

TBX5 drives *Aldh1a2* expression to regulate a RA-Hedgehog-Wnt gene regulatory network coordinating cardiopulmonary development

Scott A. Rankin¹, Jeffrey D. Steimle^{3,4,5,6}, Xinan H. Yang^{3,4,5}, Ariel B. Rydeen^{3,4,5}, Kunal Agarwal¹, Praneet Chaturvedi¹, Kohta Ikegami^{3,7}, Michael J. Herriges^{3,8}, Ivan P. Moskowitz^{3,4,5*}, and Aaron M. Zorn^{1,2*}

¹Center for Stem Cell and Organoid Medicine (CuSTOM), Division of Developmental Biology, Perinatal Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

²University of Cincinnati, College of Medicine, Department of Pediatrics, Cincinnati, OH

³Department of Pediatrics, ⁴Department of Pathology, ⁵Department of Human Genetics, University of Chicago, Chicago, IL

⁶Current address: Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX

⁷Current address: Division of Molecular and Cardiovascular Biology, Heart Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

***Correspondence:** aaron.zorn@cchmc.org and imoskowitz@peds.bsd.uchicago.edu

Keywords: Cardiopulmonary, Lung Development, Foregut, Retinoic Acid, Tbx5, Aldh1a2, Hedgehog, Second Heart Field, *Xenopus*

Running Title: Tbx5-dependent RA signaling coordinates cardiopulmonary development

KEY FINDINGS:

- Tbx5 regulates second heart field patterning and pulmonary development via retinoic acid (RA) and Hedgehog (Hh) signaling.
- Tbx5 directly maintains transcription of the RA-synthesizing enzyme *Aldh1a2* in the posterior second heart field mesoderm via an evolutionarily conserved intronic enhancer.
- Downstream of Tbx5, RA directly promotes *Shh* transcription through the evolutionarily conserved MACS1 endoderm enhancer.
- Downstream of Tbx5, RA suppresses FGF signaling to pattern the second heart field while promoting a Hedgehog-Wnt2/2b signaling cascade that induces pulmonary fate.

SUMMARY STATEMENT

Tbx5-dependent Retinoic Acid signaling regulates an evolutionarily conserved gene regulatory network that coordinates cardiac and pulmonary development.

ABSTRACT

The gene regulatory networks that coordinate the development of the cardiac and pulmonary systems are essential for terrestrial life but poorly understood. The T-box transcription factor Tbx5 is critical for both pulmonary specification and heart development, but how these activities are mechanistically integrated remains unclear. We show that *Tbx5* regulates an evolutionarily conserved retinoic acid (RA)-Hedgehog-Wnt signaling cascade coordinating cardiopulmonary development. We demonstrate that Tbx5 directly maintains expression of the RA-synthesizing enzyme *Aldh1a2* in the foregut lateral plate mesoderm via an intronic enhancer that is evolutionarily conserved among terrestrial vertebrates. *Tbx5* promotes posterior second heart field identity in a positive feedback loop with RA, antagonizing a Fgf8-Cyp regulatory module and restricting FGF activity to the anterior. Tbx5/Aldh1a2-dependent RA signaling also directly activates *Shh* transcription in the adjacent foregut endoderm through the conserved MACS1 enhancer. Epithelial Hedgehog then signals back to the mesoderm, where together with Tbx5 it activates expression of *Wnt2/2b* that ultimately induce pulmonary fate in the foregut endoderm. These results provide mechanistic insight into the interrelationship between heart and lung development informing cardiopulmonary evolution and birth defects.

INTRODUCTION

Proper integration of the cardiac and pulmonary systems begins during early embryogenesis and is essential for terrestrial life. A key feature of cardiopulmonary development is evolutionarily conserved bi-directional paracrine signaling between the foregut endoderm, which gives rise to pulmonary epithelium, and the cardiogenic mesoderm (Xie et al 2012; Rankin et al 2016; Steimle et al 2018). These signals integrate with lineage specific transcription factors (TFs) in a poorly understood gene regulatory network (GRN). A better understanding of this cardiopulmonary GRN will provide insight into how heart and lung development is orchestrated and inform the molecular basis of life-threatening cardiopulmonary birth defects.

The vertebrate heart forms from two distinct populations of cardiac progenitor cells in the anterior lateral plate mesoderm (lpm), termed the first and second heart fields (FHF and SHF; Kelly et al 2014). The FHF differentiates first and forms the early heart tube, including portions of the two atria and left ventricle. The SHF contributes to the anterior and posterior poles of the developing heart and differentiates later. The anterior SHF (aSHF) is characterized by expression of *Fgf8*, *Fgf10*, and *Tbx1* and generates the right ventricle, portions of the outflow tract and pharyngeal mesoderm (Rochais et al 2009; Kelly et al 2014). The posterior SHF (pSHF) is characterized by the expression of *Tbx5*, *Osr1*, *Foxf1*, (Xie et al 2012; Hoffmann et al 2014; Steimle et al 2018) and generates the atrial septum and sinus venosus. A subset of the pSHF marked by *Isl1*, *Gli1*, and *Wnt2* expression contains multipotent cardiopulmonary progenitors (CPPs) that give rise to lung mesenchyme, pulmonary vasculature, and myocardium of the inflow tract

(Peng et al 2013). CPPs are both the recipient and source of reciprocal signaling with the adjacent pulmonary endoderm essential for heart and lung development.

The T-box TF *Tbx5* is a key player in coordinating cardiopulmonary organogenesis; it is cell autonomously required in CPPs for heart development and non-autonomously required for pulmonary development (Xie et al 2012; Hoffman et al 2014; Steimle et al 2018; De Bono et al 2018). In its non-cell-autonomous function *Tbx5* is required to activate expression of Hedgehog (Hh) ligands in the adjacent foregut endoderm, which are essential for both heart and lung development (Steimle et al 2018). Endodermal Hh signals back to the lpm stimulating Gli TFs, which cooperate with *Tbx5* to directly activate expression of mesodermal *Wnt2/2b* signals that are essential to induce pulmonary fate in the adjacent foregut endoderm (Steimle et al 2018; Goss et al 2009; Harris-Johnson et al 2009). Thus, *Tbx5* sits atop a hierarchy establishing the reciprocal mesoderm – endoderm - mesoderm signaling loop that coordinates cardiopulmonary development. A major unanswered question is how *Tbx5* non-autonomously activates sonic hedgehog (*Shh*) ligand expression in the foregut endoderm.

A strong candidate for orchestrating the *Tbx5*-dependent cell non-autonomous activation of *Shh* in the foregut endoderm is retinoic acid (RA) signaling. RA deficient embryos have multiple cardiac and pulmonary defects (Zaffran et al 2014; Xavier-Neto et al 2015; Perl and Waxman, 2019; Sirbu et al 2020). RA is produced in the lpm by the enzyme *Aldh1a2* (Niederreither et al 1999; Metzler and Sandell, 2016) and *Aldh1a2* is co-expressed with *Tbx5* in a subset of the pSHF (Hochreb et al 2003; Ryckebusch et al

2008; DeBono et al 2018). RA regulates *Hh* ligand expression in foregut endoderm (Wang et al 2006; Rankin et al 2016) and patterns the SHF by promoting *Tbx5+*/*FoxF1+* pSHF identity whilst repressing *Tbx1+*/*Fgf8+* aSHF fate (Niederreither et al 2001; Sirbu et al 2008; Ryckebusch et al 2008; Deimling et al 2009; Ryckebusch et al 2010; Rydeen and Waxman 2016; Rankin et al 2016; DeBono et al 2018). How the regional production of RA is controlled to pattern the pSHF and regulate *shh* expression remains unknown.

In this study, we tested the hypothesis that *Tbx5* coordinates cardiopulmonary development by directly controlling expression of the RA-producing enzyme *Aldh1a2*, and that this RA signal initiates a mesenchyme-epithelial signaling cascade that controls both Hh/Wnt-dependent lung induction and SHF patterning. We show that *Tbx5* directly maintains expression of *Aldh1a2* in pSHF via an evolutionarily conserved intronic enhancer. We demonstrate that *Tbx5*/*Aldh1a2*-dependent RA signaling regulates anterior-posterior SHF patterning and directly activates *Shh* transcription in the foregut endoderm via an evolutionarily conserved MACS1 endoderm enhancer. Hh/Gli and *Tbx5* then cooperate to promote *Wnt2/2b* expression and lung induction. Our data unify previously unconnected observations in the literature and reveal that *Tbx5* regulates an *Aldh1a2*-dependent RA-Hedgehog-Wnt mesoderm-endoderm-mesoderm signaling network that coordinates pulmonary induction and SHF cardiac patterning.

RESULTS

Tbx5 regulates cardiopulmonary development and *Aldh1a2* expression in mouse

To investigate the Tbx5-regulated GRN that coordinates heart and lung development we re-examined our published RNA-seq data of cardiopulmonary (CP) tissue (containing both foregut mesoderm and endoderm) micro-dissected from wild type and *Tbx5*^{-/-} mouse embryos at E9.5 (Steimle et al 2018). Differential expression analysis revealed 1,588 up-regulated genes and 1,480 down-regulated genes and in the absence of *Tbx5* (≥ 1.5 fold change and 5%FDR) (**Fig.1A; supplemental table S1**; Steimle et al 2018). Reduced expression of Hh signaling components (*Shh*, *Ilh*, *Patch2*), Hh-targets (*Hhip* and *Gli1*), the lung-inducing *Wnt2/2b* ligands, and pulmonary progenitor marker *Nkx2-1*, indicated a loss of pulmonary fate in *Tbx5*^{-/-} mutant CPP tissue (**Fig.1B**). We examined the relationship between pSHF/lung and aSHF/pharyngeal transcripts in *Tbx5*^{-/-} embryos by intersecting the Tbx5-regulated transcriptome with gene sets from recent single cell RNA-seq studies of the developing E7.75-E9.5 mouse heart and foregut that define aSHF, pSHF, pharynx, and lung progenitor cells (de Soysa et al 2019; Han et al 2020) (**supplemental tables S2 and S3**). We found that 25% of genes specifically enriched in aSHF or pharyngeal cells (91/366), but only 5% of genes specifically enriched in pSHF or lung progenitors (10/213) from the single cell studies, overlapped with transcripts up-regulated in the *Tbx5*^{-/-} mutants (* $p < 0.0001$, hypergeometric probability test, HGT) (**Fig.1A**). On the other hand, 34% of the marker genes from the pSHF or lung primordia (72/213) but only 6% of the pSHF+pharynx enriched genes (21/366) were down-regulated in *Tbx5*^{-/-} mutants (* $p < 0.001$, HGT) (**Fig.1A**). The aSHF-enriched genes up-regulated in *Tbx5*^{-/-}

CP tissue included well-known patterning genes *Hand1*, *Irx3*, *Irx5*, *Mef2c*, *Meg3*, and *Tlx1* as well as FGF signaling components and targets including *Fgf8*, *Fgf10*, *Spry1*, *Spry2*, *Dusp6* (**Fig1. B**). Gene set enrichment analysis (GSEA) confirmed a statistically significant overrepresentation of aSHF/pharynx genes among upregulated genes (NES=1.58; $p < 0.0001$) and overrepresentation of pSHF/lung transcripts among the downregulated genes (NES= -1.99; $p < 0.0001$) in the *Tbx5*^{-/-} CP tissue (**supplemental Fig. S1A,B**). Thus, *Tbx5* mutant embryos exhibit a reduction of the pSHF transcriptional program and gain of aSHF gene expression in the pSHF domain, consistent with recent reports (De Bono et al 2018).

Tbx5-dependent genes in the pSHF indicated a possible loss of RA activity. Genes that promote RA signaling were down-regulated in *Tbx5*^{-/-} CP tissue, including *Aldh1a2*, *Crabp2*, which promotes nuclear shuttling of RA, and *Rbp1*, a cytosolic chaperone of the RA precursor retinol (**Fig.1B**). On the other hand, enzymes that attenuate RA signaling, including *Cyp26a1*, *Cyp26b1*, *Cyp26c1*, and *Dhrs3*, were increased in *Tbx5*^{-/-} pSHF/CP tissue. Reduced RA-signaling in *Tbx5*^{-/-} CP tissue was also consistent with increased expression of multiple TGF β pathway components and targets (**supplemental Fig. S1C**), which are known to be suppressed by RA during foregut and heart development (Chen et al 2007; Li et al 2010; Ma et al 2016).

The observation that *Aldh1a2* expression was reduced in the *Tbx5*^{-/-} pSHF whereas FGF signaling components and targets were increased is consistent with the well-known, evolutionarily conserved role of RA in negatively regulating *Fgf8/Fgf10*+ aSHF

fate (Rykebusch et al 2008; Sirbu et al 2008; Rydeen et al 2016). RT-qPCR of additional E9.5 dissected CP tissue validated the RNA-seq analysis with *Aldh1a2* being dramatically down-regulated in *Tbx5*^{-/-} mutants while *Fgf8* and *Fgf10* were up-regulated (**Fig. 1C**). Immunostaining of E8.5 - E9.5 wildtype embryos confirmed previous reports that *Aldh1a2* is co-expressed with *Tbx5* in a subset of the pSHF cells (**Fig. 1D-F and Supplemental Fig. 1SD**) (Hochreb et al 2003; Ryckebusch et al 2008; DeBono et al 2018; de Soysa et al 2019). To investigate the antagonism between RA and FGF signaling within the CP domain, we generated 3D reconstructions using serial sections of *Aldh1a2*, *Fgf8*, *Fgf10*, and *Shh* in-situ hybridizations through the wild-type mouse E9.5 cardiopulmonary domain. These analyses highlighted the distinct aSHF FGF and pSHF RA signal source domains, with *Aldh1a2*⁺ pSHF mesoderm surrounding the *Shh*⁺ pulmonary foregut endoderm (**Fig. 1H, I**). Together these observations were consistent with the possibility that *Tbx5* might regulate SHF patterning by controlling the expression of *Aldh1a2*, which in turn establishes or maintains a local domain of RA activity in the pSHF that subsequently suppresses FGF-mediated aSHF fate and promotes pulmonary development.

Tbx5 regulates cardiopulmonary development and maintains *aldh1a2* expression in *Xenopus*

To elucidate the underlying molecular mechanisms of the *Tbx5*-regulated cardiopulmonary GRN, we turned to *Xenopus* embryos, which facilitate the epistatic analysis of signaling pathways. We previously showed that *Tbx5*-regulated cardiopulmonary development is conserved between *Xenopus* and mouse: *Tbx5* loss of

function (LOF) in *Xenopus*, either by CRISPR/CAS9-mediated mutation or morpholino (MO) knockdown, phenocopies the mouse *Tbx5*^{-/-} phenotype with failed induction of Nkx2-1+ lung progenitors (Steimle et al 2018). Importantly we found that like in mice, *Xenopus tbx5* and *aldh1a2* are co-expressed in the foregut lpm / pSHF and *Xenopus* Tbx5 regulates *aldh1a2* expression (**Fig.2A-C**).

