The cardiac TBX5 interactome reveals a chromatin remodeling network essential for cardiac septation

Lauren Waldron[1,2], **Jeffrey D. Steimle**[8], **Todd M. Greco**[7], **Nicholas C. Gomez**[2,4], **Kerry M. Dorr**[1,2], **Junghun Kweon**[8], **Brenda Temple**[6], **Xinan Holly Yang**[8], **Caralynn M. Wilczewski**[1,2], **Ian J. Davis**[3,4], **Ileana M. Cristea**[7], **Ivan P. Moskowitz**[8], and **Frank L. Conlon**[1,2,3,4,5]

[1]University of North Carolina McAllister Heart Institute, UNC-Chapel Hill, Chapel Hill, NC 27599, USA

[2]Integrative Program for Biological & Genome Sciences, UNC-Chapel Hill, Chapel Hill, NC 27599, USA

[3]Department of Genetics, UNC-Chapel Hill, Chapel Hill, NC 27599, USA

[4]Lineberger Comprehensive Cancer Center, UNC-Chapel Hill, Chapel Hill, NC 27599, USA

[5]Department of Biology, UNC-Chapel Hill, Chapel Hill, NC 27599, USA

[6]R.L. Juliano Structural Bioinformatics Core, Department of Biochemistry and Biophysics, UNC-Chapel Hill, Chapel Hill, NC 27599, USA

[7]Department of Molecular Biology, Princeton University, NJ 08544, USA

[8]Departments of Pediatrics, Pathology, and Human Genetics, The University of Chicago, Il, 60637, USA

SUMMARY

Human mutations in the cardiac transcription factor gene TBX5 cause Congenital Heart Disease (CHD), however the underlying mechanism is unknown. We report characterization of the endogenous TBX5 cardiac interactome and demonstrate that TBX5, long considered a transcriptional activator, interacts biochemically and genetically with the Nucleosome Remodeling and Deacetylase (NuRD) repressor complex. Incompatible gene programs are repressed by TBX5 in the developing heart. CHD missense mutations that disrupt the TBX5-NuRD interaction cause depression of a subset of repressed genes. Furthermore, the TBX5-NuRD interaction is required

for heart development. Phylogenetic analysis showed that the TBX5-NuRD interaction domain evolved during early diversification of vertebrates, simultaneous with the evolution of cardiac septation. Collectively, this work defines a TBX5-NuRD interaction essential to cardiac development and the evolution of the mammalian heart, and when altered may contribute to human CHD.

INTRODUCTION

Congenital malformations, or structural birth defects, are the leading cause of infant mortality in the US and Europe (Dolk et al., 2010; Heron and Tejada-Vera, 2009). Congenital heart disease (CHD) remains the most common congenital malformation, with atrial septal defects (ASD) present in 116 per 100,000 live births and ventricular septal defects (VSD) present in 307 per 100,000 live births (Gittenberger-de Groot et al., 2014). Holt Oram Syndrome (HOS) is an autosomal disorder associated with cardiac septal defects and is caused by dominant mutations in the T-box transcription factor *Tbx5* (Basson et al., 1997; Basson et al., 1994; Basson et al., 1999; Cross et al., 2000; Holt and Oram, 1960; Li et al., 1997). Mice heterozygous for a null mutation in *Tbx5* display many of the phenotypic abnormalities of CHD, while mice homozygous for a null mutation in *Tbx5* die by E10.5 due to cardiac defects (Bruneau et al., 1999; Bruneau et al., 2001; Moskowitz et al., 2007; Moskowitz et al., 2004). In addition to its role in cardiogenesis, *Tbx5* is one of the key factors essential for cellular reprogramming of fibroblasts into induced cardiomyocytes (Ieda et al., 2010; Qian et al., 2012).

While *Tbx5* is an essential transcription factor for heart development and reprogramming, and its disease relevance is well established, there are many critical unanswered questions about the mechanism of TBX5 function. In particular, the TBX5 interactome has not been defined in-vivo, leaving unanswered the question of how TBX5 interacting proteins regulate TBX5's choice of distinct transcriptional targets, or how these interactions function to activate and/or repress target gene transcription. It is also not known how these events, when disrupted, contribute to CHD. To these ends, we initiated a directed proteomic-based approach to identify endogenous cardiac proteins that function in physical association with TBX5.

We report the discovery that TBX5 interacts biochemically and genetically with the transcriptional repression machinery of the Nucleosome Remodeling and Deacetylase (NuRD) complex to repress inappropriate expression of gene programs in the developing heart. We have further uncovered a domain of TBX5 that forms an α-Helix; we show that TBX5 missense mutations associated with septal defects in human patients map to this domain. Furthermore, HOS patient mis-sense mutations in this domain disrupt the TBX5-NuRD interaction and elicit misexpression of inappropriate gene programs that are normally repressed by TBX5. Through phylogenetic analysis across numerous vertebrate species, we find that the NuRD-interaction domain of TBX5, and hence the interaction of TBX5 with NuRD, evolved during the early diversification of vertebrates, simultaneous with the evolution of cardiac septation. Thus, the TBX5-NuRD interaction identified here plays a

central role in both cardiac development and human congenital heart defects, as well as the evolution of the mammalian heart.

RESULTS

Generation of the Tbx5[Avi] allele

Past attempts to isolate endogenous T-box protein complexes have been hampered by their relative low abundance and insolubility. Therefore, we introduced an Avitag epitope into the murine *Tbx5* locus via homologous recombination to generate a *Tbx5[Avi]* allele (Figure 1A–D). The Avitag is a small peptide tag that is biotinylated in the presence of the bacterial enzyme BirA. Thus, the Avitag/BirA system combines minimal structural invasiveness with the specificity and strength of the biotin-streptavidin interaction, the strongest noncovalent peptide-ligand interaction identified to date (Bayer and Wilchek, 1980; Maine et al., 2010; Roesli et al., 2006; van Werven and Timmers, 2006; Wang et al., 2006); *Tbx5[Avi]* mice were mated to mice ubiquitously expressing BirA (*Rosa26[BirA]*) (Driegen et al., 2005). The *Tbx5[Avi]*; *Rosa26[BirA]* compound homozygous mice show no overt phenotype, have no cardiac phenotype, and are fertile. Thus, the Avitag appears to have no effect on TBX5 activity, and expression of biotinylated TBX5[Avi] is not deleterious to embryonic development.

Isolation and characterization of the endogenous TBX5 interactome

To define the TBX5 cardiac interactome, we performed mass spectrometry analysis of affinity-purified TBX5 complexes from *Tbx5[Avi/Avi];Rosa26[BirA/BirA]* cardiac nuclei (Figure 1E). We used an unbiased gene ontology-based bioinformatics classification to screen the functions of proteins associated with TBX5 at a statistical relevance versus *Tbx5[Avi/Avi]* alone (no BirA) and *Rosa26[BirA/BirA]* alone. This analysis defined a subset of 58 candidate interactions (Table S1). We analyzed these candidates by functional network analysis; drawing upon the STRING database of protein-protein interactions, 40 out of the 58 candidates were assigned to a single interconnected network (Figure 1F, S1A).

TBX5 interacts with components of the Nucleosome Remodeling and Deacetylase (NuRD) complex in vivo

Analysis of the TBX5 interaction network identified the Heterodimeric Facilitates Chromatin Transcription (FACT) complex and several members of the transcription initiation factor TFIID complex, both of which play roles in RNA Pol II-dependent gene transcription (Berk, 1999; Saunders et al., 2003) (Figure 1F). TBX5 has been long considered to be a cardiac transcriptional activator. Surprisingly, in our analysis we identified TBX5 in association with multiple components of the NuRD chromatin-modifying complex (Figure 1F, G), which in most instances functions to repress gene transcription (Wade et al., 1998; Xue et al., 1998). We find TBX5 in association with six components of the NuRD transcriptional repression complex: CHD4, MTA1, RBBP4, GATAD2a, GATAD2b, and HDAC2. All the components required for a functional NuRD complex were identified in association with TBX5, as CHD3 commonly acts in a mutually exclusive manner with CHD4, RBBP7 and RBBP4 act redundantly, and an additional component of the NuRD complex, MBD3, has been suggested to be dispensable for

transcription factor recruitment (e.g., by hunchback (Kehle et al., 1998). We confirmed the interaction of TBX5 with multiple members of the NuRD complex in adult cardiac nuclear extracts (Figure 1G).

As TBX5 has been shown to be required for cardiogenesis at E9.5 (Bruneau et al., 2001), we established expression of components of the NuRD complex in cardiac tissue at E9.5. We found that members of the NuRD complex are strongly expressed in the heart at this stage by RT-PCR (Figure S1B). We further demonstrate that endogenous TBX5 (untagged TBX5) interacts with CHD4 at E9.5 (Figure 1H).

TBX5 and MTA1 and cardiac septation

To assess whether the TBX5-NuRD complex interaction was physiologically relevant, we analyzed mutant alleles of *Tbx5* and a central component of the NuRD complex, *Mta1*, for a genetic interaction (Figure 2, S2). At E13.5, *Mta1*$^{+/-}$ mice have no apparent cardiac defects (Figure 2C). *Tbx5*$^{+/-}$ mice exhibit partially penetrant cardiac defects as previously described, including ASD, VSD, and complete common atrioventricular canal (CCAVC), which includes both ASD and VSD components (Figure 2D) (Baban et al., 2014; Basson et al., 1994; Basson et al., 1999; Benson et al., 1996; Bruneau et al., 1999; Bruneau et al., 2001; McDermott et al., 2005). *Tbx5*$^{+/-}$; *Mta1*$^{+/-}$ heterozygous null embryos all demonstrated septal defects (19/19), a higher frequency of septal defects than either single heterozygote alone (0/14 for *Mta1*$^{+/-}$ and 12/16 for *Tbx5*$^{+/-}$) (p = 7.1E-9 vs. WT and p = 0.035 vs *Tbx5*$^{+/-}$) (Figure 2E–J). Furthermore, *Mta1*$^{-/-}$ hearts also do not have septal defects at this stage (Figure S3). We observed no change in the left ventricular chamber width or left ventricular wall thickness (Figure S2D–E) or inferior or superior cushions size (Figure 2F–I). To rule out a developmental delay in septation in the *Tbx5*$^{+/-}$; *Mta1*$^{+/-}$ heterozygous null embryos, we examined *Tbx5*$^{+/-}$; *Mta1*$^{+/-}$ heterozygous null embryos at E14.5 and consistently observe CCAVCs at this time point (Figure S2F–G). These results demonstrate a genetic interaction between *Tbx5* and *Mta1* in intro-cardiac septation providing genetic evidence for a role for the TBX5-NuRD interaction.

TBX5 binds to the same consensus DNA site in activated and repressed target genes

The finding that TBX5 and NuRD complex components interact biochemically and genetically led us to hypothesize that TBX5 can function to repress aberrant gene expression in the developing heart. To test this hypothesis, we performed RNA-seq analysis and identified genes significantly downregulated and significantly upregulated in *Tbx5* null verses wildtype heart tissue (Figure 3, Table S2). To determine which genes were direct targets of TBX5, we overlaid our data set to that of a TBX5 ChIP-seq data set (He et al., 2011) and observe that TBX5 activates and represses sets of genes that fall into distinct categories (Fig. 3A–C). This finding raises the possibility that TBX5 may distinguish activated verses repressed target genes by binding to dissimilar DNA sequences in their regulatory elements. To address this possibility we performed bioinformatics analysis on a large scale. From our ChIP-seq data, 71,624 TBX5 peak regions were identified by enrichment of signal using MACS2 (Zhang et al., 2008). These peak regions demonstrated a significant association with genes that were differentially regulated between *Tbx5* wild type and *Tbx5*$^{-/-}$ hearts.

