was dehydrated using ethanol (50%, 70%, 80%, 90%, 96%, 3 × 100%; 5 minutes each), cleared in Histo-Clear (National Diagnostics) for 30 minutes, incubated 3 times for 30 minutes in Paraplast X-tra (Leica Biosystems), and embedded for sectioning. Alternatively, tissue was dehydrated in a sucrose gradient (5%, 10%, 15%, 20%, 6–12 hours each), embedded in 20% sucrose:OCT, and frozen for cryosectioning. Sections (10- to 12- $\mu$ m thick) were used for H&E staining or immunostaining, as described previously (31).

*Cushion explant assay.* E10.5 OFT cushions were dissected and cultured on collagen matrices as described previously (72). After 72 hours, explants were fixed for 30 minutes in 4% PFA, stained with phalloidin-647 and DAPI, and imaged using a Leica TCS SP5 confocal microscope. Quantifications were performed on *Tie2-Cre* lineage-traced cells.

*Cell quantification experiments.* Cell-counting experiments were performed using Volocity Software (PerkinElmer) or manually using ImageJ (NIH). In vivo quantification of cell numbers and EdU labeling were performed by counting all lineage-traced cushion cells in every other section of E11.5 hearts. Cell-counting experiments in vivo cushion explants were performed by manually counting all lineage-traced mesenchymal cells that migrated away from the explant over the surface of the collagen gel, using a Leica SP5 confocal microscope. Quantification of invasion was performed using differential interference microscopy, counting all *Tie2-Cre* lineage-traced cells at marked distances beneath the surface of the collagen gel in

1 field of view ( $\times 5$  objective), which typically represents one-quarter of the entire explant. Migration distance was defined as the distance between the edge of the explant and the cell body. The number of biological replicates used for these experiments was 3 or more and is reported in the figure legends.

**Abs.** Abs used for immunostaining included GFP (Ab13790; 1:1,000; Abcam); ACTN2 (A7811; 1:200; Sigma-Aldrich); TNNT2 (Ab106076; 1:2,000; Abcam); cleaved caspase-3 (9664; 1:200; Cell Signaling Technology); VCAN (Ab1033; 1:200; EMD Millipore); and CD31 (73117; 1:50; BD Biosciences). EdU (100  $\mu$ l, 3 g/l) was injected i.p. into female mice 2 hours prior to isolation of the embryos. EdU staining (C10337; Molecular Probes) was performed according to the manufacturer's instructions. The following FACS Abs were obtained from eBioscience: CD31-APC (17-0311-82; 1:50); TIE2-PE (12-5987-81; 1:50); CD41 PE-Cy7 (25-0411-82; 1:100); CD45 PE-Cy7 (25-0451-82; 1:100); and Ter119 PE-Cy7 (25-5921-82; 1:100).

**FACS sorting.** Hearts were harvested in ice-cold PBS with 0.5 mM EDTA and dissociated as described previously (73), using collagenase (10 mg/ml; Worthington) and dispase (10 mg/ml, Invitrogen). Cells were stained for 30 minutes, and DAPI was used to label dead cells. *Tie2-Cre* lineage cells, negative for the blood markers CD41, CD45, and Ter119, were sorted on an Influx Cell Sorter (BD Biosciences) and collected in RNA lysis buffer (Agilent) or TRIzol reagent (Ambion) for RNA extraction, or in PBS and 5% serum for ATAC-seq.

**RNA extraction, microarray analysis, and qRT-PCR.** RNA was isolated from 10,000 FACS-sorted cells using the Agilent RNA Nanoprep Kit. Amplification, hybridization to the Affymetrix Mouse Genome 430 2.0 Array GeneChip, and analysis were performed as described previously (73). Genes "present" in controls and downregulated or "present" in mutants and upregulated by 1.5-fold or more were considered differentially expressed. Data are available in the NCBI's Gene Expression Omnibus (GEO) database (GEO GSE73858). For qRT-PCR, RNA was extracted from sorted cells using TRIzol, and cDNA was generated using SuperScript VILO (Invitrogen). qRT-PCR was performed using FastStart SYBR Green Master Mix (Roche) on a Bio-Rad CFX96 Real-Time PCR System and the primers listed in Supplemental Table 5.