Further analysis showed that Tbx5 LOF in transgenic Wnt/ β -catenin reporter embryos *Tg(WntRE:dGFP)*, where GFP is under the control of Tcf/Lef sites (Tran et al 2010), resulted in a loss of the GFP in the ventral foregut (**Fig.2B,C**) confirming the failure of Wnt/ β -catenin-dependent pulmonary induction. Moreover, cardiopulmonary marker genes that were mis-regulated in the mouse *Tbx5*^{-/-} CP tissue were also mis-regulated in *Xenopus* Tbx5 LOF embryos. Pharyngeal/aSHF markers *fgf8*, *fgf10*, *tbx1*, *cyp26a1*, *spry2*, *hand1*, *hand2*, *dhrs3*, *tgfbR2*, and *tgfbI* were all up-regulated in Tbx5-MO CP foregut tissue, whereas pSHF markers *osr1*, *foxf1*, *gli1* and *wnt2b* and pulmonary endoderm markers *shh*, *dhh*, and *nkx2-1* were reduced (**Fig.3J**). In total, all 18 transcripts tested exhibited changes in gene expression similar to *Tbx5*^{-/-} mice.

Both *X.laevis* Tbx5-MO morphant and *X.trop tbx5* CRISPR/CAS9 mutant embryos exhibited a loss or strong reduction of *aldh1a2* expression in the foregut lpm at NF34 (a timepoint similar to mouse E9.5) (**Fig. 2B-E**). Injection of human *TBX5* RNA rescued *aldh1a2* expression in both *X.laevis* Tbx5 morphants and *X.trop tbx5* mutants (**Fig.2C-E**). A time course analysis revealed that loss of Tbx5 resulted in a downregulation of *aldh1a2* expression in the foregut lpm starting at NF25, but not at early somitogenesis

stages (NF15) (**Fig.2D,E**). *Aldh1a2* transcripts were nearly absent and Aldh1a2 protein was strongly reduced in lpm of both Tbx5 morphants and mutants at NF34 (**Fig.2B-E**). These results demonstrate that Tbx5 regulates a conserved transcriptional program in *Xenopus* and mouse to coordinate SHF patterning and lung induction. Moreover *Xenopus* Tbx5 is required to maintain *aldh1a2* expression in the foregut lpm, suggesting that RA signaling is Tbx5-dependent.

Tbx5 regulates cardiopulmonary development via RA signaling

We tested if failed cardiopulmonary development in Tbx5-deficient *Xenopus* embryos was primarily caused by reduced Aldh1a2-dependent RA signaling. To do so, we asked whether blocking production of endogenous RA could phenocopy loss of Tbx5 or if addition of exogenous RA could rescue the Tbx5 LOF phenotype (**Fig.3A**). Addition of the Aldh enzyme inhibitor DEAB between NF20-34, the time when *aldh1a2* expression was Tbx5-dependent, phenocopied the Tbx5 LOF; pSHF and pulmonary fate markers were reduced/lost and aSHF gene expression was increased/expanded (**Fig.3B-J**). In addition, we treated Tbx5-MO embryos with exogenous RA between NF20-34, using a physiological concentration of 25nM (Horton and Madden 1995; Mic et al 2003; Sheikh et al 2014). As predicted, RA suppressed the expanded expression domains of aSHF markers *fgf8*, *fgf10*, *spry2*, *cyp26a1*, and *tbx1* in Tbx5-deficient embryos (**Fig. 3B-K**). RA was also sufficient to rescue endodermal expression of *shh* and *dhh* as well as expression of known Hh-target genes *gli1*, *foxf1*, and *osr1* in foregut lpm of Tbx5-MO embryos and explants (**Fig. 3B-K**). Exogenous RA treatment did not rescue expression of the pulmonary-inducing *wnt2/2b* ligands nor the lung marker *nkx2-1* in Tbx5 LOF

embryos or foregut explants. This observation was consistent with our previous report that Tbx5 is cell autonomously required to directly promote *wnt2/2b* transcription in the pSHF (Steimle et al 2018). Consistent with this interpretation, addition of recombinant WNT2B protein to Tbx5-MO foregut explants was sufficient to rescue *nkx2-1+* lung fate but it did not rescue *shh* nor *dhh* expression (**supplemental Fig.S2**).

These results combined with our previous data suggest that Tbx5 promotes cardiopulmonary development by multiple mechanisms, which are experimentally separable (**Fig. 3L**). First, by maintaining *aldh1a2* expression, Tbx5 ensures robust RA signaling required for proper SHF pattern and for induction of endodermal *Hh* expression. Second, Tbx5 directly activates *Wnt2/2b* expression in the mesoderm (Steimle et al 2018); Hh signals back to the pSHF mesoderm to via Gli TFs, which cooperate with Tbx5 to promote *Wnt2/2b* expression (Rankin et al 2016; Steimle et al 2018), and ultimately Wnt2/2b signal to the foregut endoderm to induce *Nkx2-1+* pulmonary fate.

Tbx5 directly activates *aldh1a2* transcription and indirectly represses *fgf8* via RA.

In preliminary experiments we found that mouse Tbx5 was sufficient to activate *Aldh1a2* transcription and suppress *Fgf8* and *Fgf10* expression during the directed differentiation of mouse embryonic stem cells into cardiac progenitors using a doxycycline inducible *Tbx5* transgenic cell line (Kattman et al 2011; Steimle et al 2018) (**Fig. 4A**). However, in

these experiments it was unclear whether Tbx5 regulated *Aldh1a2* or *Fgf* expression directly or indirectly.

We therefore examined whether Tbx5 was sufficient to directly activate *aldh1a2* expression in *Xenopus*. We injected RNA encoding a dexamethasone (DEX) inducible Glucocorticoid receptor(GR)-Tbx5 fusion protein (Horb et al 1999) into either the anterior or posterior mesoderm. We then induced GR-Tbx5 nuclear translocation at gastrula stage before endogenous *tbx5* is normally expressed by addition of DEX, with or without the translation inhibitor cycloheximide (CHX) to block secondary protein synthesis (**supplemental Fig.S3A**). GR-Tbx5 activated precocious *aldh1a2* transcription in both the anterior and posterior tissue even in the presence of CHX, demonstrating direct activation (**Fig.4B**). In-situ hybridization of NF34 embryos confirmed robust, ectopic activation of *aldh1a2* by GR-Tbx5 (**Fig.4C**). In contrast, suppression of *fgf8* transcription by GR-Tbx5 was sensitive to CHX, demonstrating indirect repression (**supplemental Fig. S3A,B**). We hypothesized that Tbx5 indirectly represses *fgf8* via *Aldh1a2*-dependent RA production since RA is known to directly repress *Fgf8* transcription in the mouse SHF (Kumar et al 2016). We tested this by inhibiting *Aldh* activity with addition of DEAB; this prevented the suppression of *fgf8* by GR-Tbx5 (**supplemental Fig. S3A-C**). Together, these data demonstrate that Tbx5 directly activates *aldh1a2* transcription and indirectly suppresses *fgf8* expression via RA.

Tbx5 maintains *Aldh1a2* transcription via an evolutionarily conserved intronic enhancer

We next sought to identify *Aldh1a2* enhancers that are directly regulated by Tbx5, predicting that these would be evolutionarily conserved across terrestrial vertebrates. Since a number of putative enhancers have been documented for the mouse *Aldh1a2* locus (Castillo et al 2010; Vitobello et al 2011; Huang et al 2012), we focused on the murine genome. To identify Tbx5-bound enhancers in the CP lineage, we performed Tbx5 chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) of E14.5 fetal mouse lungs as lung mesenchyme is derived from the E9.5 CPPs (Peng et al 2013). ChIP-seq uncovered five Tbx5-bound regions at the *Aldh1a2* locus. Comparing the lung ChIP-seq data to our previously published Tbx5 ChIP-seq from E14.5 heart (Steimle et al. 2018) we found that four of the five Tbx5-bound sites were lung-specific and not bound by Tbx5 in the fetal heart (**Fig. 4D**). Among the four Tbx5-bound sites, only one peak in the *Aldh1a2* first intron, which we refer to as enh1 (for “enhancer 1”; **supplemental Fig.S5A,B**), showed strong evolutionary conservation from mammals to *Xenopus* (**Fig.4D**). The enh1 region also had a strong ATAC-seq peak in E14.5 lungs indicating open enhancer chromatin (**Fig. 4D**). Sequence analysis of enh1 revealed multiple predicted Tbx5 DNA-binding motifs, one of which was perfectly conserved amongst human, mouse, chicken, and *Xenopus* (**Fig.4E**; **supplemental Fig.S5A**).

We tested the ability of both the mouse and *X.trop* enh1 intronic enhancers to drive reporter expression in *Xenopus* transgenics and luciferase reporter assays (**Fig.4F-J**). In transgenics, both the mouse and *X.trop* enh1 enhancers drove GFP expression in the foregut lpm at NF34, although the expression domain was broader than endogenous *aldh1a2* (**Fig.4G**), suggesting additional cis-regulatory elements normally refine expression *in vivo*. To quantitatively assess temporal and spatial enhancer activity, we micro-injected enh1 reporters into blastomeres targeting the future CP-foregut or hindgut regions and assayed luciferase activity at a range of developmental stages, from gastrula to tailbud (**Fig.4H; supplemental Fig.S4**). Neither the mouse nor *X.trop* enh1 enhancers drove significant reporter activity during early development at NF 10.5, 15, or 20; however, at NF25 and N34 both the mouse and *X.trop* enh1 enhancers were active in the foregut but not hindgut, (**Fig.4H; supplemental Fig.S4A,B**). This timing of enh1 enhancer activity (NF25-34) coincides with the timing at which endogenous *aldh1a2* expression is Tbx5-dependent (**Fig.2**). Together these data demonstrate that the evolutionarily conserved enh1 regulates the temporal and spatial transcription of *aldh1a2* in the foregut lpm/pSHF.

We next tested Tbx5 regulation of the enh1 enhancer by combining reporter assays with Tbx5 loss- or gain-of-function experiments. Tbx5-MO knockdown resulted in a dramatic reduction of the mouse and *X.trop* enh1 reporter activity in CP-foregut tissue at NF34, which was rescued by injection of human *TBX5* RNA (**Fig.4I**). Moreover, injection of *Xenopus* or human *TBX5* RNAs were sufficient to ectopically induce robust enh1 reporter activity in hindgut tissue, which does not express endogenous *tbx5* (**Fig.4J**).

Mutation of the single Tbx5-binding site that was perfectly conserved amongst human, mouse, chicken, and *Xenopus* enh1 resulted in a 48% ($p=0.0046$) and 60% reduction ($p=0.0031$) of the mouse and frog reporter constructs, respectively, in the foregut and also significantly blunted their responsiveness to ectopic Tbx5 in the hindgut (**supplemental Fig.S4C,D**). Mutation of all three putative Tbx5 motifs conserved amongst mammals in the mouse enh1 (**Fig. 4J**) largely abolished reporter activity in the foregut as well as Tbx5-responsiveness in hindgut injections (**Fig.4J**). We conclude Tbx5 directly maintains *Aldh1a2* expression via multiple T-box motifs found in an evolutionarily conserved first intron enhancer.

FGF gain-of-function phenocopies Tbx5-loss-of-function in *Xenopus*

In light of the finding that Tbx5-dependent RA signaling suppresses *fgf8* and *fgf10*, we tested if a temporal FGF gain-of-function (GOF) would phenocopy Tbx5 LOF (**Fig.5A**). We treated wild-type CP-foregut explants with recombinant FGF8 protein from NF20-34, the period when exogenous RA was sufficient to rescue Tbx5 LOF. As predicted FGF8 treatment largely phenocopied Tbx5 depletion with increased expression of aSHF/pharyngeal markers *tbx1*, *fgf10*, *spry2*, and *cyp26a1*, as well as reduced expression of pSHF and pulmonary endoderm genes *wnt2b*, *shh*, *gli1*, and *nkx2-1* (**Fig.5B**). We also observed reduced expression of *tbx5* and *aldh1a2* consistent with a feedback loop where FGF restricts Tbx5/Aldh1a2-mediate RA signaling (**Fig.5C**).

FGF signaling is known to promote the expression of RA-degrading Cyp26 enzymes (Shiotsugu et al 2004; Deimling et al 2011; Rydeen and Waxman 2016), but it is unclear

whether this is by direct transcriptional regulation. Therefore, we repeated the FGF8 experiments in the presence of CHX and found that indeed *cyp26a1* was still upregulated by FGF8, demonstrating direct activation (**Fig. 5D**). In contrast, the ability of FGF8 to suppress *shh* was CHX sensitive, demonstrating indirect repression (**Fig. 5D**). We hypothesized that FGF8 indirectly suppresses expression of *shh* and other RA-dependent pSHF genes by promoting Cyp26-mediated RA degradation (**Fig. 5C**). To test this, we treated CP-foregut explants with both FGF8 and the CYP enzyme inhibitor ketoconazole (keto). Keto blocked the ability of FGF8 to suppress *shh*, *dhh*, *tbx5*, *aldh1a2*, *wnt2b*, and *nkx2-1* (**Fig. 5B**), indicating that FGF indeed acts via Cyp-dependent RA degradation. Consistent with Cyp-mediated RA degradation being a major factor in endogenous cardiopulmonary patterning, keto treatment alone elevated expression of pSHF (*tbx5*, *aldh1a2*, *wnt2b*) and pulmonary endoderm genes (*shh* and *nkx2-1*), whilst decreasing aSHF markers (*fgf8*, *fgf10*, *tbx1*) (**Fig. 5B**), similar to exogenous RA treatment (**Fig. 3**). Interestingly, the CYP inhibition data indicated that RA promotes *tbx5* transcription. Indeed, we verified by in situ hybridization that *tbx5* expression in the pSHF/foregut lpm, but not *tbx5* in the FHF/heart tube, was DEAB-sensitive and required RA (**Fig. 5E**). This result is consistent with prior studies in mouse showing RA is required to activate *Tbx5* expression in the pSHF (Niederrereither et al 2001; Ryckebusch et al 2008; De Bono et al 2018). Combined with our finding that *Tbx5* directly maintains *aldh1a2* expression, these data identify a RA-*Tbx5* positive feedback loop in the pSHF.