Given a large number of peaks can skew results in gene-association analyses, we varied the number of analyzed peaks from 10,000 to 70,000 in increments of 10,000. In every instance, TBX5 was significantly associated with differentially expressed genes (data not shown). We selected the top 40,000 peaks for subsequent analyses since the inclusion of additional peaks minimally increased the number of associated genes (data not shown). We identified an association between TBX5 and 82% of genes differentially expressed between *Tbx5* wild type and *Tbx5*[−/−] hearts (2988 of 3634, Figure 3A). Of these genes, about half were activated and half were repressed by TBX5. Both groups demonstrated a statistically significant enrichment in TBX5 binding sites in the TBX5 ChIP-peak regions (activated 4.88 $\times 10^{-141}$, repressed p = 1.29×10^{-142}, KS test, Figure 3B). Genes repressed by TBX5 were enriched in pathways involved in cancer as well as cardiac diseases, while genes activated by TBX5 were enriched in pathways involved in cell cycle and DNA replication, a known function of TBX5 (Figure 3C) (Hatcher et al., 2001, Goetz et al., 2006).

TBX5 acts to directly repress inappropriate expression of genes in cardiac tissue

To validate the genes putatively repressed by TBX5, we rank ordered our list of repressed genes based on relative change in expression between wild-type and *Tbx5*[−/−] hearts and the number of reads from the ChIP-seq data set (Table S3). From this rank ordered list, we cloned the top 15 of the ChIP peak sequence elements from 14 genes and performed transcriptional assays (Figure 4, S4). Of the candidate genes tested, 11 were repressed by TBX5. This set of genes could be divided into two categories: *Casz1, Fgf11, Gad1, Kcne3, Kctd16, Klf4, Nxph4, Plekha2,* and *Sncb* are expressed in neural lineages (Diez-Roux et al., 2011; Dorr et al., 2015; Liu et al., 2011; Metz et al., 2011; Smallwood et al., 1996; Sopher et al., 2001; Trifonov et al., 2014; Wang et al., 2014), while *Casz1, Col20a1, Fgf11, Gad1, Klf4, Mef2b,* and *Plekha2* are expressed in a divergent subset of cancers (Costantini et al., 2009; Evans and Liu, 2008; Hu et al., 2015; Huang et al., 2013; Kimura et al., 2013; Ying et al., 2013). Altogether, these data suggest that TBX5 function is essential for *in vivo* transcriptional repression and prevention of ectopic expression of inappropriate gene programs in the heart.

TBX5 interaction with CHD4 does not require either the CHD4 chromo domains or the Phd domains

CHD4 is a highly conserved protein that serves as the central catalytic core component of the NuRD complex (Denslow and Wade, 2007). Studies across plants and animals have shown that CHD4 functions through interaction with DNA binding proteins by its pair of chromo domains. In addition CHD4 also contains a tandem set of Phd (plant homology domains) and an ATPase cassette (Denslow and Wade, 2007). To gain insight into the mechanisms of TBX5-NuRD mediated repression, we mapped the protein interaction domains of TBX5 of CHD4 (Figure 5A–F). Streptavidin pulldown of TBX5 complexes demonstrated that an N-terminal human CHD4 region containing the pair of Phd domains and the pair of chromo domains was required for interaction with mouse TBX5 (Figure 5A, C). Surprisingly, we found CHD4 deletions of either the Phd domains or the chromo domains still interact with TBX5 (Figure 5D) implying that TBX5 interacts with a previously undefined set of residues in the amino terminal portion of CHD4.

TBX5 interacts with the NuRD complex through a coil-α-helix domain

In vitro based approaches have led to the identification of a number of proteins that interact with TBX5 through two regions of TBX5; one in the amino-terminal portion of the protein including the most amino-terminal portion of the T-box DNA binding domain (Brown et al., 2005; Garg et al., 2003; Ghosh et al., 2009; Hiroi et al., 2001; Wang et al., 2011) and a second in the C-terminus of TBX5 containing a putative nuclear localization signal (Krause et al., 2004; Murakami et al., 2005). To determine if TBX5 associates with the NuRD complex through these regions we mapped the residues of TBX5 that was essential for interaction with CHD4. From these studies we found a region C-terminal to the T-box was necessary and sufficient for interaction with CHD4 (amino acids 238–340) (Figure 5B, E–F). A conserved function for this region is supported by our observation that the minimal region of TBX5 that interacts with CHD4 is highly conserved across 48 TBX5 orthologues (Figure 5H).

Since the region of TBX5 that interacts with CHD4 has not been included in TBX5 crystal structural analysis (Stirnimann et al., 2010) we sought to determine if the region of TBX5 that interacted with CHD4 represented a true TBX5 structural domain. We therefore conducted molecular 3D structural modeling of the entire TBX5 protein as well as the minimal region of TBX5 that we showed interacted with CHD4. Results from these analyses revealed that TBX5 residues 255–264 is highly predicted to form a small coil-to α-Helix region (Figure 5I,J). We have termed this structural-functional region of TBX5 the NuRD Interaction Domain (TBX5^{NID}) (Figure 5G).

The TBX5-NuRD interaction domain is essential for cardiac development and function in humans

Having demonstrated TBX5 and the NuRD complex genetically and biochemically interact, and having established that TBX5^{NID} represents a true structural functional domain, we next determined if the interaction between TBX5 and the NuRD complex is essential for TBX5 function. For these studies we undertook a CHD patient driven approach based on the observations that the vast majority of molecularly characterized TBX5 mutations associated with CHD are mis-sense mutations within the coding region of TBX5. Most of the CHD associated mutations are located within the TBX5 T-box causing either a loss of DNA binding or degradation by non-sense mediated decay (Fan et al., 2003a; Fan et al., 2003b; Mori and Bruneau, 2004). However in addition, there exists a subset of CHD-associated mutations that lie in the C-terminus of TBX5, outside the T-box; the resulting mutant forms of TBX5 are stable, nuclear localized and bind DNA. The molecular basis for these CHD-associated mutations has been unknown.

Multiple mis-sense mutations that correspond to TBX5-associated CHD in the C-terminal region of TBX5 that fall within the TBX5^{NID} including: TBX5^{S252I}, TBX5^{S261C}, TBX5^{V263M}, TBX5^{K266R}, and TBX5^{Q292R} (Heinritz et al., 2005) (Figure 5G). We investigated the physical location of these TBX5 amino acids on the predicted NID structural domain and observed the amino acids align along a single interaction surface of the TBX5^{NID} α-Helix (Figure 5I–J). We hypothesized the CHD mis-sense mutations in the TBX5^{NID} disrupt the TBX5-CHD4 interaction.

We tested the hypothesis that TBX5 mis-sense mutations in the TBX5^{NID} alter the TBX5-NuRD association. Three single mis-sense mutations within the TBX5^{NID} (TBX5^{S261C}, TBX5^{V263M}, and TBX5^{K266R}) each abolished the interaction between TBX5 and both CHD4 and MTA1, and an additional two other CHD single mis-sense mutations (TBX5^{S252I} and TBX5^{Q292R}) reduced the interactions (Figure 5K). These results suggest that disruption of the TBX5-NuRD complex *in vivo* causes cardiac abnormalities associated with HOS.

TBX5 can repress target genes in a NuRD dependent manner

Having demonstrated that the TBX5-NuRD interaction is disrupted by CHD-associated mutations in the NID, we queried if repression of TBX5 target genes was NuRD dependent. We performed transcriptional assays and observed, of the 11 validated targets repressed by TBX5, *Kctd16* and *Mef2b* fail to be repressed by TBX5 containing the CHD associated mutation TBX5^{S261C} (Figure 6A–B, S5). Furthermore, TBX5^{S261C} failed to repress transcription of multiple elements of the *Mef2b* gene (Figure 6B–D). In addition, we were able to ChIP CHD4 in the presence of TBX5 to these elements (Figure 6E–F). Taken together, these results imply the TBX5-NuRD complex is required to repress inappropriate gene expression in the heart and further suggest a subset of CHD mutants leads to the mis-expression of genes including *Mef2b* in the developing heart.

The TBX5-NuRD interaction arose concomitantly with the evolution of cardiac septation

Phylogenetic analysis of forty-eight TBX5 orthologues demonstrated there is complete conservation of the TBX5^{NID} amino acids mutated in TBX5-associated CHD in animals with a septated heart (TBX5^{S261}, TBX5^{V263}, TBX5^{K266}) (e.g. *Xenopus* to human) (Figure 7). In contrast, orthologues of TBX5 in vertebrates that lack a septated heart (e.g. zebrafish) have an amino acid substitution present in human CHD patients, and furthermore, this substitution ablates the TBX5-NuRD interaction (Figure 5K). We hypothesized that the *Xenopus* ortholog of Tbx5 would interact with the NuRD complex, but a *Xenopus* Tbx5 ortholog with zebrafish NID amino acid substitutions would fail to interact with the NuRD complex. To test this hypothesis we generated the K266R mutation in the *Xenopus* ortholog of Tbx5 (xTbx5^{K266R}). We observed that while mTBX5 and xTbx5 interact with the NuRD complex, xTbx5^{K266R} has a disrupted interaction with CHD4 (Figure 7B). Thus, our data suggested that the TBX5-NuRD interaction arose coincidentally with the evolution of cardiac septation and it is an essential component of an evolutionarily conserved mechanism for cardiac septum formation.

DISCUSSION

Here we report the isolation and characterization of the *in vivo* TBX5 interactome. For over 20 years, studies have defined TBX5 as an activator of transcription. However, our analysis demonstrates that TBX5 interacts biochemically and genetically with the transcriptional repression machinery of the NuRD complex. We further report TBX5 acts to repress inappropriate gene programs in cardiac tissue. We observed that only a subset of TBX5-repressed target genes failed to be repressed by the mutation of the TBX5-NuRD interacting domain (e.g. HOS mutation TBX5^{S261C}; Figure 6A–D, S5) implying that TBX5-mediated repression occurs in both a NuRD-dependent and NuRD-independent manner. We favor a

model by which TBX5 regulates growth and differentiation of cardiomyocytes in part through the repression of alternate lineage-specific gene regulatory networks.

To date, all other predicted TBX5 interactions have been identified through in vitro studies, such as GST pulldowns in bacterial or cell lysates or yeast 2-hybrid approaches (Brown et al., 2005; Garg et al., 2003; Ghosh et al., 2009; Hiroi et al., 2001; Wang et al., 2011)(Krause et al., 2004; Murakami et al., 2005). In our study we isolate and characterize the endogenous TBX5 interactome from cardiac tissue under physiological conditions. This in vivo isolation is highly significant given mutations in T-box genes are causative for over 30 human diseases. Though we found no evidence for and interaction with proteins previously identified as being associated with TBX5 (BRG1, NKX2.5, GATA4, TBX20, Myocardin, TAZ and BAF60C), it remains possible that these proteins interact with TBX5 at time points of cardiac development different from those we analyzed, or that they were not identified in our study for technical reasons (e.g. distinct lysis buffer conditions for different studies). Regarding the latter point we note that if we slightly relax the statistical cutoffs, BRG1 (SMARCA4) (Takeuchi et al., 2011) does appear to be in complex with TBX5.