**TBX20 ChIP-seq.** E11.5 TBX20-GFP mouse hearts were dissected and cross-linked in 1.8% formaldehyde for 10 minutes at room temperature, quenched in 2.5 M glycine for 5 minutes, washed, and resuspended in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris, pH 8.1). Cross-linked chromatin was sonicated using a Misonix 4000 (amplitude of 90; 30 seconds on/30 seconds off) to an average size of 200 bp. Pre-cleared chromatin extract was incubated overnight at 4°C with goat anti-GFP and immunoprecipitated with protein A sepharose beads. DNA (10 ng) was used to generate a standard Illumina sequencing library. Sequencing reads were aligned to mm9 using bowtie2 2.2.3 (74). Uniquely mapping reads with a quality greater than 1 were accepted (6.6 M reads). HOMER (32) was used to call peaks using default parameters against 10-M input reads. Data are available in the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) database under accession number E-MTAB-3967.

**ATAC-seq.** For each sample, 2 hearts were pooled prior to FACS sorting. Cell preparation, transposition reaction, and library amplification (12 cycles) were performed as previously described (36). Alternatively, libraries were made using a Kapa KK2702 Real-Time Library

Amplification Kit (Kapa Biosystems) in a 50- $\mu$ l reaction for 12 cycles (72°C for 5 min, 98°C for 30 s; 12 cycles of 98°C for 10 s, 63°C for 30 s, and 72°C for 1 min), quantified using a Kapa Library Amplification Kit, and checked on the gel as described previously (36), as well as by using a D1000 Screentape Assay (Agilent). Adapters (CTGTCTCT-TATACACATCT and AGATGTGTATAAGAGACA) were cut from reads using cutadapt (75). Paired-end alignment was performed with bowtie2 2.2.3, with a maximum distance of 2 kb between mates. Peaks were called using HOMER (32) against whole-heart input with a  $1 \times 10^{-8}$  FDR cutoff. Data are available in the EMBL-EBI database under accession number E-MTAB-3972.

**Vcan enhancer in vivo reporter assays.** *Vcan enh1* (Chr13: 90973076-90973914) and *Vcan enh2* (Chr13: 91025481-91026597) were cloned into Hsp68-*LacZ* (76), and mouse in vivo transgenic reporter assays were performed by pronuclear injection as reported previously (77). TBX20 motifs in *Vcan enh1* were modified by Gene Tailor site-directed mutagenesis (Invitrogen) using the primers listed in Supplemental Table 5. Embryos were harvested at E12.5 and subjected to X-gal staining as reported previously (77). Embryos were genotyped using *LacZ* primers (forward, 5'-CADCCTGAATGGCGAATGGCGCTT; reverse, 5'-CCCGTTGCACCACAGATGAAACGC), and embryos that were (nearly) completely stained blue (between 1 and 3 for each line) were excluded from further analysis.

**Statistics.** Values are presented as the mean  $\pm$  SEM. For ex vivo cell migration assays, a nonpaired, 2-tailed *t* test was used to analyze the migration distance between controls and mutants. Matrix invasion was analyzed using 2-way ANOVA, with Bonferroni's post tests. An unpaired, 2-tailed *t* test was used to analyze qRT-PCR data. A *P* value of less than 0.05 was considered significant.

**Study approval.** Animal care and experimental procedures were performed according to protocols approved by the IACUCs of the UCSD (protocol S04150) and the University of Chicago (protocol 71656).

## Author contributions

CJB, IA, NS, RJD, QJC, JC, MAN, and SME designed and/or performed experiments and analyzed the data. CJB, IA, NS, MAN, and SME wrote the manuscript. BZ provided reagents.

## Acknowledgments

The ATAC-seq library was prepared in collaboration with the IGM Genomics Center of the UCSD. Confocal microscopic analysis was performed at the UCSD School of Medicine's Microscopy Core, which is supported by a NINDS P30 core grant (NS047101). This work was supported by NIH grants HL119967, HL114010, and HL123857 (to M.A. Nobrega) and HL123747, HL117649, and HL074066 (to S.M. Evans); the Netherlands Organization for Scientific Research (NWO, 825.10.016, to C.J. Boogerd); the American Heart Association (13POST16480012, to C.J. Boogerd); and the California Institute for Regenerative Medicine (TG2-01154, to R.J. Dirschinger).