RA directly promotes *Shh* transcription through the evolutionarily conserved MACS1 endoderm enhancer

Our data suggest that RA from the *Aldh1a2*-expressing lpm is a likely candidate to activate Hh ligand expression in the endoderm. We tested whether exogenous RA could directly activate *shh* and *dhh* transcription in *Xenopus* foregut endoderm explants where the *tbx5/aldh1a2+* lpm, the source of endogenous RA, had been removed (**Fig.6A,B**). Without the RA-producing lpm, the foregut endoderm did not express *shh* nor *dhh*, however, addition of exogenous RA rescued their expression, even in the presence of CHX, demonstrating direct activation (**Fig.6B**). As controls RA also rescued expression of the known direct RA-target *hnf1b*, whereas the known indirect target *ptf1a* was not rescued in the presence of CHX (**Fig.6B**).

Previous work has identified an evolutionarily conserved distal *Shh* enhancer called MACS1 (for mammalian-amphibian-conserved sequences 1), which is located more than 800 Kb from *Shh*, within an intron of the *Rnf32* gene (Sagai et al 2009; Tsukiji et al 2014; Sagai et al 2017). The MACS1 enhancer is able to drive transcription in mouse foregut endoderm but the signals and TFs that control *Shh* expression via the MACS1 enhancer are unknown. An analysis of publicly available ChIP-seq data from human foregut endoderm, differentiated from pluripotent stem cells (hPSCs) in part by RA treatment (Vinckier et al 2020; Wang et al 2015), revealed binding of the RA nuclear receptor RXR at the human *SHH* MACS1 enhancer as well as H3K4me1 and H3K27ac1, epigenetic marks indicative of enhancer activation (**Fig.6C**). Sequence

analysis of the MACS1 enhancer predicted multiple RXR/RAR nuclear RA receptor half sites (Penvose et al 2019), two of which were evolutionarily conserved between human, mouse, chicken, and *Xenopus* (**Fig.6D**; **supplemental Fig.S6**). These findings raised the possibility that RA directly activates *SHH* transcription via the MACS1 enhancer.

We functionally interrogated human and *X.tropicalis SHH* MACS1 enhancer activity in *Xenopus* luciferase assays (**Fig.6E**) and found both could drive robust reporter activity in foregut but not hindgut endoderm, demonstrating spatial specificity (**Fig.6E**).

Disruption of endogenous RA signaling via DEAB treatment (NF20-34) or injection of dominant-negative RAR alpha RNA (dN-RAR α) abolished human and *X.trop* MACS1 enhancer activity (**Fig.6E**). Moreover, mutation of the 2 highly conserved RAR/RXR half sites in the MACS1 enhancers resulted in a dramatic reduction in reporter activity (**Fig.6E**). Together these data indicate that RA signaling directly regulates *shh* expression in foregut endoderm, via conserved RAR/RXR motifs in the conserved *shh* MACS1 enhancer.

DISCUSSION

Tbx5 regulates a RA-HH-Wnt GRN that coordinates SHF patterning and pulmonary specification

Our findings reveal a complex and evolutionarily conserved signaling network downstream of Tbx5 that coordinates early development of the cardiac and pulmonary systems (modeled in **Fig.7**). We identify the following aspects of this SHF mesoderm – pulmonary endoderm signaling network: 1) Tbx5 acts in a positive feedback loop with RA, maintaining *aldh1a2* transcription in pSHF mesoderm by direct activation of an *aldh1a2* enhancer; RA is in turn required to maintain *tbx5* expression in the pSHF; 2) downstream of Tbx5, Aldh1a2-dependent RA signaling is necessary to promote pSHF/PPP identity and suppress aSHF fate; 3) Tbx5-RA and FGF-Cyp form mutually antagonistic modules, with Cyp-mediated RA degradation refining the spatial domain of RA activity; and 4) Tbx5/Aldh1a2-dependent RA signaling from the pSHF is required cell non-autonomously to activate endodermal *Hh* ligand expression by direct RXR/RAR activation of the MACS1 enhancer. Ultimately Hh ligands signal back to the pSHF mesoderm to activate Gli TFs, which cooperate with Tbx5 to directly activate *wnt2/2b* transcription; Wnt2/2b then induce pulmonary fate in the foregut endoderm (Rankin et al 2016; Steimle 2018). Thus, during cardiopulmonary development, Tbx5 regulates production of three key paracrine signals, RA, Hh and Wnt. In this way the lung primordia form next to the pSHF-derived atria and pulmonary vessels, in preparation for these two organ systems to be functionally integrated during development.

Enhancers controlling the reciprocal RA-HH mesoderm and endoderm signaling

We identified enhancers driving the direct response of Tbx5 on the *aldh1a2* locus and the direct response of RA at the *shh* locus. Both of these enhancers are highly conserved amongst air breathing terrestrial vertebrate species. We found that Tbx5-dependent RA signaling directly promotes *Shh* expression in foregut endoderm via RAR/RXR half site motifs in the evolutionarily conserved *Shh* MACS1 enhancer (Sagai et al 2009; Tsukiji et al 2014; Sagai et al 2017; Penvose et al 2019). Other endodermal TFs that also contribute to *Shh* expression in the foregut have been described including FoxA2, Meis3, Islet1, Tbx2, and Tbx3 (Lin et al 2006; Dilorio et al 2007; Tamplin et al 2008; Mesbah et al 2012; Lee et al 2019), but whether or not they directly act on the *Shh* MACS1 enhancer, in cooperation with RA/RAR/RXRs, is unknown.

Previous studies have defined multiple enhancers that regulate different temporal and spatial expression domains of *Aldh1a2* during development, with input from T-box TFs as a reoccurring theme. For example, the T-box TFs VegT, Eomesodermin, and Brachyury, regulate *aldh1a2* in the *Xenopus* gastrula mesoderm via cis-regulatory elements that lie upstream of transcription start site (Gentsch et al 2013; Tomic et al 2019). Subsequent expression of *Aldh1a2* in the paraxial mesoderm and early lpm is promoted by Hox/Pbx/Meis TF complexes acting on a first intron enhancer (Vitobello et al 2011) that is distinct from the enh1 enhancer identified herein. We show that Tbx5 directly maintains *aldh1a2* transcription in the pSHF domain of the anterior lpm via the

enh1 enhancer. Interestingly Tbx1, which can act as a transcriptional repressor, is also known to spatially restrict *Aldh1a2* expression (Guris et al 2006; Aggarwal et al 2006; Ryckebusch et al 2010), although whether or not Tbx1 directly represses *Aldh1a2* transcription is unclear. We speculate that Tbx1 and Tbx5 might engage the same enh1 enhancer in different cell populations with Tbx1 inhibiting *Aldh1a2* transcription in the aSHF and Tbx5 promoting *Aldh1a2* in the pSHF.

Tbx-RA feedback loops: a regulatory node in cardiopulmonary birth defects?

We observed that Tbx5 and RA form a positive feedforward loop in the pSHF: in this domain, Tbx5 directly maintains *Aldh1a2*-dependent RA production and RA maintains *tbx5* expression. Interestingly, RA was not required for *tbx5* expression in the FHF derived heart tube. Conversely, RA signaling is required for *de novo* expression of *Tbx5* in the SHF (de Soysa et al 2019; De Bono et al 2018). We predict that this *de novo* expression in the pSHF is equivalent to the RA-dependent maintenance of *tbx5* that we observed in *Xenopus*. The molecular mechanism by which RA promotes *Tbx5* expression is unclear. Although enhancers that regulate *Tbx5* transcription in the heart and limb have been identified (Minguillon et al 2012; Smemo et al 2012), whether these enhancers control pSHF expression and if they are directly regulated by RA/RAR/RXR complexes remains to be determined.

Interestingly, Tbx1 and RA have an antagonistic relationship during SHF patterning. In mouse, loss of *Tbx1* and aberrant retinoic acid (RA) synthesis result in cardiovascular

defects similar to the phenotype found in human DiGeorge syndrome patients. *Tbx1* represses *Aldh1a2* expression in the mouse aSHF (Guris et al 2006; Aggarwal et al 2006; Ryckebusch et al 2010) and RA suppresses *tbx1* expression in both *Xenopus* pSHF (our study) and in mice (Ryckebusch et al 2010). Moreover, genetically removing one copy of *Aldh1a2*, thereby reducing the level of RA, decreases cardiovascular malformations found in *Tbx1* heterozygous embryos (Ryckebusch et al 2010).

Integrating these *Tbx5* and *Tbx1* results suggests that T-box TF control of RA signaling acts as a mechanistic toggle, wherein RA is activated by *Tbx5* to promote the cardiopulmonary program in the posterior SHF while RA is suppressed by *Tbx1* to repress the cardiopulmonary program in the anterior SHF. *Tbx5*-RA feedforward loops are also required for limb bud development. Limb defects and AVSDs are both a facet of the phenotypic spectrum observed in Holt Oram syndrome in human patients with *TBX5* mutations (Nishimoto et al 2015). This raises the intriguing possibility that *Tbx5*-RA interactions were an evolutionary innovation in both limb and cardiopulmonary mesoderm in the adaptation to terrestrial life and that disrupting this feedforward loop is central to *TBX5*-associated birth defects. Overall, this work provides a framework for understanding the developmental basis of the human birth defects observed in DiGeorge Syndrome and Holt–Oram Syndrome.

MATERIALS AND METHODS

Ethics Statement

Mouse and *Xenopus* experiments were performed according to Institutional Animal Care and Use Committee (IACUC) protocols (University of Chicago protocol 71737; Cincinnati Children's Hospital protocol 2019-0053).

Xenopus Methods

Xenopus Embryo Injections

Wild-type adult *X.laevis* and *X.tropicalis* frogs were purchased from Nasco (Fort Atkinson, WI). Adult transgenic *X.laevis* and *X.tropicalis* Wnt/B-catenin reporter (*Xla.Tg(WntREs:dEGFP)^{Vlemx}*, NXR_0064; and *Xtr.Tg(WntREs:dEGFP)^{Vlemx}*, NXR_1094) were purchased from the National *Xenopus* Resource (RRID:SCR_013713). Ovulation, in-vitro fertilization and natural mating, embryo de-jellying, and microinjection was performed as described (Sive et al 2000). Plasmids for *Xenopus* GR-Tbx5 (Addgene 117248) *Xenopus* Tbx5 (Addgene 117247) (Horb et al 1999) and *Xenopus* dominant-negative RARa (Sharp and Goldstein, 1997) were previously described. Human *TBX5* (Horizon Discovery OHS5894-202500411) was gateway sub-cloned from its entry vector pENTR223 into the expression vector pCSf107mT-Gateway-3'myc (Addgene 67617) using clonase (ThermoFisher 11791020) according to manufacturer's instructions. Linearized plasmid templates were used to make mRNA for injection using the Ambion mMessage mMachine SP6 RNA synthesis kit (ThermoFisher AM1340). Total amounts of injected mRNA were as

follows: GR-Tbx5 RNA, 125 pg; dN-RARa, 200pg; human TBX5-myc,100pg. Previously validated Tbx5 translation-blocking morpholinos (Brown et al 2005; Steimle et al 2018) were injected at the eight-cell stage (mixture of 2.5 ng each MO1+2 per dorsal marginal zone (dmz) in *X.laevis*; mixture of 0.5ng each MO1+2 per dmz in *X.tropicalis*). MOs were purchased from GeneTools (Philmath, OR) and were as follows: Tbx5-MO1: 5'TTA GGA AAG TGT CTC TGG TGT TGC C 3'; a negative control Tbx5 mismatch MO1 with three nucleotides mutated: 5'**T**CA **G**IA AAG **T**AT CTC TGG TGT TGC C 3'; Tbx5-MO2: 5'CAT AAG CCT CCT CTG TGT CCG CCA T 3'; Tbx5 3bp mismatch MO2: 5'**T**AT **C**AG **A**CT CCT CTG TGT CCG CCA T 3' (mismatch bases are indicated in bold, underlined).

For F0 CRISPR-mediated indel mutations, a sgRNA targeting *X.trop tbx5* exon 5 (DNA-binding domain) that causes frameshift mutations was synthesized in-vitro as previously described (Steimle et al 2018). This exon5 sgRNA causes approximately 40% of injected embryos to have a phenotype (**supplemental table S4**). Briefly, 2nL of a mixture containing 50pg/nl sgRNA with 0.5ng/nl Cas9 protein (PNA Bio CP01-20) was injected on either side of the sperm entry point at the 1-cell stage (total of 200pg sgRNA and 2ng Cas9 protein per embryo).

For *Xenopus* whole embryo small molecule treatments, embryos were cultured in 0.1X MBS + 50 ug/mL gent with concentrations of 1uM dexamethasone, 25nM RA, or 10uM DEAB (Sigma D86256). In all experiments, corresponding amounts of vehicle controls (DMSO or 0.2% fatty-acid free BSA) were used.

Gastrula or CP-foregut explants (containing both endoderm and lpm) were micro-dissected in 1X MBS + 50 µg/mL gentamycin sulfate (gent; MP Biochemicals 1676045) +/- 10U/mL dispase (Corning Life Sciences 354235; to help remove the lpm) and were cultured in 0.5X MBS + 0.2% fatty acid free BSA (Fisher BP9704) + 50 µg/mL gent with the following concentrations of factors: 1µM dexamethasone (DEX; Sigma D4902); 1µM cycloheximide (CHX; Sigma C4859); 25nM all-trans RA (Sigma R2625); 100 ng/mL WNT2B (R&D systems 3900-WN-025); 1µM DEAB (Sigma D86256); 100ng/mL FGF8b (R&D systems 423-F8-025); 0.5µM ketoconazole (Tocris 1103). In CHX experiments, explants were treated for 2 hours in CHX prior to DEX+CHX treatment for 6 hours.

Xenopus RT-qPCR

Xenopus explants were dissected from embryos of 2 to 3 separate fertilization/injection experiments, frozen on dry ice in 200µL of TRIzol (ThermoFisher 15596018) and stored at -80°C. RNA was extracted using TRIzol and purified using the Direct-zol RNA miniprep plus kit (ZymoResearch R2070); 500ng RNA was used in cDNA synthesis reactions using Superscript Vilo Mastermix (ThermoFisher 11755050), all according to manufacturer's instructions. qPCR reactions were carried out using Powerup mastermix (ThermoFisher A25742) on ABI StepOnePlus or QuantStudio3 machines. *Xenopus* RT-qPCR primer sequences are listed in **supplemental table S5**. Relative expression, normalized to ubiquitously expressed *odc*, was determined using the $2^{-\Delta\Delta C_t}$ method. Graphs display the average $2^{-\Delta\Delta C_t}$ value +/- standard deviation. Statistical significance ($p < 0.05$) was determined using parametric two-tailed paired T-test, relative to uninjected, untreated explants. Each black dot in the RT-qPCR graphs represents an

independent biological replicate containing 4 explants. Heat map of *Xenopus* RT-qPCR gene expression was generated using Morpheus software (<https://software.broadinstitute.org/morpheus/>) and shows the average $2^{-\Delta\Delta Ct}$ value from 3 biological replicates for each condition.