TBX5 and cardiac septation

Our data suggest the TBX5-NuRD interaction arose coincidentally with cardiac septation. This hypothesis is further supported by our observations that $Tbx5^{+/-}; Mta1^{+/-}$ compound heterozygous mice have a higher frequency of cardiac septal defects than either single heterozygote alone (Figure 2J). AVSDs have been reported in 18 HOS patients, and 10 have been characterized to date (Baban et al., 2014). Of these, 6 of the mutations lead to the introduction of a stop codon before the TBX5[NID], and one mis-sense mutation maps within the TBX5[NID] and ablates the TBX5-NuRD interaction (Figure 5K). Thus, we find a correlation between the phenotype of patients displaying intracardiac septal defects and the specific genotype of TBX5 patient mutations that causes absent or altered TBX5[NID].

CONCLUSIONS

Our studies demonstrate that a proteomic-based approach coupled with protein modeling provides a powerful predictive strategy in prioritizing patient mutations. This type of approach should be broadly applicable to the analysis of protein complexes in any cell or tissue type. The power of such an approach is underscored by recent high-throughput sequencing studies that have led to the identification of a vast number of somatic mutations in coding regions of potential disease causing genes. Exploiting such large datasets has proved problematic because it has not been possible to determine which mutations are naturally occurring, non-disease associated variants and which are disease-causing. Results presented here demonstrate a proteomic-modeling based strategy that can serve to prioritize mis-sense mutations and, therefore, correlate disease associated phenotypes with specific mutations.

EXPERIMENTAL PROCEDURES

Generation of *Tbx5^{Avi}* mice

The targeting construct to create *Tbx5^{Avi}* was created in collaboration with the UNC Animal Models Core and the UNC BAC Core (Chapel Hill). Briefly, the biotin acceptor peptide (Avi) targeting cassette was inserted in frame to the terminal exon of the *Tbx5* gene. Targeted ES cells were selected and introduced into B6 mouse blastocysts. High-chimera males were mated with B6 females to produce F1 heterozygotes (*Tbx5^{Avi-Neo/+}*). Incorporation of the *Tbx5^{Avi}* construct was confirmed by PCR and Southern analysis (Figure 1B), presence of *Tbx5^{Avi}* mRNA confirmed by RT-PCR (Figure 1D). F1 heterozygotes were crossed with *Sox2-Cre* mice (Hayashi et al., 2002) to remove the neomycin cassette in the F2 generation (Figure 1C). F2 compound heterozygous mice were intercrossed to remove the *Sox2-Cre* allele and establish *Tbx5^{Avi/Avi}* homozygous lines. *Tbx5^{Avi/Avi}* mice were crossed to *Rosa26^{BirA}* mice (JAX), (Driegen et al., 2005). Resultant progeny were intercrossed to obtain compound homozygotes. Genotyping was performed by PCR. Genotyping primers in Supplemental Experimental Procedures. Southern probe was designed to *Tbx5* intron 8. Genomic DNA was digested with EcoRV or BamHI. Animal care and animal experiments were conducted in accordance with the Animal Care Committee at the University of North Carolina, Chapel Hill.

Preparation of cardiac nuclei

Frozen hearts from 4 week old *Tbx5^{Avi/Avi}; Rosa26^{BirA/BirA}* (n=3, 2 replicates) and *Tbx5^{Avi/Avi}* or *Rosa26^{BirA/BirA}* control mice (n=3, 1 replicate per genotype) were homogenized using a mortar and pestle in liquid nitrogen. Nuclei were prepared as previously described (Franklin et al., 2011) and snap frozen in liquid N_2.

Solubilization of protein complexes

Plasmids were transfected into HEK-293 cells, harvested and lysed as previously described (Kaltenbrun et al., 2013). E9.5 hearts were harvested in PBS and lysed as described (Kaltenbrun et al., 2013). Cells and nuclear pellets were resuspended in lysis buffer (200mM K-HEPES pH 7.4, 1.1M KOAc, 20mM $MgCl_2$, 1% Tween-20, 10μM $ZnCl_2$, 10mM $CaCl_2$, 0.5% Triton X-100, 500 mM NaCl, 1X protease inhibitors (Roche), 1X phosphatase inhibitors (Sigma), optimized for TBX5 complexes. Nuclei were homogenized using a Polytron (Kinematica), and processed for affinity purification or immunoprecipitation as described (Kaltenbrun et al., 2013).

Conjugation of magnetic beads and affinity purification

Conjugation of V5 (Invitrogen R960-CUS) or TBX5 (Aviva ARP33403-P050) antibody to magnetic beads (Invitrogen), and immunoisolation of protein complexes was performed as described (Cristea et al., 2005; Kaltenbrun et al., 2013), with following modifications for streptavidin-coated beads. Affinity purifications performed using streptavidin conjugated Dynabeads (Invitrogen) by slow rocking at RT, 30 minutes. The isolated protein complex was eluted for 20 min at 95°C in 40 μL (80mM NaOAc pH 9.0, 95% formamide). Western blots were probed with mouse anti-V5 (1:5000, Invitrogen R960-CUS), rabbit anti-CHD4

(1:500, Active Motif 39289), mouse anti-HDAC2 (1:1000, Abcam Ab51832), rabbit anti-MTA1 (1:5000, Bethyl A300-280A), rabbit anti-TBX5 (1:1000, Aviva ARP33403-P050, using light chain-specific rabbit secondary), goat anti-TBX5 (1:500, Santa Cruz SC-17866), and mouse anti-GFP (1:10000, Clontech 63281).

Mass spectrometry

Isolated protein complexes were analyzed by mass spectrometry as previously described (Kaltenbrun et al., 2013). Briefly, tandem mass spectra were extracted by Proteome Discoverer (Thermo Fisher Scientific), and all MS/MS samples were analyzed with SEQUEST (Thermo Fisher Scientific; version 1.2.0.208), set up to search the human and mouse UniProt-SwissProt protein sequence database, assuming digestion pattern with trypsin. Scaffold (version Scaffold_3_00_06, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide sequences were deemed a match if they could be established at >95.0% probability as specified by the Peptide Prophet algorithm (Keller et al., 2002). In turn, protein identifications were deemed a match if they could be established at >99.0% probability by the ProteinProphet algorithm and have at least two sequenced peptides. Protein identifications and associated spectral counts were exported to Excel for further data processing.

Interaction bioinformatics analysis

Protein identifications were filtered to exclude non-specific associations. Proteins were retained as specific candidates if they (1) had at least two spectral counts in both $Tbx5^{Avi/Avi}$; $Rosa26^{BirA/BirA}$ replicates with at least 4 spectral counts in one replicate, and (2) were uniquely identified or had at least a 2.5-fold spectral count enrichment in the $Tbx5^{Avi/Avi}$; $Rosa26^{BirA/BirA}$ condition versus the controls. Candidates were assigned gene ontology classification using the UniProt GOA annotations from within the Cytoscape (v3.2) platform (Cline et al., 2007). Gene symbols were submitted to the web-based STRING database (Franceschini et al., 2013) for interaction network analysis. Interactions with a combined STRING score of > 0.4 (medium confidence) were retained, exported, and further visualized in Cytoscape.

RNA extraction and RT-PCR

Adult (n=2) hearts were cryolysed (2.5 min/cycle, 30 Hz, 20 cycles) in Trizol (Invitrogen) and RNA extracted using Trizol and purified on RNeasy columns (Qiagen). cDNA synthesis was performed from 0.5–1 µg of RNA using random primers and SuperScript II reverse transcriptase (Invitrogen). Expression levels assessed using GoTaq Green Master Mix (Promega) and Taq polymerase by GeneAmp PCR System (Applied Biosystems). PCR products were analyzed by gel electrophoresis. RT-PCR primer sequences are provided in supplemental Experimental Procedures.

Transcriptional assays

ChIP peak regions from candidate genes were amplified from mouse genomic DNA and subcloned into the PGL3-Promoter Luciferase vector (Promega). Primer sequences in Supplemental Experimental Procedures. HEK-293 cells were transfected with 600ng total

DNA in 12 well plates using FuGene6 (Promega) and harvested 48 hours later for transcription assay which were performed using Dual Reporter Assay System (Promega) according to the manufacturer's protocol in three biological and three technical replicates per assay.

DNA constructs

Generation of the $Tbx5^{Avi}$ targeting construct for embryonic stem cells and generation of chimeric mice was performed by JrGang Chen (UNC BAC Core-Chapel Hill). $Tbx5$-Avi-$V5$ was generated by synthesizing an oligo encoding the Avi tag (5' ggc ctg aac gac atc ttc gag gct cag aaa atc gaa tgg cac gaa 3') and subcloning it into the pcDNA3.1/V5-His TOPO vector (Invitrogen). Full length and truncated versions of mouse $Tbx5$ were then subcloned into pcDNA3.1-Avi-V5. pEF1α-BirA-V5-Neo was generously provided by Stuart Orkin (Harvard University) (Kim et al., 2009). GFP tagged full length, C-terminal, and N-terminal human CHD4 constructs were generously provided by Stephen Jackson (Cambridge) (Polo et al., 2010). Site-directed mutagenesis primers are available in Supplemental Experimental Procedures.

Genetic interaction and histology

$Tbx5^{fl/fl}$ mice were generously provided by Jon and Christine Seidman (Harvard Medical School). The $Tbx5^{+/-}$ mouse has been reported (Bruneau et al., 2001). The $Mta1^{+/-}$ mouse line was generated by Harold Olivey (Indiana University Northwest) and generously provided by Eric Svensson (Novartis Institutes for Biomedial Research). All mice are in a mixed B6/129/SvEv/CD-1 background and experiments were performed according to a protocol reviewed and approved by the Institutional Animal Care and Use Committee of the University of Chicago, in compliance with the USA Public Health Service Policy on Humane Care and Use of Laboratory Animals. E13.5 and E14.5 wild type, $Tbx5^{+/-}$, $Mta1^{+/-}$, and $Tbx5^{+/-};Mta1^{+/-}$ embryos were dissected from timed matings. Embryos fixed in 4% PFA overnight and sent to the Human Tissue Resource Center at the University of Chicago for paraffin embedding, sectioning, and H&E staining. Slides were imaged on a Leica DM2500 microscope and QImaging Retiga 2000R 1394 Color Cooled camera. $Mta1^{-/-}$ null hearts were sectioned and prepared as previously described (Dorr et al., 2015).

RNA-seq

Heart tube dissections were performed as described (Hoffmann et al., 2014). Briefly, E9.5 $Tbx5^{+/+}$ and $Tbx5^{-/-}$ heart tubes were dissected away from the body nearest the myocardial reflections and four heart tubes were pooled to isolate sufficient RNA. Genotypes assigned via PCR. Pools of tissue were mechanically homogenized in TRIzol Reagent (Invitrogen, Carlsbad CA) and RNA was isolated in accordance with the manufacturer's instructions. RNA resuspended in water was further purified with 500 μl of 1-butanol (4 times) and 500 μl diethyl ether (twice) (Krebs et al., 2009). Isolated RNA underwent library preparation and 51-bp single-ended RNA sequencing at the Genomics Core Facility at The University of Chicago. Briefly, mRNA was selected using an oligo-dT pulldown, and barcoded libraries were prepared according to Illumina's instructions accompanying the TruSeq RNA Sample prep kit v2 (Part# RS-122–2001). Library fragments of ~275 bps (insert plus adaptor and PCR primer sequences) were quantitated using the Agilent Bio-analyzer 2100 and pooled in

equimolar amounts. The pooled libraries were sequenced on the HiSeq2500 in Rapid Run Mode following the manufacturer's protocols (Invitrogen, 2013).