Address correspondence to: Marcelo A. Nobrega, Department of Human Genetics, University of Chicago, 920 E. 58th Street CLSC 319, Chicago, Illinois 60637, USA. Phone: 773.834.7919; E-mail: nobrega@uchicago.edu. Or to: Sylvia M. Evans, Skaggs School of Pharmacy, UCSD, 9500 Gillman Road, La Jolla, California 92093, USA. Phone: 858.822.2452; E-mail: syevans@ucsd.edu.

1. Hoffman JI, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol.* 2002;39(12):1890–900.
2. Töpfer A, et al. Functionally significant, rare transcription factor variants in tetralogy of Fallot. *PLoS One.* 2014;9(8):e95453.
3. Qian L, et al. Transcription factor neuro-mancer/TBX20 is required for cardiac function in *Drosophila* with implications for human heart disease. *Proc Natl Acad Sci U S A.* 2008;105(50):19833–19838.
4. Kirk EP, et al. Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. *Am J Hum Genet.* 2007;81(2):280–291.
5. Posch MG, et al. A gain-of-function TBX20 mutation causes congenital atrial septal defects, patent foramen ovale and cardiac valve defects. *J Med Genet.* 2010;47(4):230–235.
6. Pan Y, et al. TBX20 loss-of-function mutation contributes to double outlet right ventricle. *Int J Mol Med.* 2015;35(4):1058–1066.
7. Takeuchi JK, et al. Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. *Development.* 2005;132(10):2463–2474.
8. Cai CL, et al. T-box genes coordinate regional rates of proliferation and regional specification during cardiogenesis. *Development.* 2005;132(10):2475–2487.
9. Singh MK, et al. Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2. *Development.* 2005;132(12):2697–2707.
10. Stennard FA, et al. Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation. *Development.* 2005;132(10):2451–2462.
11. Cai X, et al. Tbx20 acts upstream of Wnt signaling to regulate endocardial cushion formation and valve remodeling during mouse cardiogenesis. *Development.* 2013;140(15):3176–3187.
12. Anderson RH, Mohun TJ, Brown NA. Clarifying the morphology of the ostium primum defect. *J Anat.* 2015;226(3):244–257.
13. Mommersteeg MT, et al. Two distinct pools of mesenchyme contribute to the development of the atrial septum. *Circ Res.* 2006;99(4):351–353.
14. Lin CJ, Lin CY, Chen CH, Zhou B, Chang CP. Partitioning the heart: mechanisms of cardiac septation and valve development. *Development.* 2012;139(18):3277–3299.
15. Creazzo TL, Godt RE, Leatherbury L, Conway SJ, Kirby ML. Role of cardiac neural crest cells in cardiovascular development. *Annu Rev Physiol.* 1998;60:267–286.
16. Wu B, et al. Endocardial cells form the coronary arteries by angiogenesis through myocardial-endocardial VEGF signaling. *Cell.* 2012;151(5):1083–1096.
17. Yamamura H, Zhang M, Markwald RR, Mjaatvedt CH. A heart segmental defect in the anterior-posterior axis of a transgenic mutant mouse. *Dev Biol.* 1997;186(1):58–72.
18. van den Hoff MJ, et al. Myocardialization of the cardiac outflow tract. *Dev Biol.* 1999;212(2):477–490.
19. Phillips HM, et al. Disruption of planar cell polarity signaling results in congenital heart defects and cardiomyopathy attributable to early cardiomyocyte disorganization. *Circ Res.* 2007;101(2):137–145.
20. Mjaatvedt CH, Yamamura H, Capehart AA, Turner D, Markwald RR. The *Cspg2* gene, disrupted in the *hdf* mutant, is required for right cardiac chamber and endocardial cushion formation. *Dev Biol.* 1998;202(1):56–66.
21. Yoon SH, et al. Versican secretion from endocardium is required for chamber formation via the recruitment of cardiac progenitors. *Circulation.* 2012;126(21 Supplement):A16246.
22. Burns TA, et al. Imbalanced expression of Vcan mRNA splice form proteins alters heart morphology and cellular protein profiles. *PLoS One.* 2014;9(2):e89133.
23. Wirrig EE, et al. Cartilage link protein 1 (Crtl1), an extracellular matrix component playing an important role in heart development. *Dev Biol.* 2007;310(2):291–303.
24. Mommersteeg MT, Yeh ML, Parnavelas JG, Andrews WD. Disrupted Slit-Robo signalling results in membranous ventricular septum defects and bicuspid aortic valves. *Cardiovasc Res.* 2015;106(1):55–66.
25. Tullio AN, et al. Nonmuscle myosin II-B is required for normal development of the mouse heart. *Proc Natl Acad Sci U S A.* 1997;94(23):12407–12412.