Xenopus In-situ Hybridization

In-situ hybridization of *Xenopus* embryos was performed as described (Sive et al 2000) with minor modifications. Briefly, embryos were fixed overnight at 4°C in MEMFA (0.1 M MOPS, 2 mM EGTA, 1 mM MgSO₄, 3.7% formaldehyde), washed 3x5 minutes in MEMFA without formaldehyde, dehydrated directly into 100% ethanol, washed 5-6 times in 100% ethanol, and stored at -20°C for at least 24 hours. Proteinase K (ThermoFisher AM2548) on day 1 was used at 2 ug/mL for 10 minutes on stage NF15,NF25 embryos and 5 ug/mL on NF34 embryos; hybridization buffer included 0.1%SDS; RNase A (ThermoFisher 12091021) used at 0.5ug/mL; and anti-DIG-alkaline phosphatase antibody (Sigma 11093274910) used at 1:5,000 in MAB buffer (100 mM Maleic acid, 150 mM NaCl, pH7.5) + 10% heat-inactivated lamb serum (Gibco 16070096) + 2% blocking reagent (Sigma 11096176001). Anti-sense DIG labeled in-situ probes were generated using linearized plasmid cDNA templates with the 10X DIG RNA labeling mix (Sigma 11277073910) according to manufacturer's instructions.

Xenopus Immunofluorescence

Embryos were fixed in 100 mM HEPES (pH7.5), 100 mM NaCl, 2.7% methanol-free formaldehyde for 2 hours at room temperature, dehydrated directly into Dent's post-

fixative (80% Methanol / 20% DMSO), washed 5 times in Dent's, and stored in Dent's at -20°C for at least 48 hours. Embryos were serially rehydrated (75%, 40%, 25% methanol) into PBS+0.1% TritonX-100 (PBSTr). Embryos were then cut in a transverse plane through the pharynx and posterior to the liver to create a foregut sample using a fine razor blade on a 2% agarose-coated dish in PBSTr. Foreguts were subjected to antigen retrieval in 1X R-universal epitope recovery buffer (Electron Microscopy Sciences 62719-10) for 1 hour at $60-65^{\circ}\text{C}$, washed 2×10 min in PBSTr, blocked for 1-2 hr in PBSTr + 10% normal donkey serum (Jackson ImmunoResearch 017-000-001) + 1% DMSO at room temperature, and incubated overnight at 4°C in this blocking solution + primary antibodies: chicken anti-GFP (Aves GPF-1020; diluted 1:1,000), mouse anti-Sox2 (Abcam ab79351; 1:1,000), rabbit anti-Aldh1a2 (Abcam ab96060; 1:500); goat anti-Tbx5 (Santa Cruz Biotechnology sc-17866, 1:350). Secondary antibodies were donkey anti-Chicken 488, donkey anti-rabbit Cy3, and donkey anti-mouse Cy5, donkey anti-goat 405 (Jackson ImmunoResearch 703-546-155, 711-166-152, 715-175-151, and 705-476-147, respectively; all used at 1:1,000 dilution). After extensive washing in PBSTr, samples were incubated overnight at 4°C in PBSTr + 0.2% DMSO + secondary antibodies. Samples were again extensively washed in PBSTr, dehydrated into 100% methanol, washed 5 times in 100% methanol, cleared and imaged in Murray's Clear (2 parts benzyl benzoate, 1 part benzyl alcohol) on a metal slide with glass coverslip bottom using a Nikon A1R confocal microscope to obtain optical sections.

Xenopus Luciferase Assays

The *Xenopus tropicalis* v9.1 genome on Xenbase.org; RRID:SCR_003280 (Karimi et al 2018) was used to define *Xenopus* enhancers. For *Xenopus* luciferase assays, the sequences of the mouse *Aldh1a2* Enh1 (chr9:71241739-71242765; mm10 genome), *X.trop aldh1a2* enh1 (chr3:89631924-89632943; v9.1/xenTro9 genome), *X.trop shh* MACS1(chr6:9535614-9536245; v9.1/xenTro9 genome), and human *SHH* MACS1 (chr7:156459384-156460049; hg19 genome) enhancers, as well as their respective mutant forms, were commercially synthesized (Genscript USA, Piscataway NJ; or IDT DNA, Coralville, IA) and cloned into the pGL4.23 firefly luc2/miniP vector (Promega E8411). For enh1 enhancer assays, embryos were co-injected with 5 pg of pRL-TK:renilla luciferase plasmid (Promega E2241) + 50 pg of the pGL4.23 luc2/miniP enhancer:luciferase plasmid and the following amounts of MOs or mRNAs into each dorsal marginal zone (dmz) region of 4-8cell embryos: 3.5 ng of tbx5-MO or 3bp mismatch-MO; 62.5 pg GR-Tbx5 RNA; 50pg *Xenopus* Tbx5 RNA; 50pg human TBX5-myc RNA. For hindgut mesendoderm injections, the luciferase reporters were injected into the ventral-posterior marginal zone at the 4-8cell stage +/- Tbx5 RNA. For analysis of Shh MACS1 enhancer activity in endoderm, C1 (foregut) or C4 (hindgut) blastomeres were injected at the 16- or 32-cell stage with 5pg pRL-TK+ 50pg MACS1:luc +/- 100pg dnRARa RNA.

Each biological replicate contained a pool of 5 embryos, obtained from 2 to 3 separate fertilization/injection experiments which were frozen on dry ice in a minimal volume of 0.1xMBS and stored at -80°C. To assay luciferase activity samples were lysed in 100uL of 100 mM TRIS-Cl pH7.5, centrifuged for 10 min at ~13,000 x g and then 25uL of the

clear supernatant lysate was used separately in firefly (Biotium #30085–1) and renilla (Biotium 300821) luciferase assays according to the manufacturer's instructions. Relative luciferase activity was determined by normalizing firefly to renilla levels for each sample. Graph show the average relative luciferase activity +/- standard deviation with dots showing values of biological replicates. Statistical significance was determined by parametric two-tailed paired T-test, * $p < 0.05$.

Xenopus transgenesis

Transgenesis was carried out using the I-SceI meganuclease procedure (Ogino et al 2006; Pan et al 2006). *Xenopus* transgenic plasmids were constructed using the pI-SceI-d2EGFP plasmid backbone (Addgene 32674). First, a fragment containing the mouse or *X.trop* enh1 enhancers upstream of a minimal TATA box promoter (Tran et al 2010) flanked by duplicated copies of the 250bp chick B-globin HS4 insulator (Allen and Weeks, 2005; Rankin et al 2011) was commercially synthesized (Genscript USA) and cloned into the Apal/XhoI sites of pBluescript II KS+ (Agilent 212207). Apal/XhoI digestion released this fragment, and it was ligated into Apal/XhoI digested pI-SceI-d2EGFP plasmid. The meganuclease reaction contained 200ng DNA, 2.5 μ l I-SceI enzyme (New England Biolabs R0694S; kept at -80°C and used within 1 month of purchase) in 20 μ l total volume and was incubated at 37°C for 30 minutes. 5nl was then injected twice into 1-cell embryos on either side of the sperm entry point (10nl total of meganuclease reaction injected per embryo). We observed 14/102 (~13%) and 21/183 (11%) GFP+ full transgenic embryos using the mouse and *X.trop* enh1 constructs, respectively, from two independent injection experiments. As a negative control, 0/87

embryos were GFP positive when injected using reactions that omitted the I-SceI enzyme.

Mouse Methods

RNA-seq

RNA-seq of the micro-dissected e9.5 wild-type and *Tbx5*^{-/-} pSHF/CPP was previously published and is available on GEO (Steimle et al 2018; GSE75077). Heat maps were generated using Morpheus software (<https://software.broadinstitute.org/morpheus/>). Columns in the Fig.1 heat map represent biological replicates (*Tbx5*^{+/+} wt N=5, *Tbx5*^{-/-} N=2), and each column replicate contained n=4 pooled CP dissected regions

RT-qPCR, In situ hybridization, and Immunofluorescence

RT-qPCR of dissected, pooled (n=4) mouse e9.5 pSHF/CPP regions was performed as described (Steimle et al 2018), cDNA generated using SuperScript III First-Strand Synthesis SuperMix (ThermoFisher 18080051), and qPCR was performed using PowerUp mastermix (ThermoFisher A25742). Gene-expression levels were normalized by *Gapdh* and RT-qPCR primers are listed in **supplemental table S5**. In situ hybridization on mouse embryos was performed as described (Hoffmann et al. 2009). *Shh* probe was provided by Elizabeth Grove (University of Chicago).

Immunofluorescence of mouse embryos was performed as described (Rankin et al 2016) using rabbit anti-Aldh1a2 (Abcam ab96060; 1:300) and goat anti-Tbx5 (Santa Cruz Biotechnology sc-17866, 1:300).

Reconstructions of whole mount in-situ hybridizations were generated using previously published methods (Steimle et al. 2018). In brief images were obtained and pre-processed using Adobe Photoshop CS3 Extended (version 10.0.1, <http://www.adobe.com>) and reconstructed with AMIRA (version 5.3.2, <http://www.amira.com>). Manual review of each image in the stack was performed and corrections made when necessary. LabelFields for gene expression and tissue were generated from the same series of sections using separate CastField and LabelVoxel modules. The SurfaceGen module was used to generate surfaces from these LabelFields. Gene expression models for two different genes were initially aligned using the Landmark (2sets) module, and a minimum of three landmarks were used to align the separate models. These landmarks were located using the pharyngeal endoderm and ventral edge of the SHF. Final alignments were fine-tuned manually using the Transform editor.

Mouse ESCs

The inducible *Tbx5*OE-mESC line was previously generated (Steimle et al., 2018) and differentiated to the cardiac lineage as described (Kattman et al 2011). Doxycycline (Sigma D9891; concentrations of 0, 5, 10, 25, 50, 100, 250, 500ng/mL) was applied at the cardiac progenitor-like stage (day 6) and cells were harvested for RNA 24 hours later.

ChIP-seq

ChIP-seq was performed using dissected whole lungs from E14.5 CD-1 mouse embryos obtained from Charles River. Chromatin was prepared as previously described (Steimle et al., 2018). For immunoprecipitation, the chromatin extract was incubated with 5ug of the anti-TBX5 antibody (Santa Cruz Biotechnology sc-17866; Lot #G1516) at 4°C for >12 hours in a total volume of 200 µL. The immune-complexes were captured by Protein G-conjugated magnetic beads (Life Technologies, 1003D) and washed as previously described (Steimle et al., 2018). The captured chromatin was eluted in ChIP Elution Buffer (10 mM TrisHCl, pH 8.0; 1 mM EDTA; 1% SDS; 250 mM NaCl) at 65°C. After RNase and proteinase K treatment and reverse cross-linking, DNA were purified. High-throughput sequencing libraries from ChIP and input DNA were prepared using NEBNext Ultra DNA Library Prep Kit (New England Biolabs, E7370S). During library preparation, adaptor-ligated DNA fragments of 200-650 bp in size were selected before PCR amplification using Sera-Mag magnetic beads (GE, 6515-2105-050-250). DNA libraries were sequenced using Illumina Hi-seq instruments (single-end 50 base) by the Genomics Core Facility at the University of Chicago.

Bioinformatics

ChIP-seq Analysis

Raw sequencing reads were aligned to the mm10 genome using Bowtie2 (Langmead and Salzberg, 2012) and SAMtools (Li et al., 2009) requiring a minimum mapping quality of 10 (-q 10). Pooled peak calling was performed using default settings of MACS2 callpeak (Zhang et al., 2008) with a q-value set to 0.05 and tag size set to 6 (-q 0.05 -s 6). A fold-enrichment track was generated using MACS2 with the bdgcmp

function (-m FE) for visualization on the IVG genome browser (Thorvaldsdottir et al 2013). Public data reanalyzed in this study was downloaded from GEO either as Bigwig files or raw reads which were processed as described above.

RNA-seq analysis

RNA-seq of the micro-dissected e9.5 wild-type and *Tbx5*^{-/-} pSHF/CPP was previously published and is available on GEO (Steimle et al 2018; GSE75077). This RNA-seq data was re-analyzed using Computational Suite for Bioinformaticians and Biologists (CSBB – v3.0.0) using *ProcessPublicData* module

[<https://github.com/praneet1988/Computational-Suite-For-Bioinformaticians-and-Biologists>].

Differentially expressed genes (DEGs) between *Tbx5*^{-/-} and wild-type were identified using RUVSeq, with a threshold of 1.5 fold change and 5% FDR. Expression heat maps were generated using Morpheus (<https://software.broadinstitute.org/morpheus/>).

DEGs were compared with gene sets from single cell RNA-seq defining aSHF versus pSHF (de Soysa et al 2019; GSE126128) and pharynx versus CPP + lung progenitor cells (Han et al 2020; GSE136689) from the early mouse embryo. We created a cardio-pharyngeal enriched gene set by combining marker genes of aSHF and pharynx mesendoderm, and a cardio-pulmonary gene set by combining markers pSHF, pulmonary mesoderm and lung endoderm. Overlap in gene sets were visualized by Venn diagrams and significant overlaps were defined by hypergeometric tests (HGT). In addition, we assessed the enrichment of upregulated and downregulated DEGs from

the *Tbx5*^{-/-} embryos compared to the single cell data sets by gene set enrichment analysis (GSEA) (Subramanian et al 2005).

ACKNOWLEDGEMENTS

We thank members of the Zorn and Wells labs for suggestions and Lisa Sandell (University of Louisville) for discussions.

COMPETING INTERESTS

No competing interests declared

FUNDING

This work was funded by NIH Grants P01 HD093363 to A.M.Z. and R01 HL092153 and R01 HL124836 to I.P.M. Support was also provided by NIH Grants T32 GM007183 for J.D.S. R21AG054770 to K.I and T32 HL007381 for J.D.S. and A.B.R.

Data availability

E14.5 mouse lung *Tbx5* ChIP-seq data generated in this study is available from the Gene Expression Omnibus (GEO) accession number GSE167207.

Analysis of public datasets

Previously published data that were reanalyzed include: *Tbx5*-regulated transcriptome, GSE75077 (Steimle et al 2018); single cell RNA-seq defining mouse aSHF and pSHF, GSE126128 (de Soysa et al 2019); single cell RNA-seq defining mouse E9.5 pharynx, CPP, and lung progenitor cells, GSE136689 (Han et al 2020); from the early mouse

embryo, E14.5 mouse heart Tbx5 ChIP-seq GSE139803 (Burnicka-Turek et al 2020); E14.5 mouse lung ATAC-seq (<https://www.encodeproject.org>; Davis et al 2018; ENCSR335VJW); ChIP-seq of human PSC-derived foregut endoderm for RXR, GSE104840 (Vinckier et al 2020); human PSC-derived foregut H3K4me1 and H3K427Ac, GSE54471 (Wang et al 2015).