The raw data output from the Illumina Genome Analyzer was in phredscaled base-64 fastq format, representing the sequence and quality scores for each read. The single-ended 51-bp reads were filtered based on their quality scores using FastQC toolkit (Andrews S, 2010), selecting for the 37-bp reads on the right with a minimum quality score of 35 in 75% of the bases of the read. The reads passing the filter were aligned to the mouse genome mm10 assembly by Tophat (version 2.0.6) using default parameters (Kim et al., 2013). The output SAM files were filtered for alignment sorted, and converted to BAM files using Samtools (Li et al., 2009). Hit counts were estimated and normalized by both each transcript's length and the total yield of the sequencer using Cufflinks (version 2.1.1, default parameters and "-p 2") (Trapnell et al., 2012). Among all Cufflinks reported transcript abundances, the Fragments Per Kilobase of exon per Million fragments mapped (FPKM), values larger than 0.125 in all samples were kept for the following statistical analysis. This filter guaranteed the median correlation coefficient >0.96 between any two biological replicates while kept ~65% of all transcripts, including 11789 coding genes (Ensembl annotation, mm10, 2014 April version). Differential expression between mutant samples (n=2) and wild types (n=4) was tested using Cuffdiff (v2.1.1) (Trapnell et al., 2012) Although Cuffdiff lacks power for designs with fewer than three replicates per group, the low variability of the two biological replicates in all cases (r>0.96) guarantees the statistic power of a fold change to reveal biologically meaningful results. All genes with an adjusted p-value less than 0.05 and the absolute fold change larger than 1.57 were called significant.

Large Scale bioinformatics analysis of TBX5 binding sites

Differentially expressed genes were determined by q-value (<0.05) and status ("OK"). Those genes that did not meet the status threshold were given a non-significant q-value (qval=1). TBX5 ChIP-Seq FASTQ data was downloaded from GEO (GSE21529). Tags were filtered using Tag-dust (Lassmann et al., 2009) and aligned to the mouse genome (mm9) using bowtie (Langmead et al., 2009). Peaks were called using MACS2 (Zhang et al., 2008) with default parameters. Peaks were associated with differentially expressed genes and differential motif enrichment was determined using BETA (Wang et al., 2013) with the following options –-pn 40000 and –-df 0.05 –-da 1. Because BETA uses a curated database that lacks the TBX5 motif, TBX5 peaks associated with differentially regulated genes were also analyzed for motif enrichment using HOMER (Heinz et al., 2010). Genes associated with TBX5 regulation were then analyzed for pathway enrichment using DAVID (Huang da et al., 2009). The top 10 (by p-value) KEGG pathways are shown for both activated and repressed genes.

Chromatin IP

HEK-293 cells were transfected with the *Kctd16-Luciferase* construct used in transcriptional assays, as well as wild type TBX5 or TBX5[S261C] (20 μg total DNA) using FuGene 6 (Promega). Cells were fixed in 1% PFA for 10 minutes at RT. 10 million cells were used per ChIP. Cells were lysed in ChIP Buffer (50mM Tris pH 7.5, 140 mM NaCl, 1mM EDTA, 1 mM EGTA, 0.1% sodium deoxycholate, 0.1% SDS), sonicated using a Branson 450d, 12

cycles, 30s/cycle (1s on, 0.5s off), 30% amplitude), and spun at top speed to isolate chromatin. Triton X-100 was added to 1%, and the chromatin was added to CHD4-conjugated Protein G beads (Life Technologies) at 4°C O/N. ChIP was resumed using ChIP Protocol (Agilent). ChIP primer sequence is available in Supplemental Experimental Procedures.

TBX5 NID Modelling

The mouse TBX5 sequence was submitted to the DisProt PONDR-FIT disorder predictor in order to identify disordered regions in the protein. Both the entire protein and just the NID region were submitted to the fold recognition server HHpred (http://toolkit.tuebingen.mpg.de/hhpred) to determine if any suitable templates were available for structural modeling. A small coil-to-helical region involving residues 255-264 was predicted by HHpred. This region of TBX5 was modeled based on the correspondingly small region of the structure of the CRISPR-associated protein [PDB ID 3VZI].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010 Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Baban A, Pitto L, Pulignani S, Cresci M, Mariani L, Gambacciani C, Digilio MC, Pongiglione G, Albanese S. Holt-Oram syndrome with intermediate atrioventricular canal defect, and aortic coarctation: Functional characterization of a de novo TBX5 mutation. Am J Med Genet A. 2014

Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soults J, Grayzel D, Kroumpouzou E, Traill TA, Leblanc-Straceski J, et al. Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome. Nat Genet. 1997; 15:30–35. [PubMed: 8988165]

Basson CT, Cowley GS, Solomon SD, Weissman B, Poznanski AK, Traill TA, Seidman JG, Seidman CE. The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). N Engl J Med. 1994; 330:885–891. [PubMed: 8114858]

Basson CT, Huang T, Lin RC, Bachinsky DR, Weremowicz S, Vaglio A, Bruzzone R, Quadrelli R, Lerone M, Romeo G, et al. Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. Proc Natl Acad Sci U S A. 1999; 96:2919–2924. [PubMed: 10077612]

Bayer EA, Wilchek M. The use of the avidin-biotin complex as a tool in molecular biology. Methods Biochem Anal. 1980; 26:1–45. [PubMed: 7392958]

Benson DW, Basson CT, MacRae CA. New understandings in the genetics of congenital heart disease. Curr Opin Pediatr. 1996; 8:505–511. [PubMed: 8946132]

Berk AJ. Activation of RNA polymerase II transcription. Curr Opin Cell Biol. 1999; 11:330–335. [PubMed: 10395559]

Brown DD, Martz SN, Binder O, Goetz SC, Price BM, Smith JC, Conlon FL. Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. Development. 2005; 132:553–563. [PubMed: 15634698]

Bruneau BG, Logan M, Davis N, Levi T, Tabin CJ, Seidman JG, Seidman CE. Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. Dev Biol. 1999; 211:100–108. [PubMed: 10373308]

Bruneau BG, Nemer G, Schmitt JP, Charron F, Robitaille L, Caron S, Conner DA, Gessler M, Nemer M, Seidman CE, et al. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. Cell. 2001; 106:709–721. [PubMed: 11572777]

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, et al. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc. 2007; 2:2366–2382. [PubMed: 17947979]

Costantini JL, Cheung SM, Hou S, Li H, Kung SK, Johnston JB, Wilkins JA, Gibson SB, Marshall AJ. TAPP2 links phosphoinositide 3-kinase signaling to B-cell adhesion through interaction with the cytoskeletal protein utrophin: expression of a novel cell adhesion-promoting complex in B-cell leukemia. Blood. 2009; 114:4703–4712. [PubMed: 19786618]

Cristea IM, Williams R, Chait BT, Rout MP. Fluorescent proteins as proteomic probes. Mol Cell Proteomics. 2005; 4:1933–1941. [PubMed: 16155292]

Cross SJ, Ching YH, Li QY, Armstrong-Buisseret L, Spranger S, Lyonnet S, Bonnet D, Penttinen M, Jonveaux P, Leheup B, et al. The mutation spectrum in Holt-Oram syndrome. J Med Genet. 2000; 37:785–787. [PubMed: 11183182]

Denslow SA, Wade PA. The human Mi-2/NuRD complex and gene regulation. Oncogene. 2007; 26:5433–5438. [PubMed: 17694084]

Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, Magen A, Canidio E, Pagani M, Peluso I, et al. A high-resolution anatomical atlas of the transcriptome in the mouse embryo. PLoS Biol. 2011; 9:e1000582. [PubMed: 21267068]

Dolk H, Loane M, Garne E. The prevalence of congenital anomalies in Europe. Adv Exp Med Biol. 2010; 686:349–364. [PubMed: 20824455]

Dorr KM, Amin NM, Kuchenbrod LM, Labiner H, Charpentier MS, Pevny LH, Wessels A, Conlon FL. Casz1 is required for cardiomyocyte G1-to-S phase progression during mammalian cardiac development. Development. 2015; 142:2037–2047. [PubMed: 25953344]

Driegen S, Ferreira R, van Zon A, Strouboulis J, Jaegle M, Grosveld F, Philipsen S, Meijer D. A generic tool for biotinylation of tagged proteins in transgenic mice. Transgenic Res. 2005; 14:477–482. [PubMed: 16201414]

Evans PM, Liu C. Roles of Krupel-like factor 4 in normal homeostasis, cancer and stem cells. Acta Biochim Biophys Sin (Shanghai). 2008; 40:554–564. [PubMed: 18604447]

Fan C, Duhagon MA, Oberti C, Chen S, Hiroi Y, Komuro I, Duhagon PI, Canessa R, Wang Q. Novel TBX5 mutations and molecular mechanism for Holt-Oram syndrome. J Med Genet. 2003a; 40:e29. [PubMed: 12624158]

Fan C, Liu M, Wang Q. Functional analysis of TBX5 missense mutations associated with Holt-Oram syndrome. J Biol Chem. 2003b; 278:8780–8785. [PubMed: 12499378]

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013; 41:D808–D815. [PubMed: 23203871]

Franklin S, Zhang MJ, Chen H, Paulsson AK, Mitchell-Jordan SA, Li Y, Ping P, Vondriska TM. Specialized compartments of cardiac nuclei exhibit distinct proteomic anatomy. Mol Cell Proteomics. 2011; 10 M110 000703.