26. Ma X, Adelstein RS. A point mutation in *Myh10* causes major defects in heart development and body wall closure. *Circ Cardiovasc Genet.* 2014;7(3):257–265.
27. Lee MP, Yutzey KE. *Twist1* directly regulates genes that promote cell proliferation and migration in developing heart valves. *PLoS One.* 2011;6(12):e29758.
28. Chen YH, Ishii M, Sucov HM, Maxson RE. *Msx1* and *Msx2* are required for endothelial-mesenchymal transformation of the atrioventricular cushions and patterning of the atrioventricular myocardium. *BMC Dev Biol.* 2008;8:75.
29. Zhang L, et al. Characterization of a t(5;8)(q31;q21) translocation in a patient with mental retardation and congenital heart disease: implications for involvement of *RUNX1T1* in human brain and heart development. *Eur J Hum Genet.* 2009;17(8):1010–1018.
30. Sakabe NJ, et al. Dual transcriptional activator and repressor roles of TBX20 regulate adult cardiac structure and function. *Hum Mol Genet.* 2012;21(10):2194–2204.
31. Shen T, et al. Tbx20 regulates a genetic program essential to adult mouse cardiomyocyte function. *J Clin Invest.* 2011;121(12):4640–4654.
32. Heinz S, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell.* 2010;38(4):576–589.
33. Stennard FA, et al. Cardiac T-box factor Tbx20 directly interacts with *Nkx2-5*, *GATA4*, and *GATA5* in regulation of gene expression in the developing heart. *Dev Biol.* 2003;262(2):206–224.
34. Furtado MB, et al. Cardiogenic genes expressed in cardiac fibroblasts contribute to heart development and repair. *Circ Res.* 2014;114(9):1422–1434.
35. Ashburner M, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25–29.
36. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.* 2013;10(12):1213–1218.
37. Booger CJ, Moorman AF, Barnett P. Expression of muscle segment homeobox genes in the developing myocardium. *Anat Rec (Hoboken).* 2010;293(6):998–1001.
38. Gao Z, et al. *Ets1* is required for proper migration and differentiation of the cardiac neural crest. *Development.* 2010;137(9):1543–1551.
39. Timmerman LA, et al. Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes Dev.* 2004;18(1):99–115.
40. Lindsley A, et al. Identification and characterization of a novel Schwann and outflow tract endocardial cushion lineage-restricted periostin enhancer. *Dev Biol.* 2007;307(2):340–355.
41. Blow MJ, et al. ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet.* 2010;42(9):806–810.
42. Schachterle W, Rojas A, Xu SM, Black BL. ETS-dependent regulation of a distal *Gata4* cardiac enhancer. *Dev Biol.* 2012;361(2):439–449.
43. Gross DS et al. Nuclease hypersensitive sites in chromatin. *Annu Rev Biochem.* 1988;57:159–97.
44. Holwerda SJ, de Laat W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond B Biol Sci.* 2013;368(1620):20120369.
45. Graves BJ, Petersen JM. Specificity within the ets family of transcription factors. *Adv Cancer Res.* 1998;75:1–55.
46. Hollenhorst PC, Jones DA, Graves BJ. Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic Acids Res.* 2004;32(18):5693–5702.
47. De Val S, Black BL. Transcriptional control of endothelial cell development. *Dev Cell.* 2009;16(2):180–195.
48. de Laat W, Grosveld F. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.* 2003;11(5):447–459.
49. Smallwood A, Ren B. Genome organization and long-range regulation of gene expression by enhancers. *Curr Opin Cell Biol.* 2013;25(3):387–394.
50. Smemo S, et al. Obesity-associated variants within FTO form long-range functional connections with *IRX3*. *Nature.* 2014;507(7492):371–375.
51. Jin F, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature.* 2013;503(7475):290–294.
52. Roadmap Epigenomics Consortium, et al. Integrative analysis of 111 reference human epigenomes. *Nature.* 2015;518(7539):317–330.
53. Briggs LE, Kakarla J, Wessels A. The pathogenesis of atrial and atrioventricular septal defects with special emphasis on the role of the dorsal mesenchymal protrusion. *Differentiation.* 2012;84(1):117–130.
54. Zhang T, et al. Jun is required in *Isl1*-expressing progenitor cells for cardiovascular development. *PLoS One.* 2013;8(2):e57032.