REFERENCES

- Aggarwal, V.S., Liao, J., Bondarev, A., Schimmang, T., Lewandoski, M., Locker, J., Shanske, A., Campione, M., and Morrow, B.E. (2006). Dissection of Tbx1 and Fgf interactions in mouse models of 22q11DS suggests functional redundancy. *Hum. Mol. Gen.* 15, 3219–3228.
- Allen, B. G., and Weeks, D. L. (2005). Bacteriophage phiC31 integrase mediated transgenesis in *Xenopus laevis* for protein expression at endogenous levels. *Methods in molecular biology* 518, 113–122.
- Burnicka-Turek, O., Broman, M. T., Steimle, J. D., Boukens, B. J., Petrenko, N. B., Ikegami, K., Nadadur, R. D., Qiao, Y., Arnolds, D. E., Yang, X. H., Patel, V. V., Nobrega, M. A., Efimov, I. R., and Moskowitz, I. P. (2020). Transcriptional Patterning of the Ventricular Cardiac Conduction System. *Circulation research*, 127 (3), e94–e106.
- Castillo, H. A., Cravo, R. M., Azambuja, A. P., Simões-Costa, M. S., Sura-Trueba, S., Gonzalez, J., Slonimsky, E., Almeida, K., Abreu, J. G., de Almeida, M. A., Sobreira, T. P., de Oliveira, S. H., de Oliveira, P. S., Signore, I. A., Colombo, A., Concha, M. L., Spengler, T. S., Bronner-Fraser, M., Nobrega, M., Rosenthal, N., and Xavier-Neto, J. (2010). Insights into the organization of dorsal spinal cord pathways from an evolutionarily conserved raldh2 intronic enhancer. *Development* 137 (3), 507–518.
- Chen F., Desai T.J., Qian J., Niederreither K., Lü J., and Cardoso W.V. (2007). Inhibition of Tgf β signaling by endogenous retinoic acid is essential for primary lung bud induction. *Development* 134, 2969–2979.
- De Bono C., Thellier C., Bertrand N., Sturny R., Jullian E., Cortes C., Stefanovic S., Zaffran S., Théveniau-Ruissy M., and Kelly R.G. (2018). T-box genes and retinoic acid signaling regulate the segregation of arterial and venous pole progenitor cells in the murine second heart field. *Hum. Mol. Genet.* 27, 3747–3760.
- de Soysa, T. Y., Ranade, S. S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H. T., Schrick, A., Del Sol, A., Gifford, C. A., and Srivastava, D. (2019). Single-cell

analysis of cardiogenesis reveals basis for organ-level developmental defects. *Nature*, 572 (7767), 120–124.

Deimling S.J., and Drysdale T.A. (2009). Retinoic acid regulates anterior-posterior patterning within the lateral plate mesoderm of *Xenopus*. *Mech. Dev.* 126, 913–923.

Deimling S.J., and Drysdale T.A. (2011). Fgf is required to regulate anterior-posterior patterning in the *Xenopus* lateral plate mesoderm. *Mech. Dev.* 128, 327–41.

Dilorio, P., Alexa, K., Choe, S. K., Etheridge, L., and Sagerström, C. G. (2007). TALE-family homeodomain proteins regulate endodermal sonic hedgehog expression and pattern the anterior endoderm. *Dev. Biol.*, 304 (1), 221–231.

Faial, T., Bernardo, A. S., Mendjan, S., Diamanti, E., Ortmann, D., Gentsch, G. E., Mascetti, V. L., Trotter, M. W., Smith, J. C., and Pedersen, R. A. (2015). Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic stem cells. *Development* 142 (12), 2121–2135.

Goddeeris, M. M., Rho, S., Petiet, A., Davenport, C. L., Johnson, G. A., Meyers, E. N., and Klingensmith, J. (2008). Intracardiac septation requires hedgehog-dependent cellular contributions from outside the heart. *Development* 135 (10), 1887–1895.

Genga, R., Kernfeld, E. M., Parsi, K. M., Parsons, T. J., Ziller, M. J., and Maehr, R. (2019). Single-Cell RNA-Sequencing-Based CRISPRi Screening Resolves Molecular Drivers of Early Human Endoderm Development. *Cell Rep.* 27 (3), 708–718.

Gentsch, G. E., Owens, N. D., Martin, S. R., Piccinelli, P., Faial, T., Trotter, M. W., Gilchrist, M. J., and Smith, J. C. (2013). In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. *Cell Rep.* 4 (6), 1185–1196.

Goss, A. M., Tian, Y., Tsukiyama, T., Cohen, E. D., Zhou, D., Lu, M. M., Yamaguchi, T. P., and Morrisey, E. E. (2009). Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. *Dev. cell*, 17 (2), 290–298.

Guris, D. L., Duester, G., Papaioannou, V. E., and Imamoto, A. (2006). Dose-dependent interaction of Tbx1 and Crkl and locally aberrant RA signaling in a model of del22q11 syndrome. *Dev. cell*, 10 (1), 81–92.

Harris-Johnson, K. S., Domyan, E. T., Vezina, C. M., & Sun, X. (2009). beta-Catenin promotes respiratory progenitor identity in mouse foregut. *Proc. Natl. Acad. Sci. USA* 106, 16287–16292.

Hochgreb, T., Linhares, V. L., Menezes, D. C., Sampaio, A. C., Yan, C. Y., Cardoso, W. V., Rosenthal, N., and Xavier-Neto, J. (2003). A caudorostral wave of RALDH2 conveys anteroposterior information to the cardiac field. *Development* 130 (22), 5363–5374.

Hoffmann, A. D., Peterson, M. A., Friedland-Little, J. M., Anderson, S. A., and Moskowitz, I. P. (2009). Sonic hedgehog is required in pulmonary endoderm for atrial septation. *Development* 136 (10), 1761–1770.

Hoffmann, A. D., Yang, X. H., Burnicka-Turek, O., Bosman, J. D., Ren, X., Steimle, J. D., Vokes, S. A., McMahon, A. P., Kalinichenko, V. V., and Moskowitz, I. P. (2014). Foxf genes integrate tbx5 and hedgehog pathways in the second heart field for cardiac septation. *PLoS Gen.* 10 (10), e1004604.

Horb, M.E, and Thomsen, G.H. (1999). Tbx5 is essential for heart development. *Development* 126, 1739–1751.

Horton, C., and Maden, M. (1995). Endogenous distribution of retinoids during normal development and teratogenesis in the mouse embryo. *Dev. Dyn.* 202, 312-323.

Huang, G. N., Thatcher, J. E., McAnally, J., Kong, Y., Qi, X., Tan, W., DiMaio, J. M., Amatruda, J. F., Gerard, R. D., Hill, J. A., Bassel-Duby, R., and Olson, E. N. (2012). C/EBP transcription factors mediate epicardial activation during heart development and injury. *Science* 338 (6114), 1599–1603.

Karimi, K., Fortriede, J. D., Lotay, V. S., Burns, K. A., Wang, D. Z., Fisher, M. E., Pells, T. J., James-Zorn, C., Wang, Y., Ponferrada, V. G., Chu, S., Chaturvedi, P., Zorn, A. M., & Vize, P. D. (2018). Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic acids research*, 46 (D1), D861–D868.

Kattman, S.J., Witty, A.D., Gagliardi, M., Dubois, N.C., Niapour, M., Hotta, A., Ellis, J., and Keller, G. (2011). Stage-Specific Optimization of Activin/Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines. *Cell Stem Cell* 8, 228–240.

Kelly, R. G., Buckingham, M. E., and Moorman, A. F. (2014). Heart fields and cardiac morphogenesis. *Cold Spring Harb Perspect Med.* 4 (10), a015750.

Klaus, A., Saga, Y., Taketo, M. M., Tzahor, E., and Birchmeier, W. (2007). Distinct roles of Wnt/beta-catenin and Bmp signaling during early cardiogenesis. *Proc. Natl. Acad. Sci. USA* 104 (47), 18531–18536.

Kumar, S., Cunningham, T. J., and Duyster, G. (2016). Nuclear receptor corepressors Ncor1 and Ncor2 (Smrt) are required for retinoic acid-dependent repression of Fgf8 during somitogenesis. *Dev. Biol.* 418 (1), 204–215.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9 (4), 357–359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25 (16), 2078–2079.

Li, P., Pashmforoush, M., and Sucov, H. M. (2010). Retinoic acid regulates differentiation of the secondary heart field and TGFbeta-mediated outflow tract septation. *Dev. Cell* 18 (3), 480–485.

Lin, L., Bu, L., Cai, C.L., Zhang, X., and Evans, S. (2007). *Isl1* is upstream of sonic hedgehog in a pathway required for cardiac morphogenesis. *Dev Biol.* 295, 756–763.

Lee, J.H., Protze, S.I., Laksman Z, Backx PH, and Keller GM. (2017). Human pluripotent stem cell-derived atrial and ventricular cardiomyocytes develop from distinct mesoderm populations. *Cell Stem Cell* 21, 179–194.e4.

Lee, K., Cho, H., Rickert, R. W., Li, Q. V., Pulecio, J., Leslie, C. S., and Huangfu, D. (2019). *FOXA2* Is Required for Enhancer Priming during Pancreatic Differentiation. *Cell Rep.* 28 (2), 382–393.e7.

Luna-Zurita, L., Stirnimann, C. U., Glatt, S., Kaynak, B. L., Thomas, S., Baudin, F., Samee, M. A., He, D., Small, E. M., Mileikovskiy, M., Nagy, A., Holloway, A. K., Pollard, K. S., Müller, C. W., and Bruneau, B. G. (2016). Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. *Cell*, 164 (5), 999–1014.

Ma, M., Li, P., Shen, H., Estrada, K. D., Xu, J., Kumar, S. R., and Sucov, H. M. (2016). Dysregulated endocardial TGFβ signaling and mesenchymal transformation result in heart outflow tract septation failure. *Dev. Biol.* 409 (1), 272–276.

Mesbah, K., Rana, M. S., Francou, A., van Duijvenboden, K., Papaioannou, V. E., Moorman, A. F., Kelly, R. G., and Christoffels, V. M. (2012). Identification of a *Tbx1/Tbx2/Tbx3* genetic pathway governing pharyngeal and arterial pole morphogenesis. *Hum. Mol. Gen.* 21 (6), 1217–1229.

Metzler, M.A.; Sandell, L.L. Enzymatic Metabolism of Vitamin A in Developing Vertebrate Embryos. *Nutrients* 2016, 8, 812.

Mic, F. A., Molotkov, A., Benbrook, D. M., and Duester, G. (2003). Retinoid activation of retinoic acid receptor but not retinoid X receptor is sufficient to rescue lethal defect in retinoic acid synthesis. *Proc. Natl. Acad. Sci. USA* 100 (12), 7135–7140.

Niederreither K., Subbarayan V., Dollé P., and Chambon P. (1999). Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet.* 21, 444-448.

Niederreither, K., Vermot, J., Messaddeq, N., Schuhbaur, B., Chambon, P., and Dolle, P. (2001). Embryonic retinoic acid synthesis is essential for heart morphogenesis in the mouse. *Development* 128, 1019–31.

Ogino, H., McConnell, W.B., and Grainger, R.M. (2006). High-throughput transgenesis in *Xenopus* using I-SceI meganuclease. *Nat Protoc.* 1 (4), 1703-10.

Pan, F. C., Chen, Y., Loeber, J., Henningfeld, K., and Pieler, T. (2006). I-SceI meganuclease-mediated transgenesis in *Xenopus*. *Dev. Dyn.* 235 (1), 247–252.

Peng, T., Tian, Y., Boogerd, C. J., Lu, M. M., Kadzik, R. S., Stewart, K. M., Evans, S. M., and Morrisey, E. E. (2013). Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor. *Nature* 500 (7464), 589–592.

Penvose, A., Keenan, J. L., Bray, D., Ramlall, V., and Siggers, T. (2019). Comprehensive study of nuclear receptor DNA binding provides a revised framework for understanding receptor specificity. *Nat. commun.* 10 (1), 2514.

Perl, E., and Waxman, J. S. (2019). Reiterative Mechanisms of Retinoic Acid Signaling during Vertebrate Heart Development. *J. Dev. Biol.* 7 (2), 11.

Rankin, S. A., Hasebe, T., Zorn, A. M., and Buchholz, D. R. (2009). Improved cre reporter transgenic *Xenopus*. *Dev Dyn.* 238 (9), 2401–2408.

Rankin, S.A., Zorn, A.M., and Buchholz, D.R. (2011). New doxycycline-inducible transgenic lines in *Xenopus*. *Dev Dyn.* 240 (6):1467-1474.

Rankin, S. A., Han, L., McCracken, K. W., Kenny, A. P., Anglin, C. T., Grigg, E. A., Crawford, C. M., Wells, J. M., Shannon, J. M., and Zorn, A. M. (2016). A Retinoic Acid-Hedgehog Cascade Coordinates Mesoderm-Inducing Signals and Endoderm Competence during Lung Specification. *Cell Reports*, 16 (1), 66–78.

Rankin, S. A., McCracken, K. W., Luedeke, D. M., Han, L., Wells, J. M., Shannon, J. M., and Zorn, A. M. (2018). Timing is everything: Reiterative Wnt, BMP and RA signaling regulate developmental competence during endoderm organogenesis. *Dev. Biol.* 434 (1), 121–132.

Rochais, F., Mesbah, K., and Kelly, R. G. (2009). Signaling pathways controlling second heart field development. *Circ. Res.*, 104 (8), 933–942.

Ryckebusch, L., Wang, Z., Bertrand, N., Lin, S. C., Chi, X., Schwartz, R., Zaffran, S., and Niederreither, K. (2008). Retinoic acid deficiency alters second heart field formation. *Proc. Natl. Acad. Sci. USA* **105** (8), 2913–2918.

Ryckebusch, L., Bertrand, N., Mesbah, K., Bajolle, F., Niederreither, K., Kelly, R. G., and Zaffran, S. (2010). Decreased levels of embryonic retinoic acid accelerate recovery from arterial growth delay in a mouse model of DiGeorge syndrome. *Circ. Res.* **106** (4), 686–694.

Rydeen, A. B. and Waxman, J. S. (2016). Cyp26 Enzymes Facilitate Second Heart Field Progenitor Addition and Maintenance of Ventricular Integrity. *PLoS biology* **14** (11), e2000504.

Sagai, T., Amano, T., Tamura, M., Mizushima, Y., Sumiyama, K., and Shiroishi, T. (2009). A cluster of three long-range enhancers directs regional Shh expression in the epithelial linings. *Development* **136**, 1665–1674.