Garg V, Kathiriya IS, Barnes R, Schluterman MK, King IN, Butler CA, Rothrock CR, Eapen RS, Hirayama-Yamada K, Joo K, et al. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature. 2003; 424:443–447. [PubMed: 12845333]

Ghosh TK, Song FF, Packham EA, Buxton S, Robinson TE, Ronksley J, Self T, Bonser AJ, Brook JD. Physical interaction between TBX5 and MEF2C is required for early heart development. Mol Cell Biol. 2009; 29:2205–2218. [PubMed: 19204083]

Gittenberger-de Groot AC, Calkoen EE, Poelmann RE, Bartelings MM, Jongbloed MR. Morphogenesis and molecular considerations on congenital cardiac septal defects. Ann Med. 2014; 46:640–652. [PubMed: 25307363]

Goetz SC, Brown DD, Conlon FL. TBX5 is required for embryonic cardiac cell cycle progression. Development. 2006; 133:2575–2584. [PubMed: 16728474]

Hatcher CJ, Kim MS, Mah CS, Goldstein MM, Wong B, Mikawa T, Basson CT. TBX5 transcription factor regulates cell proliferation during cardiogenesis. Dev Biol. 2001; 230:177–188. [PubMed: 11161571]

Hayashi S, Lewis P, Pevny L, McMahon AP. Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. Mech Dev. 2002; 119(Suppl 1):S97–S101. [PubMed: 14516668]

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. [PubMed: 21415370]

Heinritz W, Shou L, Moschik A, Froster UG. The human TBX5 gene mutation database. Hum Mutat. 2005; 26:397. [PubMed: 16134140]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010; 38:576–589. [PubMed: 20513432]

Heron M, Tejada-Vera B. Deaths: leading causes for 2005. Natl Vital Stat Rep. 2009; 58:1–97. [PubMed: 20361522]

Hiroi Y, Kudoh S, Monzen K, Ikeda Y, Yazaki Y, Nagai R, Komuro I. Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. Nat Genet. 2001; 28:276–280. [PubMed: 11431700]

Hoffmann AD, Yang XH, Burnicka-Turek O, Bosman JD, Ren X, Steimle JD, Vokes SA, McMahon AP, Kalinichenko VV, Moskowitz IP. Foxf genes integrate tbx5 and hedgehog pathways in the second heart field for cardiac septation. PLoS Genet. 2014; 10:e1004604. [PubMed: 25356765]

Holt M, Oram S. Familial heart disease with skeletal malformations. Br Heart J. 1960; 22:236–242. [PubMed: 14402857]

Hu S, Li L, Yeh S, Cui Y, Li X, Chang HC, Jin J, Chang C. Infiltrating T cells promote prostate cancer metastasis via modulation of FGF11-->miRNA-541-->androgen receptor (AR)-->MMP9 signaling. Mol Oncol. 2015; 9:44–57. [PubMed: 25135278]

Huang CC, Tu SH, Lien HH, Jeng JY, Huang CS, Huang CJ, Lai LC, Chuang EY. Concurrent gene signatures for han chinese breast cancers. PLoS One. 2013; 8:e76421. [PubMed: 24098497]

Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4:44–57. [PubMed: 19131956]

Ieda M, Fu JD, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, Srivastava D. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell. 2010; 142:375–386. [PubMed: 20691899]

Kaltenbrun E, Greco TM, Slagle CE, Kennedy LM, Li T, Cristea IM, Conlon FL. A Gro/TLE-NuRD corepressor complex facilitates Tbx20-dependent transcriptional repression. J Proteome Res. 2013; 12:5395–5409. [PubMed: 24024827]

Kehle J, Beuchle D, Treuheit S, Christen B, Kennison JA, Bienz M, Muller J. dMi-2, a hunchback-interacting protein that functions in polycomb repression. Science. 1998; 282:1897–1900. [PubMed: 9836641]

Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 2002; 74:5383–5392. [PubMed: 12403597]

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14:R36. [PubMed: 23618408]

Kim J, Cantor AB, Orkin SH, Wang J. Use of in vivo biotinylation to study protein-protein and protein-DNA interactions in mouse embryonic stem cells. Nat Protoc. 2009; 4:506–517. [PubMed: 19325547]

Kimura R, Kasamatsu A, Koyama T, Fukumoto C, Kouzu Y, Higo M, Endo-Sakamoto Y, Ogawara K, Shiiba M, Tanzawa H, et al. Glutamate acid decarboxylase 1 promotes metastasis of human oral cancer by beta-catenin translocation and MMP7 activation. BMC Cancer. 2013; 13:555. [PubMed: 24261884]

Krebs S, Fischaleck M, Blum H. A simple and loss-free method to remove TRIzol contaminations from minute RNA samples. Anal Biochem. 2009; 387:136–138. [PubMed: 19146820]

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10:R25. [PubMed: 19261174]

Lassmann T, Hayashizaki Y, Daub CO. TagDust--a program to eliminate artifacts from next generation sequencing data. Bioinformatics. 2009; 25:2839–2840. [PubMed: 19737799]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]

Li QY, Newbury-Ecob RA, Terrett JA, Wilson DI, Curtis AR, Yi CH, Gebuhr T, Bullen PJ, Robson SC, Strachan T, et al. Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. Nat Genet. 1997; 15:21–29. [PubMed: 8988164]

Liu Z, Yang X, Li Z, McMahon C, Sizer C, Barenboim-Stapleton L, Bliskovsky V, Mock B, Ried T, London WB, et al. CASZ1, a candidate tumor-suppressor gene, suppresses neuroblastoma tumor growth through reprogramming gene expression. Cell Death Differ. 2011; 18:1174–1183. [PubMed: 21252912]

Maine GN, Li H, Zaidi IW, Basrur V, Elenitoba-Johnson KS, Burstein E. A bimolecular affinity purification method under denaturing conditions for rapid isolation of a ubiquitinated protein for mass spectrometry analysis. Nat Protoc. 2010; 5:1447–1459. [PubMed: 20671728]

McDermott DA, Bressan MC, He J, Lee JS, Aftimos S, Brueckner M, Gilbert F, Graham GE, Hannibal MC, Innis JW, et al. TBX5 genetic testing validates strict clinical criteria for Holt-Oram syndrome. Pediatr Res. 2005; 58:981–986. [PubMed: 16183809]

Metz M, Gassmann M, Fakler B, Schaeren-Wiemers N, Bettler B. Distribution of the auxiliary GABAB receptor subunits KCTD8, 12, 12b, and 16 in the mouse brain. J Comp Neurol. 2011; 519:1435–1454. [PubMed: 21452234]

Mori AD, Bruneau BG. TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. Curr Opin Cardiol. 2004; 19:211–215. [PubMed: 15096952]

Morris, C. Epidemiology of congenital heart disease. In: DJ; Crawford, MH.; Paulus, WJ., editors. Cardiology. Philadelphia: Mosby; 2010. p. 1381-1389.

Moskowitz IP, Kim JB, Moore ML, Wolf CM, Peterson MA, Shendure J, Nobrega MA, Yokota Y, Berul C, Izumo S, et al. A molecular pathway including Id2, Tbx5, and Nkx2-5 required for cardiac conduction system development. Cell. 2007; 129:1365–1376. [PubMed: 17604724]

Moskowitz IP, Pizard A, Patel VV, Bruneau BG, Kim JB, Kupershmidt S, Roden D, Berul CI, Seidman CE, Seidman JG. The T-Box transcription factor Tbx5 is required for the patterning and maturation of the murine cardiac conduction system. Development. 2004; 131:4107–4116. [PubMed: 15289437]

Polo SE, Kaidi A, Baskcomb L, Galanty Y, Jackson SP. Regulation of DNA-damage responses and cell-cycle progression by the chromatin remodelling factor CHD4. EMBO J. 2010; 29:3130–3139. [PubMed: 20693977]

Qian L, Huang Y, Spencer CI, Foley A, Vedantham V, Liu L, Conway SJ, Fu JD, Srivastava D. In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. Nature. 2012; 485:593–598. [PubMed: 22522929]

Roesli C, Neri D, Rybak JN. In vivo protein biotinylation and sample preparation for the proteomic identification of organ- and disease-specific antigens accessible from the vasculature. Nat Protoc. 2006; 1:192–199. [PubMed: 17406232]

Saunders A, Werner J, Andrulis ED, Nakayama T, Hirose S, Reinberg D, Lis JT. Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. Science. 2003; 301:1094–1096. [PubMed: 12934007]

Sedmera D, Thompson RP. Myocyte proliferation in the developing heart. Dev Dyn. 2011; 240:1322–1334. [PubMed: 21538685]

Smallwood PM, Munoz-Sanjuan I, Tong P, Macke JP, Hendry SH, Gilbert DJ, Copeland NG, Jenkins NA, Nathans J. Fibroblast growth factor (FGF) homologous factors: new members of the FGF family implicated in nervous system development. Proc Natl Acad Sci U S A. 1996; 93:9850–9857. [PubMed: 8790420]

Sopher BL, Koszdin KL, McClain ME, Myrick SB, Martinez RA, Smith AC, La Spada AR. Genomic organization, chromosome location, and expression analysis of mouse beta-synuclein, a candidate for involvement in neurodegeneration. Cytogenet Cell Genet. 2001; 93:117–123. [PubMed: 11474193]

Stirnimann CU, Ptchelkine D, Grimm C, Muller CW. Structural basis of TBX5-DNA recognition: the T-box domain in its DNA-bound and - unbound form. J Mol Biol. 2010; 400:71–81. [PubMed: 20450920]

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012; 7:562–578. [PubMed: 22383036]

Trifonov S, Yamashita Y, Kase M, Maruyama M, Sugimoto T. Glutamic acid decarboxylase 1 alternative splicing isoforms: characterization, expression and quantification in the mouse brain. BMC Neurosci. 2014; 15:114. [PubMed: 25322942]

van Werven FJ, Timmers HT. The use of biotin tagging in Saccharomyces cerevisiae improves the sensitivity of chromatin immunoprecipitation. Nucleic Acids Res. 2006; 34:e33. [PubMed: 16500888]

Wade PA, Jones PL, Vermaak D, Wolffe AP. A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. Curr Biol. 1998; 8:843–846. [PubMed: 9663395]

Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH. A protein interaction network for pluripotency of embryonic stem cells. Nature. 2006; 444:364–368. [PubMed: 17093407]

Wang C, Cao D, Wang Q, Wang DZ. Synergistic activation of cardiac genes by myocardin and tbx5. PLoS One. 2011; 6:e24242. [PubMed: 21897873]

Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, Tang Q, Meyer CA, Zhang Y, Liu XS. Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nat Protoc. 2013; 8:2502–2515. [PubMed: 24263090]

Wang W, Kim HJ, Lee JH, Wong V, Sihn CR, Lv P, Perez Flores MC, Mousavi-Nik A, Doyle KJ, Xu Y, et al. Functional significance of K+ channel beta-subunit KCNE3 in auditory neurons. J Biol Chem. 2014; 289:16802–16813. [PubMed: 24727472]

Xue Y, Wong J, Moreno GT, Young MK, Cote J, Wang W. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol Cell. 1998; 2:851–861. [PubMed: 9885572]

Ying CY, Dominguez-Sola D, Fabi M, Lorenz IC, Hussein S, Bansal M, Califano A, Pasqualucci L, Basso K, Dalla-Favera R. MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. Nat Immunol. 2013; 14:1084–1092. [PubMed: 23974956]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]

- Isolation and characterization of the endogenous cardiac TBX5 interactome.

- TBX5 interacts biochemically and genetically with the NuRD repressor complex.

- TBX5 represses inappropriate gene programs incompatible with cardiac development.

- Human TBX5 mutations alter the NuRD interaction and correlate with septal defects.