55. De Val S, et al. Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. *Cell*. 2008;135(6):1053–1064.
56. Bouveret R, et al. NKX2-5 mutations causative for congenital heart disease retain functionality and are directed to hundreds of targets. *Elife*. 2015;4.
57. Rivera-Feliciano J, et al. Development of heart valves requires Gata4 expression in endothelial-derived cells. *Development*. 2006;133(18):3607–3618.
58. Pu WT, Ishiwata T, Juraszek AL, Ma Q, Izumo S. GATA4 is a dosage-sensitive regulator of cardiac morphogenesis. *Dev Biol*. 2004;275(1):235–244.
59. Laforest B, Nemer M. GATA5 interacts with GATA4 and GATA6 in outflow tract development. *Dev Biol*. 2011;358(2):368–378.
60. Rao SS, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665–1680.
61. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012;485(7398):376–380.
62. Hatano S, et al. Versican/PG-M is essential for ventricular septal formation subsequent to cardiac atrioventricular cushion development. *Glycobiology*. 2012;22(9):1268–1277.
63. Theocharis AD, et al. Cell-matrix interactions: focus on proteoglycan-proteinase interplay and pharmacological targeting in cancer. *FEBS J*. 2014;281(22):5023–5042.
64. Yeung TL, et al. TGF- $\beta$  modulates ovarian cancer invasion by upregulating CAF-derived versican in the tumor microenvironment. *Cancer Res*. 2013;73(16):5016–5028.
65. Arslan F, Bosserhoff AK, Nickl-Jockschat T, Doerfelt A, Bogdahn U, Hau P. The role of versican isoforms V0/V1 in glioma migration mediated by transforming growth factor- $\beta$ 2. *Br J Cancer*. 2007;96(10):1560–1568.
66. Said N, Sanchez-Carbayo M, Smith SC, Theodorou D. RhoGDI2 suppresses lung metastasis in mice by reducing tumor versican expression and macrophage infiltration. *J Clin Invest*. 2012;122(4):1503–1518.
67. Havre PA, Dang LH, Ohnuma K, Iwata S, Morimoto C, Dang NH. CD26 expression on T-anaplastic large cell lymphoma (ALCL) line Karpas 299 is associated with increased expression of versican and MT1-MMP and enhanced adhesion. *BMC Cancer*. 2013;13:517.
68. Ween MP, Hummitzsch K, Rodgers RJ, Oehler MK, Ricciardelli C. Versican induces a pro-metastatic ovarian cancer cell behavior which can be inhibited by small hyaluronan oligosaccharides. *Clin Exp Metastasis*. 2011;28(2):113–125.
69. Kisanuki YY, Hammer RE, Miyazaki J, Williams SC, Richardson JA, Yanagisawa M. Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis in vivo. *Dev Biol*. 2001;230(2):230–242.
70. Madisen L, et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci*. 2010;13(1):133–140.
71. Muzumdar MD, Tasic B, Miyamichi K, Li L, Luo L. A global double-fluorescent Cre reporter mouse. *Genesis*. 2007;45(9):593–605.
72. Runyan RB, Markwald RR. Invasion of mesenchyme into three-dimensional collagen gels: a regional and temporal analysis of interaction in embryonic heart tissue. *Dev Biol*. 1983;95(1):108–114.
73. Van Handel B, et al. Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. *Cell*. 2012;150(3):590–605.
74. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–359.
75. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011;17(1):10–12.
76. Gotea V, Visel A, Westlund JM, Nobrega MA, Pennacchio LA, Ovcharenko I. Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. *Genome Res*. 2010;20(5):565–577.
77. Smemo S, Campos LC, Moskowitz IP, Krieger JE, Pereira AC, Nobrega MA. Regulatory variation in a TBX5 enhancer leads to isolated congenital heart disease. *Hum Mol Genet*. 2012;21(14):3255–3263.