Sagai, T., Amano, T., Maeno, A., Kimura, T., Nakamoto, M., Takehana, Y., Naruse, K., Okada, N., Kiyonari, H., and Shiroishi, T. (2017). Evolution of Shh endoderm enhancers during morphological transition from ventral lungs to dorsal gas bladder. *Nat. commun.* **8**, 14300.

Sharpe, C.R., and Goldstone, K. (1997). Retinoid receptors promote primary neurogenesis in *Xenopus*. *Development* **124**, 515–523.

Sheikh B.N., Downer N.L., Kueh A.J., Thomas T., and Voss A.K. (2014). Excessive versus physiologically relevant levels of retinoic acid in embryonic stem cell differentiation. *Stem Cells* **32**, 1451–1458.

Shiotsugu, J., Katsuyama, Y., Arima, K., Baxter, A., Koide, T., Song, J., Chandraratna, R. A., and Blumberg, B. (2004). Multiple points of interaction between retinoic acid and FGF signaling during embryonic axis formation. *Development* **131** (11), 2653–2667.

Sirbu, I. O., Zhao, X., and Duester, G. (2008). Retinoic acid controls heart anteroposterior patterning by down-regulating Isl1 through the Fgf8 pathway. *Dev. Dyn.* **237** (6), 1627–1635.

Sirbu, I. O., Chiş, A. R., and Moise, A. R. (2020). Role of carotenoids and retinoids during heart development. *Biochimica et biophysica acta. Molecular and cell biology of lipids*, **1865** (11), 158636.

Sive, H.L., Grainger, R.M., and Harland, R.M. (2000). *Early Development of Xenopus laevis: A Laboratory Manual*. Cold Spring Harbor Lab Press, Cold Spring Harbor, NY.

Steimle, J. D., and Moskowitz, I. P. (2017). TBX5: A Key Regulator of Heart Development. *Curr. Top Dev. Biol.* **122**, 195–221.

Steimle, J. D., Rankin, S. A., Slagle, C. E., Bekeny, J., Rydeen, A. B., Chan, S. S., Kweon, J., Yang, X. H., Ikegami, K., Nadadur, R. D., Rowton, M., Hoffmann, A. D., Lazarevic, S., Thomas, W., Boyle Anderson, E., Horb, M. E., Luna-Zurita, L., Ho, R. K., Kyba, M., Jensen, B., Zorn, A.M., Conlon, F.L., and Moskowitz, I. P. (2018). Evolutionarily conserved *Tbx5-Wnt2/2b* pathway orchestrates cardiopulmonary development. *Proc. Natl. Acad. Sci. USA* 115 (45), E10615–E10624.

Stefanovic, S., Laforest, B., Desvignes, J. P., Lescroart, F., Argiro, L., Maurel-Zaffran, C., Salgado, D., Plaindoux, E., De Bono, C., Pazur, K., Théveniau-Ruissy, M., Bérout, C., Puceat, M., Gavalas, A., Kelly, R. G., and Zaffran, S. (2020). *Hox*-dependent coordination of mouse cardiac progenitor cell patterning and differentiation. *eLife* 9, e55124.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102 (43), 15545-50.

Tamplin, O. J., Kinzel, D., Cox, B. J., Bell, C. E., Rossant, J., and Lickert, H. (2008). Microarray analysis of *Foxa2* mutant mouse embryos reveals novel gene expression and inductive roles for the gastrula organizer and its derivatives. *BMC genomics* 9, 511.

Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*, 14 (2), 178–192.

Tosic, J., Kim, G., Pavlovic, M. Schroder, C.M, Mersiowsky, S.L., Barg, M., Hofherr, A., Probst, S., Kottgen, M., Hein, L., and Arnold, S.J. (2019). *Eomes* and *Brachyury* control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nat Cell Biol.* 21, 1518–1531 (2019).

Tran, H. T., Sekkali, B., Van Imschoot, G., Janssens, S., and Vleminckx, K. (2010). *Wnt*/beta-catenin signaling is involved in the induction and maintenance of primitive hematopoiesis in the vertebrate embryo. *Proc. Natl. Acad. Sci. USA*, 107 (37), 16160–16165.

Tsukiji, N., Amano, T., and Shiroishi, T. (2014). A novel regulatory element for *Shh* expression in the lung and gut of mouse embryos. *Mech. Dev.* 131,127–136.

Vermot, J., Niederreither, K., Garnier, J. M., Chambon, P., and Dollé, P (2003). Decreased embryonic retinoic acid synthesis results in a DiGeorge syndrome phenotype in newborn mice. *Proc. Natl. Acad. Sci. USA*, 100 (4), 1763–1768.

Vinckier, N. K., Patel, N. A., Geusz, R. J., Wang, A., Wang, J., Matta, I., Harrington, A. R., Wortham, M., Wetton, N., Wang, J., Jhala, U. S., Rosenfeld, M. G., Benner, C. W.,

Shih, H. P., and Sander, M. (2020). LSD1-mediated enhancer silencing attenuates retinoic acid signalling during pancreatic endocrine cell development. *Nat. comm.* 11(1), 2082.

Vitobello, A., Ferretti, E., Lampe, X., Vilain, N., Ducret, S., Ori, M., Spetz, J. F., Selleri, L., and Rijli, F. M. (2011). Hox and Pbx factors control retinoic acid synthesis during hindbrain segmentation. *Dev. cell* 20 (4), 469–482.

Waldron, L., Steimle, J. D., Greco, T. M., Gomez, N. C., Dorr, K. M., Kweon, J., Temple, B., Yang, X. H., Wilczewski, C. M., Davis, I. J., Cristea, I. M., Moskowitz, I. P., and Conlon, F. L. (2016). The Cardiac TBX5 Interactome Reveals a Chromatin Remodeling Network Essential for Cardiac Septation. *Dev. Cell* 36, 262–275.

Wang, Z., Dolle, P., Cardoso, W. V. and Niederreither, K. (2006). Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives. *Dev. Biol.* 297,433 -445.

Wang, A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N. A., Muth, K., Palmer, J., Qiu, Y., Wang, J., Lam, D. K., Raum, J. C., Stoffers, D. A., Ren, B., & Sander, M. (2015). Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell stem cell*, 16 (4), 386–399.

Weirauch, M. T., Yang, A., Albu, M., Cote, A. G., Montenegro-Montero, A., Drewe, P., Najafabadi, H. S., Lambert, S. A., Mann, I., Cook, K., Zheng, H., Goity, A., van Bakel, H., Lozano, J. C., Galli, M., Lewsey, M. G., Huang, E., Mukherjee, T., Chen, X., Reece-Hoyes, J. S., Govindarajan S., Shaulsky, G., Walhout A.J.M., Bouget, F.Y., Ratsch, G., Larrondo, L.F., Ecker, J.R., and Hughes, T. R. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* 158 (6), 1431–1443.

Xavier-Neto, J., Sousa Costa, Â. M., Figueira, A. C., Caiaffa, C. D., Amaral, F. N., Peres, L. M., da Silva, B. S., Santos, L. N., Moise, A. R., and Castillo, H. A. (2015). Signaling through retinoic acid receptors in cardiac development: Doing the right things at the right times. *Biochem. Biophys. Acta.* 1849 (2), 94–111.

Xie, L., Hoffmann, A. D., Burnicka-Turek, O., Friedland-Little, J. M., Zhang, K., and Moskowitz, I. P. (2012). Tbx5-hedgehog molecular networks are essential in the second heart field for atrial septation. *Dev. Cell* 23 (2), 280–291.

Yamagishi, H., Maeda, J., Hu, T., McAnally, J., Conway, S. J., Kume, T., Meyers, E. N., Yamagishi, C., and Srivastava, D. (2003). Tbx1 is regulated by tissue-specific forkhead proteins through a common Sonic hedgehog-responsive enhancer. *Genes Dev.* 17 (2), 269–281.

Zaffran S, Robrini N, and Bertrand, N. (2014). Retinoids and cardiac development. *J. Dev. Biol.* 2, 50–71.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., & Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome biology*, 9(9), R137.

FIGURES and LEGENDS

Figure 1

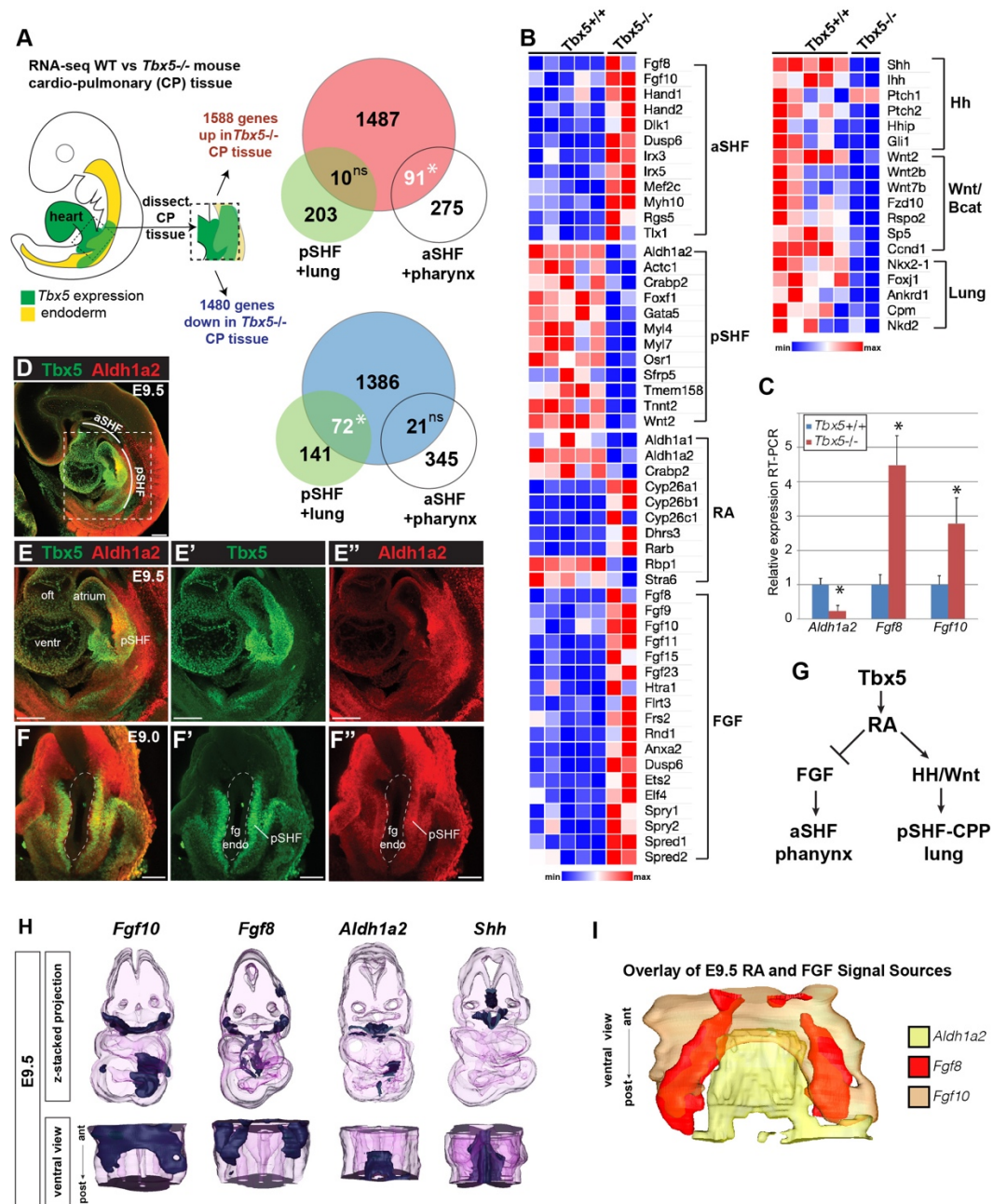


Figure 1. *Tbx5* is required for mouse cardiopulmonary development.

(A) Schematic of an E9.5 mouse embryo highlighting the dissected cardio-pulmonary (CP) tissue (containing mesoderm and foregut endoderm) profiled by bulk RNA-seq.

Venn diagrams show genes differentially expressed in wildtype versus *Tbx5*^{-/-} CP tissue (≥ 1.5 fold change, 5% FDR; Steimle et al 2018; GSE75077) intersected with gene sets from single cell RNA-seq studies defining aSHF + pharynx cells versus pSHF/CP + lung progenitor cells (de Soysa et al 2019; GSE126128, Han et al 2020; GSE136689) (**supplemental table S2 and S3**). Statistically significant intersection based on hypergeometric tests * $p < 0.0001$. ns= not significant. (B) Transcriptome analysis of *Tbx5*^{-/-} CP tissue suggests disrupted SHF patterning and failed pulmonary development with reduced RA and increased FGF signaling. Heat map of selected differentially expressed genes in wild-type *Tbx5*^{+/+} (n=5) and *Tbx5*^{-/-} mutant (n=2) CP tissue grouped by domain of expression or pathway. (C) RT-qPCR validation of decreased *Aldh1a2* and increased *Fgf8*, *Fgf10* expression in E9.5 wild-type and *Tbx5*^{-/-} CP tissue. Relative mean expression + S.D. * $p < 0.05$ Student T-test relative to wild type littermates. (D-F) Wholemount immunostaining of E9.5 (D-E”) or E9 (F-F”) mouse embryos show that *Tbx5* (green) and *Ald1a2* (red) expression overlaps in a subset of the pSHF. Abbreviations: aSHF=anterior second heart field; pSHF = posterior second heart field; oft = outflow tract; ventr = ventricle; fg endo= foregut endoderm. Scale bar = 100 μ M. (G) Model of the proposed *Tbx5*-RA signaling networks in the cardiopulmonary tissue. (H) 3-D reconstructions from section in-situ hybridizations of e9.5 mouse cardiopulmonary foregut region showing expression domains of *Fgf10*, *Fgf8*, *Aldh1a2*, and *Shh* transcripts. Top row is a z-stack 3-D projection; bottom row is a rotated ventral view; (I) shows a ventral view, merged projection of the expression domains of *Fgf10*, *Fgf8*, and *Shh*.