Waldron et al. isolate the cardiac interactome of TBX5, revealing that this transcription factor, implicated in congenital heart disease (CHD) and thought to be an activator, interacts with NuRD to represses inappropriate gene expression in the heart. CHD mutations disrupting the TBX5-NuRD interaction cause derepression of some of these genes.
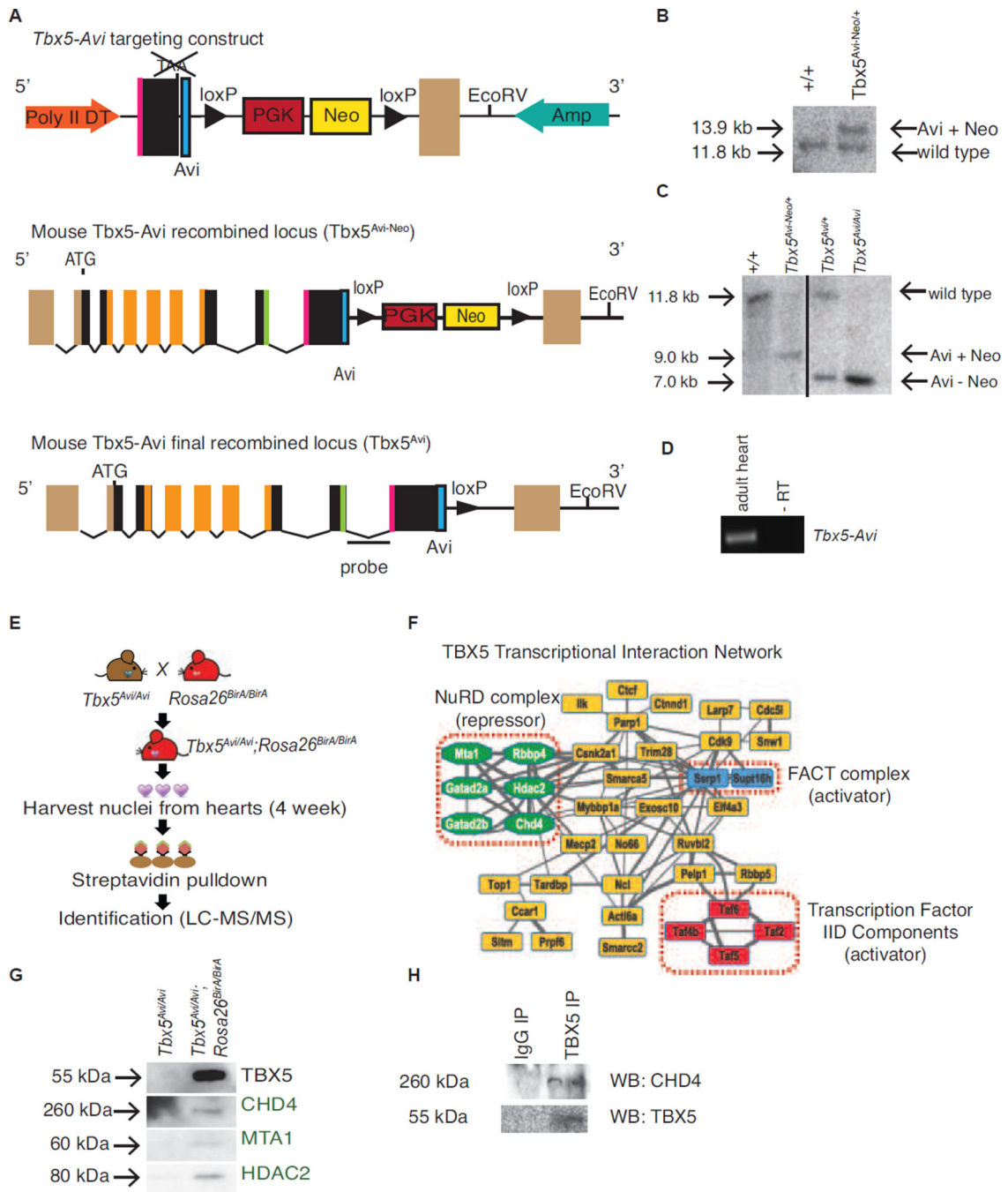
Figure 1. TBX5 interacts with the NuRD complex

(**A**) (Top) The *Tbx5-Avi* targeting cassette contains the biotin acceptor peptide (Avi) fused in frame to the terminal coding exon of *Tbx5*. Neo-neomycin resistance gene. PGK-phosphoglycerate kinase-1 promoter. Amp-Ampicillan resistance gene. Exons colored brown represent untranslated exons (5' and 3' UTR). Exons colored orange represent the T-box DNA binding domain. Exons colored green represent the nuclear localization signal. Exons colored pink represent the activation domain. (Middle) Schematic of the *Tbx5* recombined locus (*Tbx5^Avi-Neo^*) upon homologous recombination. (Bottom) Schematic of

the *Tbx5* locus upon introduction of *Cre* (*Tbx5^Avi^*). (**B**) Southern blot of positive embryonic stem cell clone confirms homologous recombination. (**C**) Southern blot of F2 mice confirms removal of Neo cassette upon introduction of *Cre*. Samples were run on the same gel but were noncontiguous. (**D**) RT-PCR analysis of *Tbx5^Avi/+^* adult hearts demonstrates presence of *Tbx5^Avi^* mRNA. (**E**) Overview of the isolation and characterization of the endogenous cardiac TBX5 interactome. (**F**) TBX5 transcriptional interaction network. Network nodes, labeled with mouse gene symbols, are candidate direct or indirect TBX5 protein interactions identified from affinity enrichment of biotinylated TBX5. Network edges represent known and/or predicted functional interactions in the STRING database. Edge thickness reflects the combined STRING evidence score for each binary relationship. Thicker edges represent increased interaction evidence. Selected transcriptional complexes within the network are highlighted in red dashed boxes. (**G**) Affinity isolation of endogenous TBX5^Avi^ from 4 week cardiac nuclei confirms interaction of TBX5 with endogenous components of the NuRD complex. (**H**) Affinity isolation of TBX5 from E9.5 hearts further confirms interaction of TBX5 with the NuRD component CHD4 in the embryonic heart.
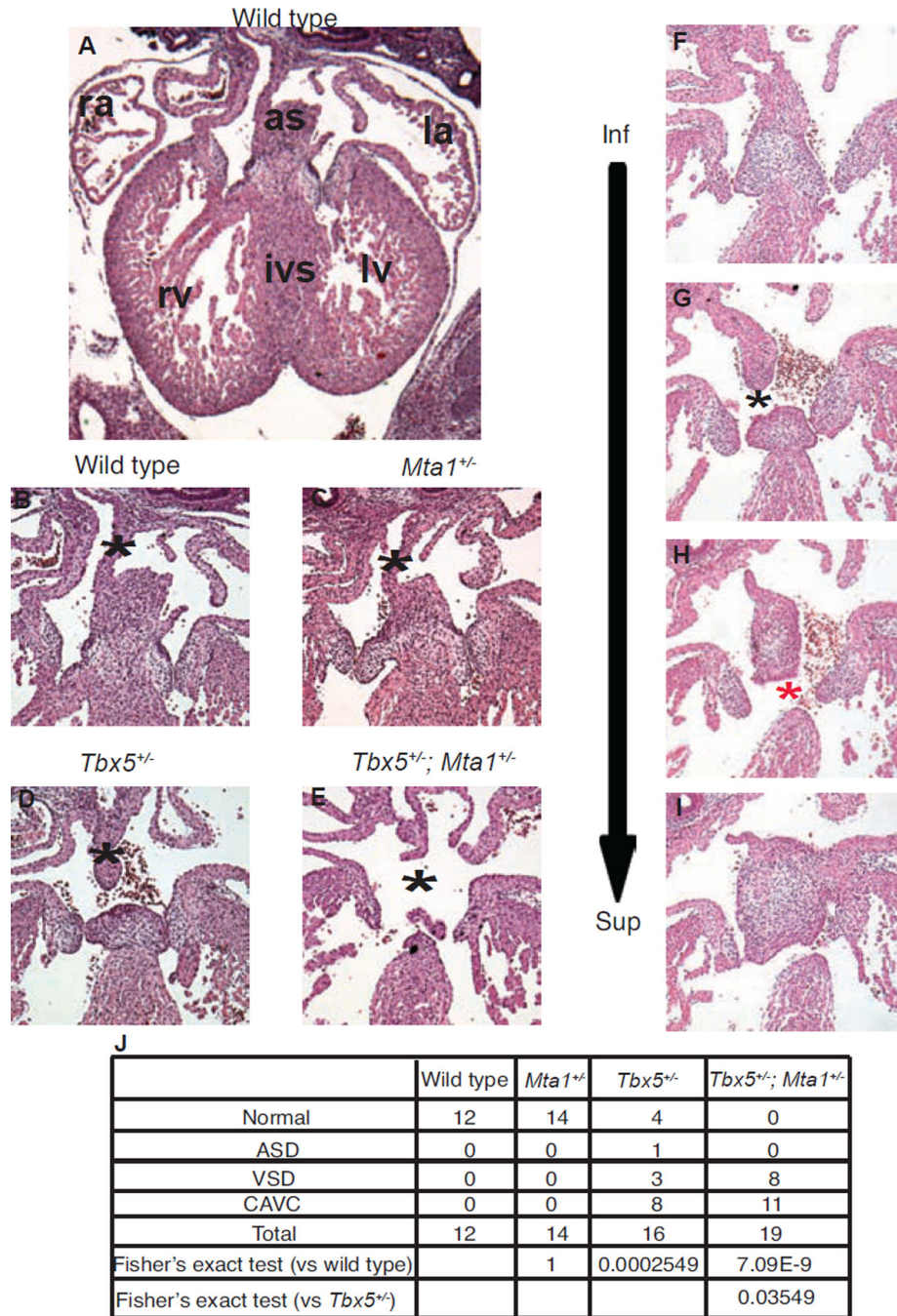
Figure 2. TBX5 and the NuRD complex genetically interact

Histology (Hematoxylin and Eosin) of embryonic hearts in transverse section at E13.5 (**A**) Low magnification of a wild type heart. High magnification of (**B**) wild type, (**C**) $Mta1^{+/-}$, (**D**) $Tbx5^{+/-}$ and, (**E**) $Tbx5^{+/-}$; $Mta1^{+/}$ hearts. $Tbx5^{+/-}$; $Mta1^{+/}$ hearts exhibited cardiac septal defects including ASD, VSD, and CCAVC (asterisk). (**F–I**) Histology at E13.5 through a single $Tbx5^{+/-}$; $Mta1^{+/-}$ heart from inferior to superior revealing ASD (**G**) (black asterisk) and VSD (red asterisk) (**H, I**) components of CCAVC. (**J**) Quantification of cardiac defects observed in each genotype. Fisher's exact test showed that $Tbx5^{+/-}$; $Mta1^{+/-}$ mice