Figure 2

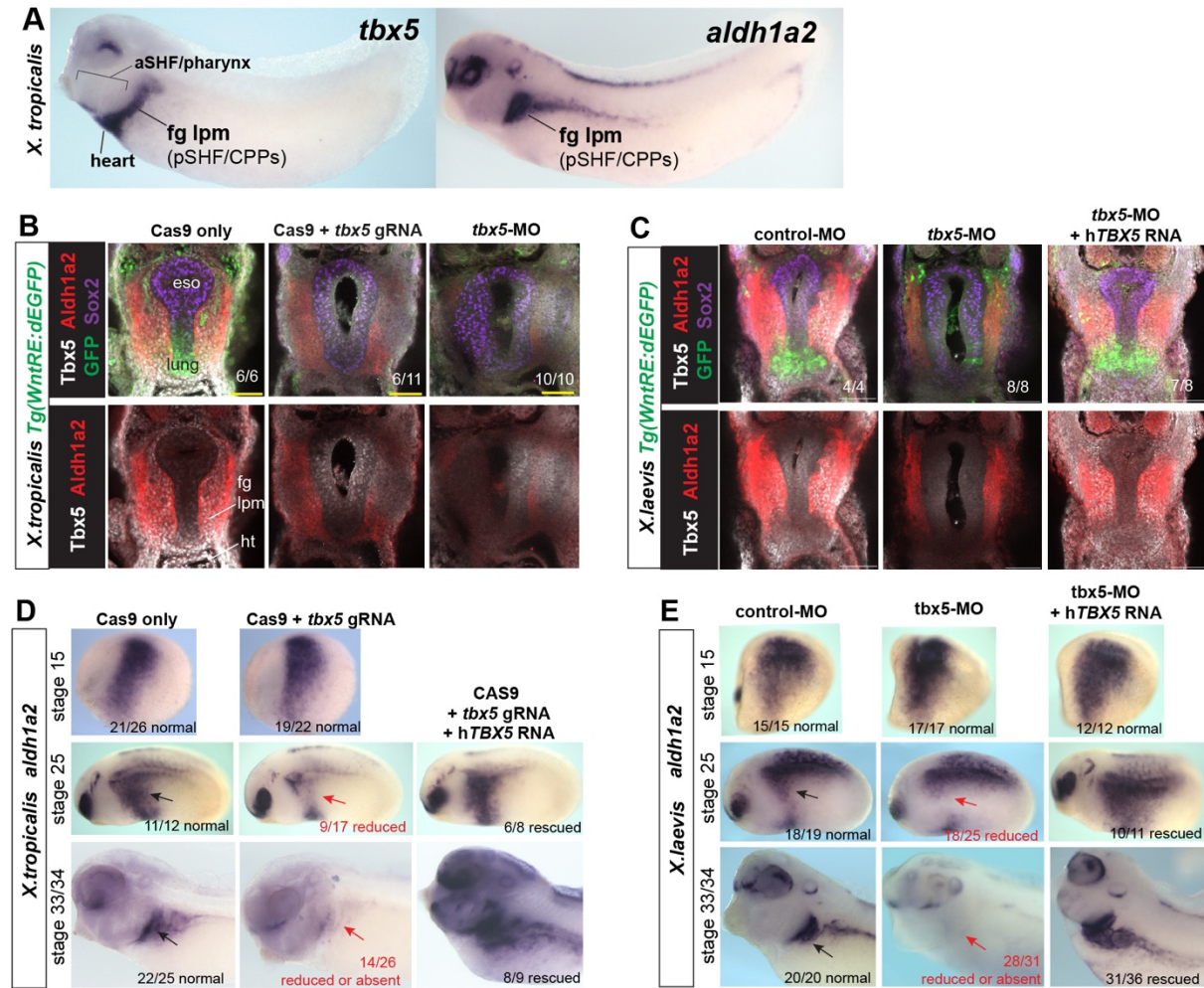


Figure 2. Tbx5 maintains *aldh1a2* expression in *Xenopus* foregut lpm.

(A) Whole-mount in-situ hybridization of *tbx5* and *aldh1a2* expression in *X. tropicalis* embryos (NF32), shows co-expression in the foregut lpm that contains pSHF and CPPs.

(B-C) Immunofluorescence of transgenic Wnt-reporter embryos *Tg(WntRE:dEGFP)* from (B) *Xenopus tropicalis* or (C) *Xenopus laevis* confirm that like in mouse Tbx5 and Aldh1a2 are co-expressed, and that depletion of Tbx5 by CRISPR/CAS9-mediated mutation or Morpholino (MO) injection results in loss of Aldh1a2 and GFP+ pulmonary fate, which is rescued by co-injection of human *TBX5* RNA. Number of imaged embryos

with the observed staining pattern is indicated. Abbreviations: eso, esophagus; ht, heart; pSHF, posterior second heart field. (D-E) Tbx5 is required to maintain *aldh1a2* expression between NF25-34. Wholemount in-situ hybridization of *aldh1a2* expression in *X. tropicalis* (D) or *X. laevis* (E) in control, Tbx5-deficient or Tbx5-rescue embryos at the indicated stage. Expression of *aldh1a2* is unaffected at NF15, reduced by NF25, and largely absent in the fg lpm / pSHF (arrows) at NF33/34. Injection of Human TBX5 RNA (*hTBX5*) rescues *aldh1a2* expression in Tbx5-deficient embryos.

Figure 3

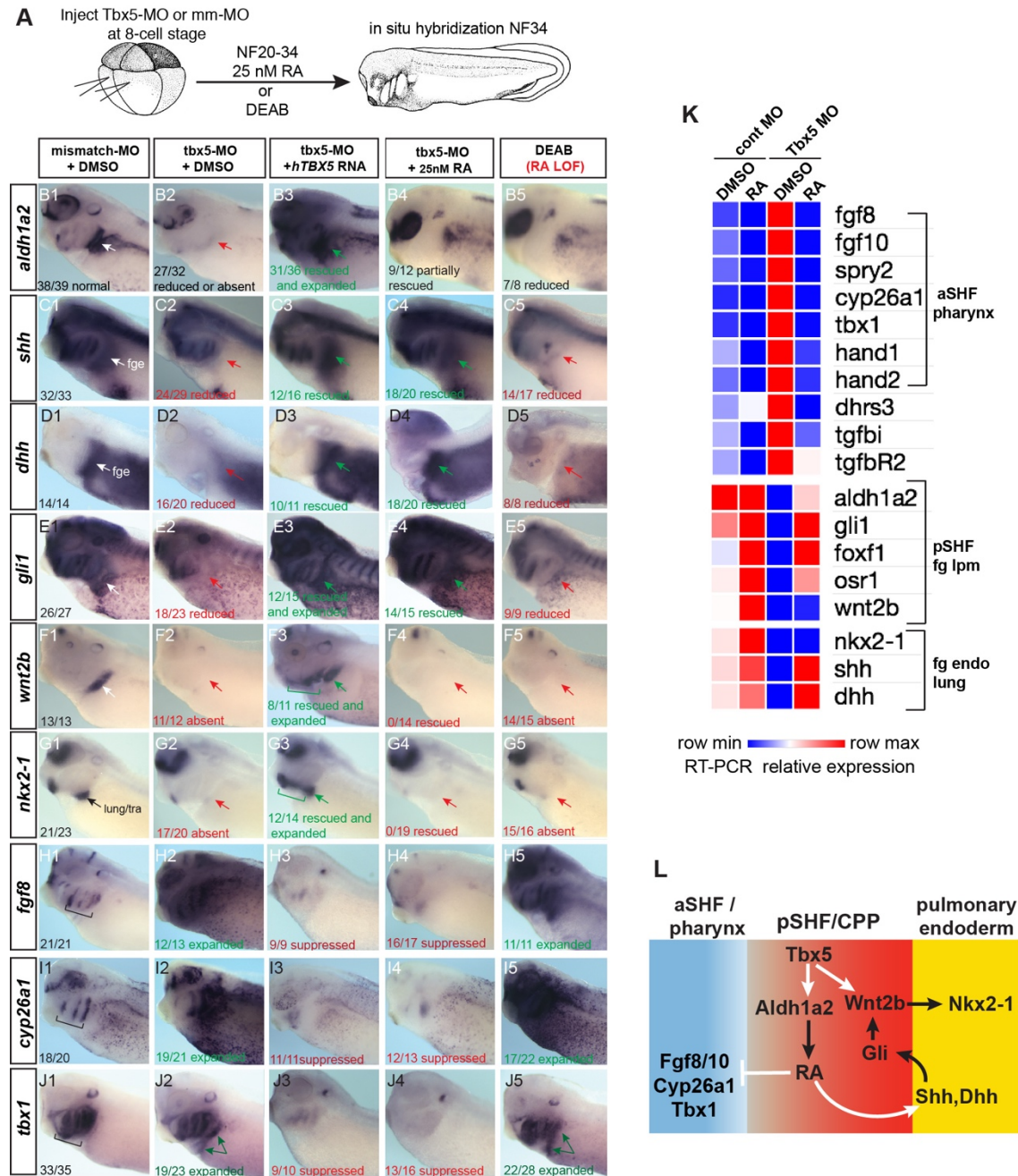


Figure 3. Tbx5 regulates *Xenopus* cardiopulmonary development in part via RA

(A) Schematic of the experimental design. (B-I) Exogenous RA rescues Tbx5 LOF, while inhibition of RA phenocopies Tbx5 LOF. Wholemount in-situ hybridization of NF34 *X. laevis* embryos after the indicated experimental treatments: injection of negative

control 3bp mismatch-MO (10ng), Tbx5-MO (10 ng), Human *TBX5* RNA (*hTBX5*; 100pg), and/or 25nM RA, 10uM DEAB, DMSO vehicle control from NF20-34. The numbers of embryos with the observed expression pattern are indicated. Arrows indicate the relevant expression domain in the CP tissue. Brackets indicate the aSHF/pharyngeal domain. (J) Heatmap showing relative expression from RT-PCR analysis of NF34 Cardiopulmonary (CP)- foregut (fg) tissue dissected from control or Tbx5-MO injected embryos and treated with or without RA from NF20-34. Each row is the average from the 3 biological replicates (n=4 explants per replicate). (K) Diagram of the GRN model showing the key role of *Aldh1a2*-dependent RA signaling downstream of *Tbx5*. White arrows indicate relationships tested in the above experiments and black arrows are inferred from previous publications.

Figure 4

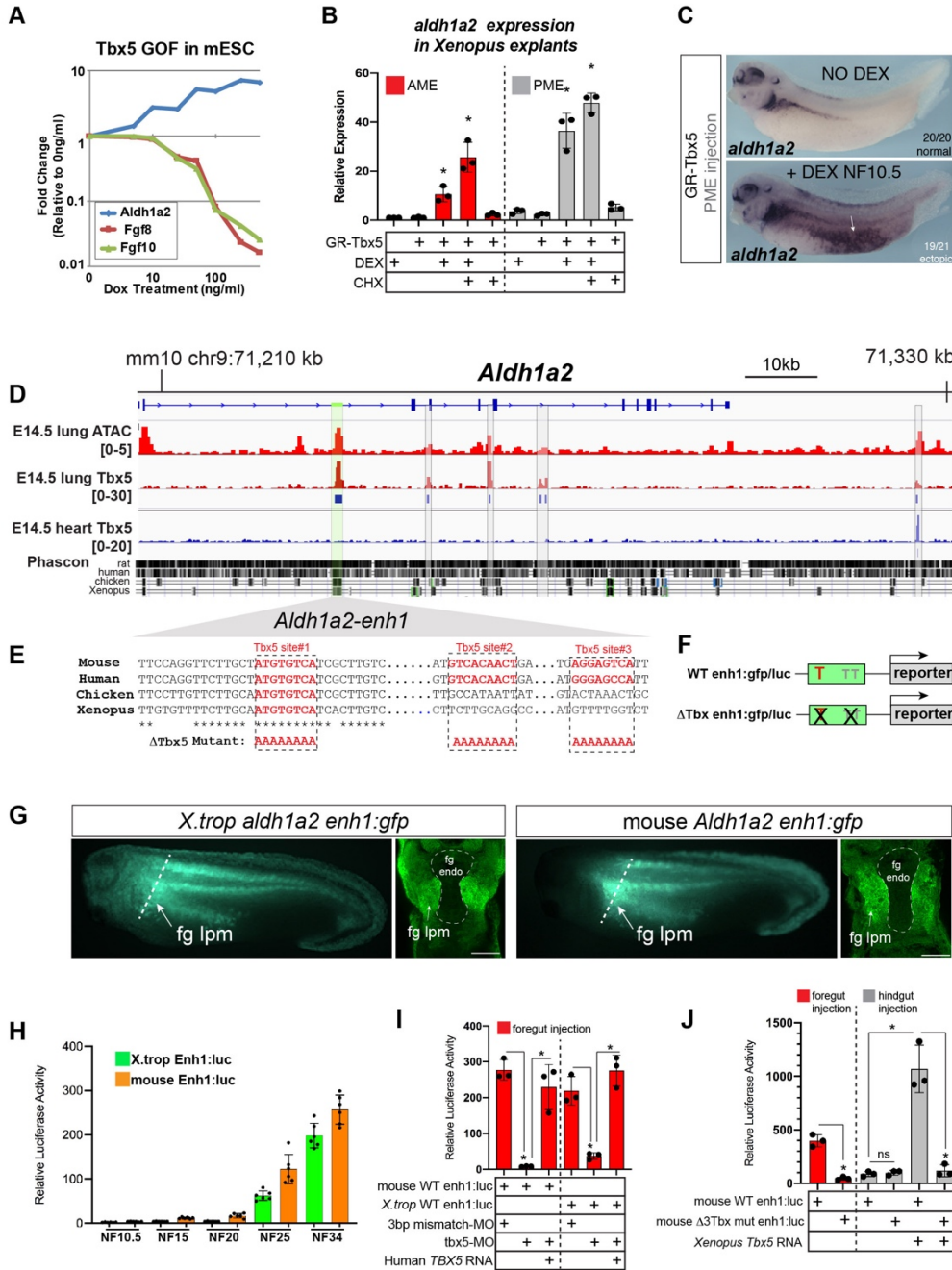


Figure 4. Tbx5 directly activates *aldh1a2* transcription via an evolutionarily conserved first intron enhancer.

(A) Dox inducible Tbx5 activated *Aldh1a2* expression and represses *Fgf8* and *Fgf10* mouse embryonic stem cells (mESC) differentiated into cardiac fate in a does

dependent manner. Graph shows fold change expression relative to uninduced control.