demonstrated significantly more septal defects than either wild type or *Tbx5*$^{+/-}$ mice. as = atrial septum, ivs = intervectricular septum, la = left atrium, lv = left ventricle, ra = right atrium, rv = right ventricle.
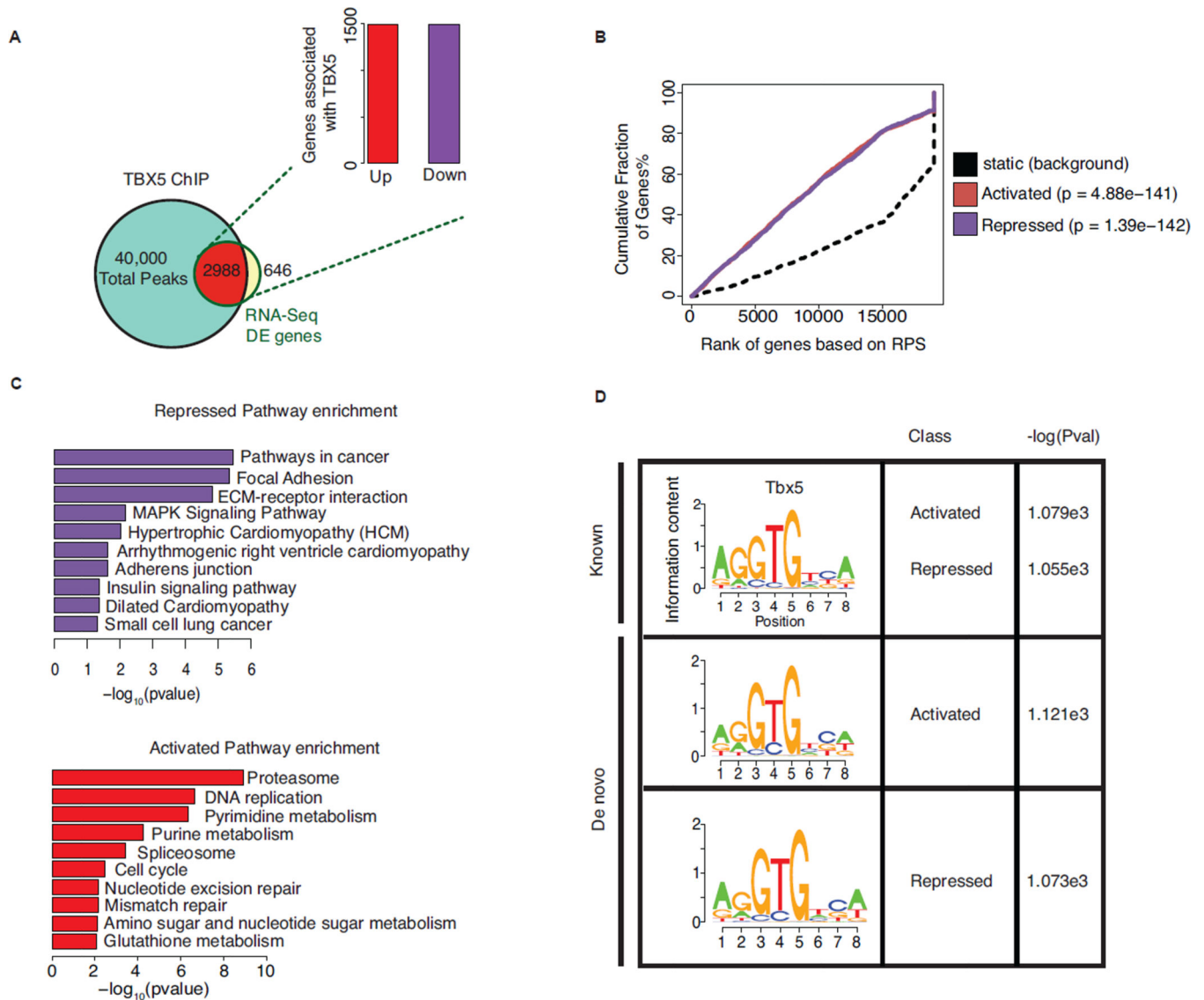
Figure 3. Analysis of TBX5 binding motifs in activated vs repressed genes

(**A**) Schematic overlay between TBX5 peaks (He et al., 2011) and differentially expressed genes between wild-type and *Tbx5* null heart tissue. 2988 genes are differentially expressed and associated with a TBX5 peak. Inset is number of up and downregulated genes. (**B**) Activating/Repressive function prediction of TBX5 peaks. Red and purple lines represent up and down-regulated genes. Black dashed line represents background (static genes). Genes are ranked from high to low based on their regulatory potential score (RPS) determined by BETA. P values represent significance of up or down-regulated group relative to background (static genes) by Kolmogorov-Smirnov test. (**C**) Kegg Pathway enrichment of down-regulated genes associated with TBX5 binding ranked by $-\log_{10}$(pvalue). (**D**) TBX5 consensus motif (top) and top motif from de novo identification (bottom) and its presence at ChIP-seq peaks associated with activated or repressed genes.

A

| Gene	Log2 fold change	ChIP peak location	ChIP peak value(s)	Size of element
Sncb	5.01	intron 4	18	321 bp
Nxph4	4.93	5' UTR (5.7 kb from TSS)	41	490 bp
Mef2b	4.61	5' UTR (7.4 kb from TSS)	27	347 bp
Kctd16	4.53	5' UTR (1.4 kb from TSS)	17	426 bp
Gad1	4.19	intron 1	17	276 bp
Kcne3	3.63	5' UTR (100bp from TSS)	23	350 bp
Fgf11	3.49	intron 1	45	1,200 bp
Col20a1	1.95	5' UTR (2.7 kb from TSS)	17, 35	1,100 bp
Klf4	1.74	exon 3	57, 56	3,291 bp
Fgf8	1.71	5' UTR (2.8 kb from TSS)	80	384 bp
Fbxl20	1.37	5' UTR/ exon 1/ intron 1	25, 56, 30	1,600 bp
Casz1-1	1.26	exon 1/ intron 1	60, 29, 43	3,200 bp
Casz1-2	1.26	intron 1	32, 22, 18	2,400 bp
Kcnn2	1.23	exon 1	13	140 bp
Plekha2	1.21	5' UTR (1.3 kb from TSS)	49, 33	1,700 bp
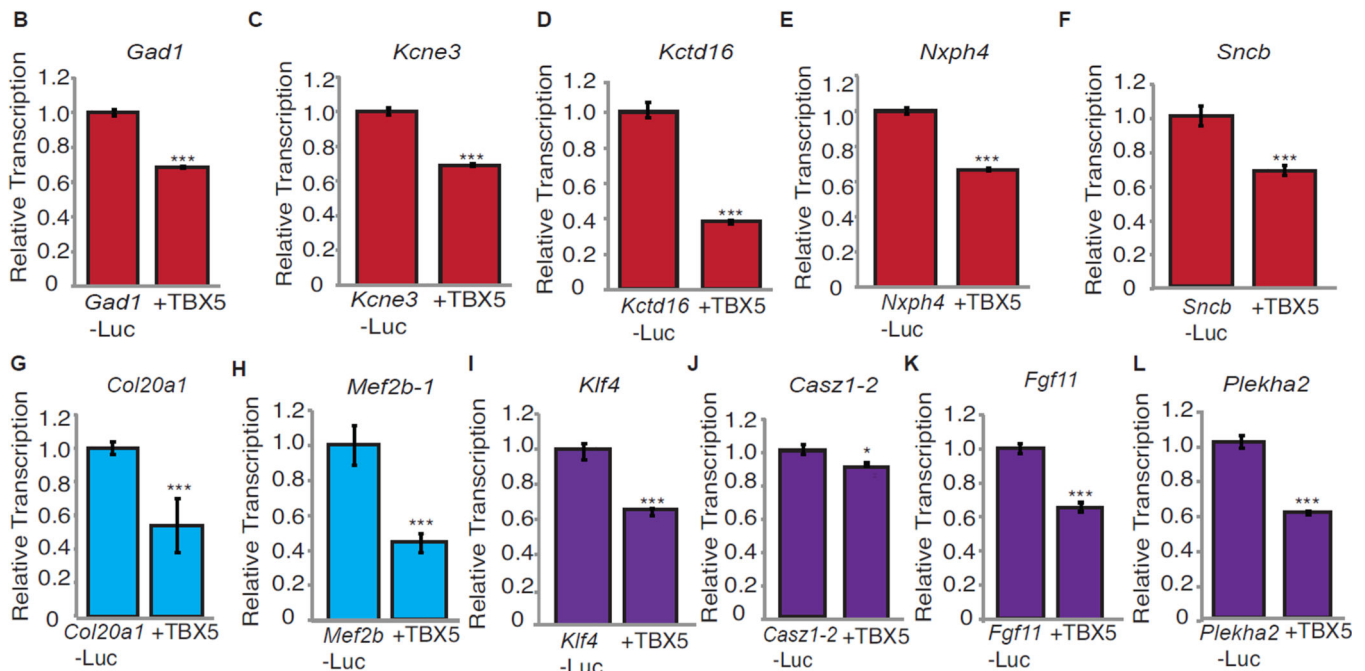
Figure 4. TBX5 functions to represses misexpression of genes in cardiac tissue
(**A**) Summary of fold change (RNA-seq) and ChIP-seq peak values of the 15 top rank ordered genes. (**B–L**) Gene reporter elements cloned from potential TBX5 targets in the presence or absence of TBX5. Transcriptional assays of target genes show TBX5 acts as a transcriptional repressor in 11 of 15 elements tested (additional data in Figure S4). Genes defined as neural are shown in red, expressed in cancer as blue, or both as purple. Graphs are plotted as the mean luciferase value ± SEM. Student's two tailed t-test was used to perform statistical analysis. * p< 0.05 *** p < 0.001.
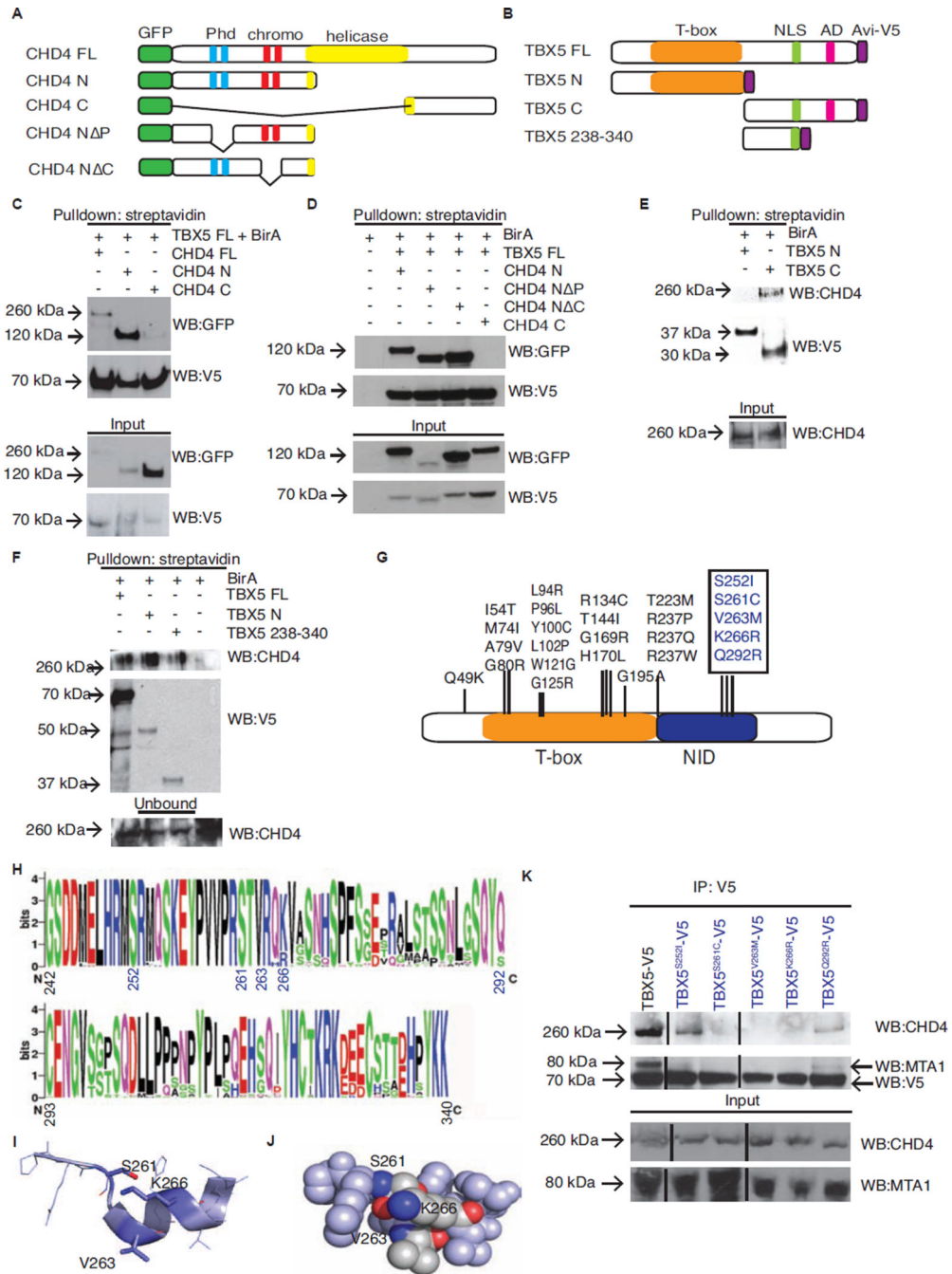
Figure 5. Congenital heart disease associated mutations of TBX5 disrupt TBX5-NuRD complex interaction and activity

(**A**) Schematic of CHD4 deletion constructs. Phd= Plant homeo domain. CHD4 FL = aa 1–1937, CHD4 N = aa 1–758, CHD4 C = aa 1183–1937, TBX5 FL = aa 1-518, TBX5 N = aa 1–237, TBX5 C = aa 238–518. (**B**) Schematic of TBX5 deletion constructs. NLS = Nuclear Localization Signal. AD = Activation domain. (**C–D**) Affinity isolation of full-length TBX5-Avi-V5 with GFP-tagged CHD4 deletions reveals that CHD4 interacts with TBX5 via sequences in the N-terminus of CHD4, but that the Phd and chromo domains are not

individually required for the TBX5-CHD4 interaction. (**E–F**) Affinity isolation of CHD4 with a TBX5-Avi-V5 deletion series demonstrates TBX5 sequences aa 238–340 are both necessary and sufficient for interaction with CHD4. This region has been termed the NuRD interaction domain (NID). (**G**) Schematic of human TBX5 protein showing location of a subset of CHD associated mis-sense mutations. NID= NuRD interacting domain (aa 238–340). (**H**) Sequence alignment of the TBX5$^{\text{NID}}$ from 48 TBX5 ortholgues. Height of letters is relative conservation at that residue. (**I**) Structural prediction of the TBX5$^{\text{NID}}$ shows the TBX5$^{\text{NID}}$ is comprised of a coil region followed an α-helix. (**J**) CHD associated mis-sense mutations align along a single surface of the α-helix and are predicted to disrupt protein-protein interactions (representative residues TBX5$^{\text{S261}}$, TBX5$^{\text{V263}}$, and TBX5$^{\text{K266}}$ shown in dark blue). (**K**) Immunoisolation of V5-tagged CHD associated mis-sense mutations of TBX5 probed for CHD4 and MTA1.
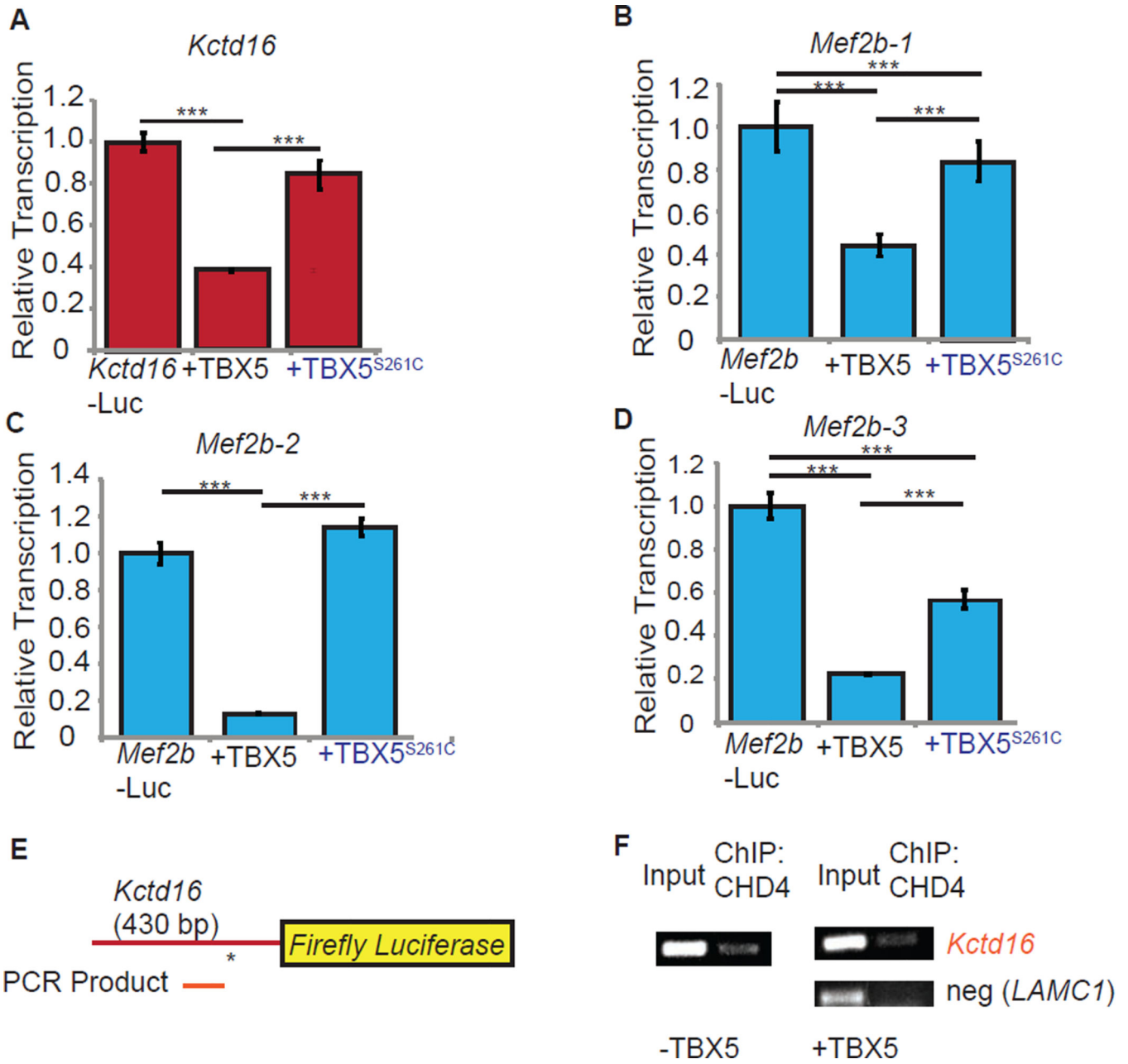
Figure 6. Congenital heart disease associated mutations of TBX5 disrupt TBX5-NuRD complex activity

(**A–D**) Transcription of the regulatory regions of TBX5 target genes in response to wild type TBX5 (second column) or TBX5^{S261C} (third column). Additional gene targets shown in Figure S7. n= 9 per transcriptional assay. Graphs are plotted as the mean luciferase value ± SEM. Student's two tailed t-test was used to perform statistical analysis. *** $p < 0.001$. (**E**) Schematic of ChIP primer location on *Kctd16* element. The T-box binding element (asterisk) is noted. (**F**) ChIP of CHD4 from cells co-transfected with *Kctd16-Luciferase* in the presence of absence of TBX5.
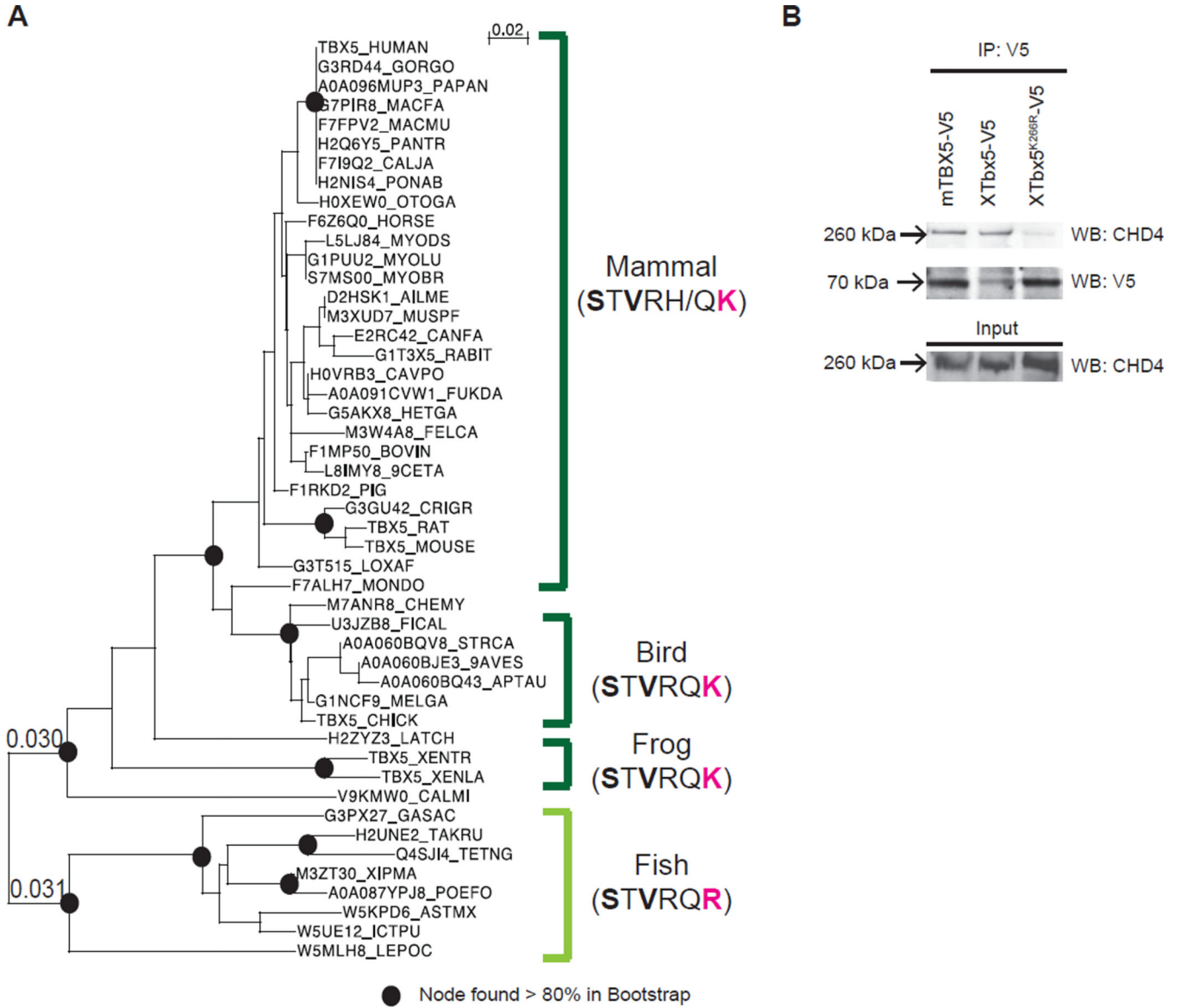
Figure 7. The TBX5-NuRD interaction evolved concurrently with cardiac septation
(A) Phylogenetic comparison of TBX5 orthologues with boot-strap values. Dark green brackets represent animals with a septated heart. Light green brackets represent animals without a septated heart. Sequence of core amino acid residues required for the TBX5-NuRD interaction are shown to the right. Amino acid residues mutated in TBX5-associated CHD are bolded, and a K->R substitution from fish to frog is highlighted in pink. (**B**) Immunoisolation of V5-tagged TBX5 orthologs probed for CHD4.