(B) Tbx5 directly activated *aldh1a2* expression in *Xenopus* anterior or posterior mesendoderm (AME, PME) explants. RT-qPCR shows that *aldh1a2* transcription was induced by DEX activated GR-Tbx5 in the presence of CHX. Graphs show mean relative expression \pm standard deviation from N=3 biological replicates, each containing 4 explants/replicate. * $p < 0.05$, pair wise Student T-test relative to uninjected, untreated explants. (C) Wholemount in-situ hybridization of *aldh1a2* expression of *X. laevis* NF34 embryos injected with GR-Tbx5 (100pg) into the PME with or without DEX. (D) Genome browser of the mouse *Aldh1a2* locus showing Tbx5 ChIP-seq tracks from E14.5 mouse lung (GSE167207) and E14.5 mouse heart (Burnecka-Turik et al 2020, GSE139803) as well as ATAC-seq track from the ENCODE project (Davis et al 2018; ENCSR335VJW). Tbx5 ChIP-seq peaks in the E14.5 lung are indicated in blue. The multiple species sequence alignment shows that the prominent Tbx5-bound enhancer the first intron (enh1) is evolutionarily conserved from mammals to *Xenopus*. (E) Multiple species sequence alignment of enh1 reveals one Tbx5 DNA-binding site conserved from mammals to *Xenopus* and two additional mammalian-specific Tbx5 sites, which were mutated in reporter constructs. (F) Schematic of the Wild type (WT) and mutant (Δ Tbx) enh1:gfp and enh1:luciferase reporter constructs. (G) Both the *Xenopus* and mouse intronic enh1 enhancer are sufficient to drive GFP expression in the foregut lpm in *Xenopus* transgenic assays. Wholemounts show GFP epifluorescence and section views of GFP immunostaining. (H) Time course of *Xenopus* and mouse enh1:luc reporter activity injected into *X. laevis* CP-foregut tissue, reflects endogenous Tbx5-dependent *aldh1a2* expression between NF25-34. Graphs show

mean relative luciferase activity \pm standard deviation. N=5 biological replicates/time point with 5 embryos/replicate. *p<0.05, parametric two-tailed paired T-test. (I) The *Xenopus* and mouse *alldh1a2 enh1* reporter constructs are regulated by Tbx5. Graphs show relative mean luciferase activity \pm standard deviation of reporters injected into CP-foregut tissue with control mm-MO, Tbx5-MO, and/or human *TBX5* RNA. N=3 biological replicates/time point with 5 embryos/replicate. *p<0.05, parametric two-tailed paired T-test. (J) The 3 putative Tbx5 motifs in the mouse *alldh1a2-enh1* enhancer are required for reporter activity in the CP-foregut tissue and Tbx5-dependent activation in the hindgut. Graphs show mean relative luciferase activity \pm standard deviation. N=5 biological replicates/time point with 5 embryos/replicate. *p<0.05, parametric two-tailed paired T-test.

Figure 5

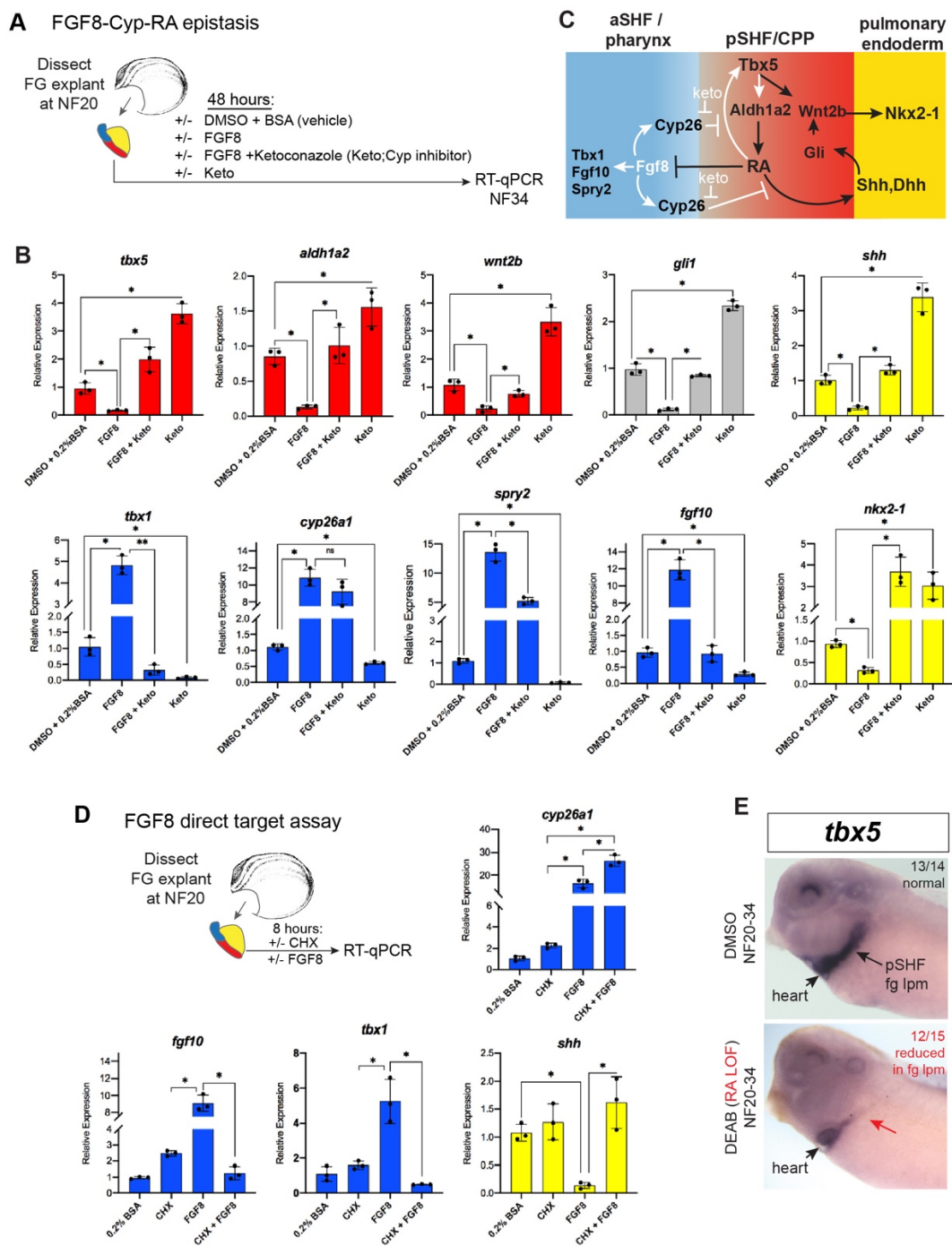


Figure 5. FGF8 gain-of-function phenocopies Tbx5 loss-of-function in *Xenopus*

(A) Schematic of FGF8 gain of function assay in *Xenopus* cardiopulmonary foregut (FG) explants. Explants were micro-dissected at NF20, treated with vehicle controls (DMSO

+0.2%BSA) or the indicated combinations of 100 ng/mL FGF8b and/or 0.5 μ M ketoconazole (Cyp-inhibitor), harvested at NF34, and analyzed via RT-qPCR. (B) RT-qPCR showing mean relative expression of genes for pSHF (red), aSHF (blue) and pulmonary endoderm (yellow), \pm standard deviation from N=3 biological replicates (4 explants/replicate). * p <0.05, parametric two-tailed paired T-test. (C) Model depicting the observed FGF8 gain of function results. White arrows indicate relationships tested in these experiments. (D) FGF8 direct target gene assay in *Xenopus* cardiopulmonary foregut explants, demonstrating FGF8 directly activates *cyp26a1* and indirectly suppresses *shh*. Explants were micro-dissected at NF20, pre-treated with 1 μ M cycloheximide (CHX) for 2 hours prior to culture in 100ng/mL FGF8b + CHX for 6 hours followed by RT-qPCR analysis. Graphs display mean relative expression \pm standard deviation from N=3 biological replicates that contained 4 explants/replicate. * p <0.05, parametric two-tailed paired T-test. (E) RA signaling is required for the *tbx5* expression in the fg lpm / pSHF domain, but not the heart. Embryos were cultured in 10 μ M DEAB from NF20-34 and assayed by in situ hybridization. Number of embryos assayed and with the observed expression pattern is indicated.

Figure 6

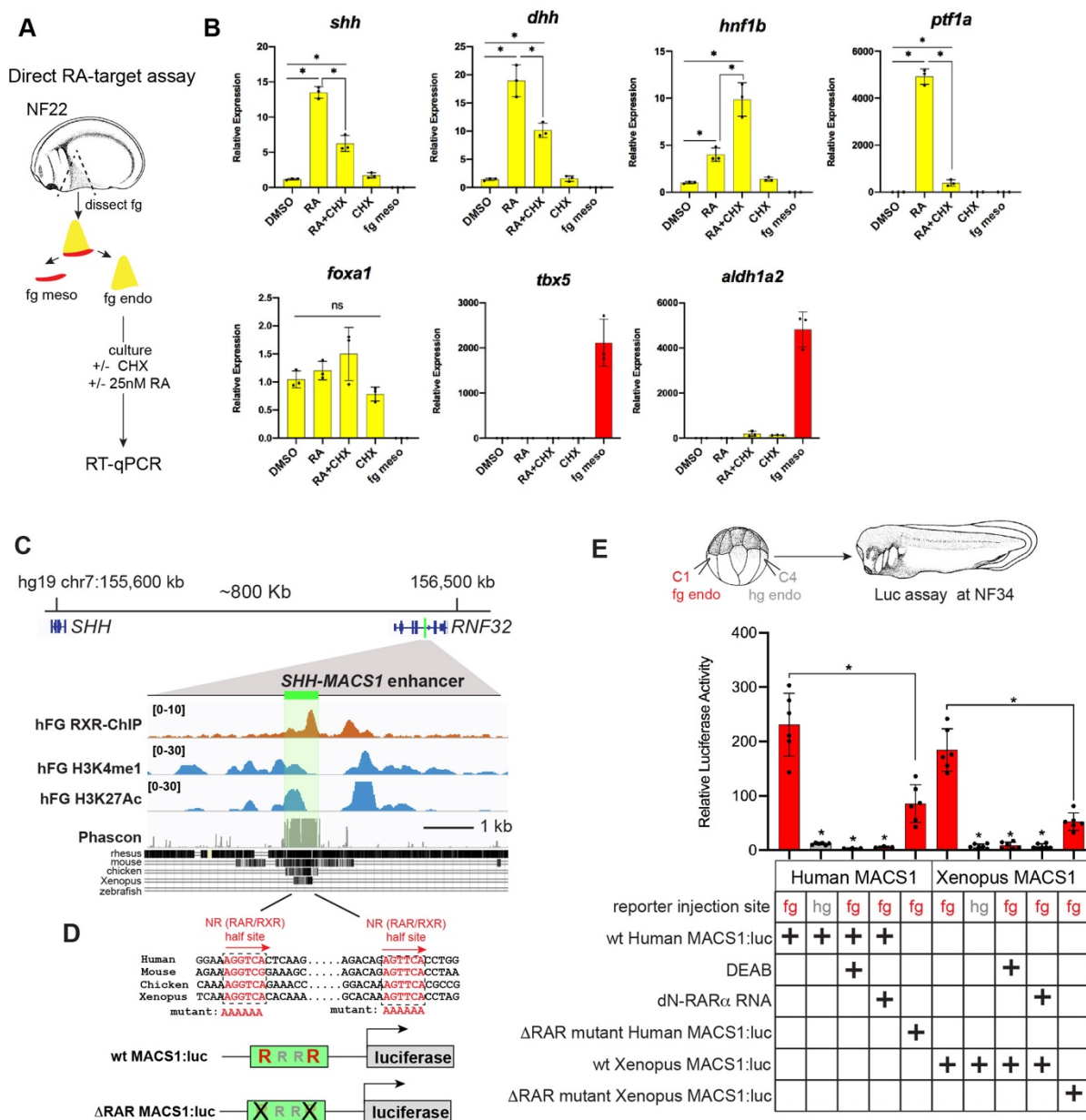


Figure 6. RA-RAR directly activates *shh* transcription in the *Xenopus* foregut endoderm via an evolutionarily conserved MACS1 enhancer.

(A) Schematic of direct RA target gene assay in *Xenopus*. Foregut endoderm (fg endo; yellow) was micro-dissected away from foregut mesoderm (fg meso; red) at NF25, pre-treated with 1uM cycloheximide (CHX) for 2 hours prior to culture in 25nM RA + CHX (or

DMSO vehicle control) for 6 hours followed by RT-qPCR analysis. (B) RA directly activates *shh* and *dhh* expression in the presence of CHX. Graphs show mean relative expression \pm standard deviation from N=3 biological replicates (4 explants/replicate). Endoderm genes are shown in yellow, mesoderm makers in red which confirm clean dissections. * $p < 0.05$, parametric two-tailed paired T-test; ns=not significant. (C) Genome browser of the human *SHH* locus showing the evolutionarily conserved MACS1 distal enhancer known to regulate foregut endoderm expression, which is embedded in an intron of the *RNF32*. Published ChIP-seq tracks of RXR, H3K4me1, and H3K27ac1 from hPSC-derived foregut endoderm (Vinckier et al 2020, GSE104840; Wang et al 2015, GSE54471). Green shading highlights a highly conserved ~634 basepair MACS1 enhancer region. (D) MACS1 enhancer contains multiple nuclear receptor RAR/RXR DNA-binding half sites, 2 of which are highly conserved. Schematics show the wildtype and mutant MACS1:luciferase reporter constructs. (E) Luciferase reporter assay in *Xenopus* show that the Human and *X. tropicalis* MACS1 enhancers are activated by RA via the RAR/RXR DNA-binding sites. 50pg of MACS1:luciferase reporter+ 5pg pRL-TK reporter were microinjected +/- 250pg of dominant-negative RAR α RNA into either the C1 foregut (fg; red bars) or C4 hindgut (hg; grey bars) blastomeres and luciferase activity was assayed at NF34. 10 μ M DEAB treatment was from NF20-NF34. Mean relative luciferase activity \pm standard deviation, from N=6 biological replicates/time point with 5 embryos/replicate. * $p < 0.05$, parametric two-tailed paired T-test relative to WT MACS1:luc in the foregut (fg).

Figure 7

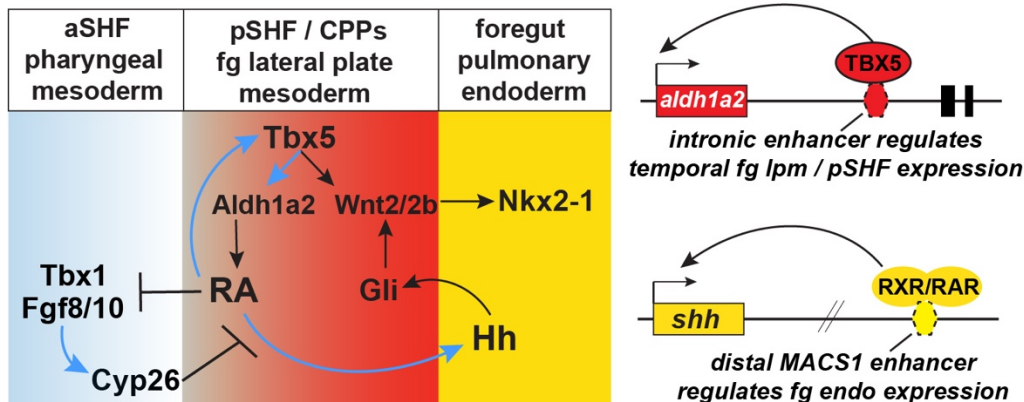
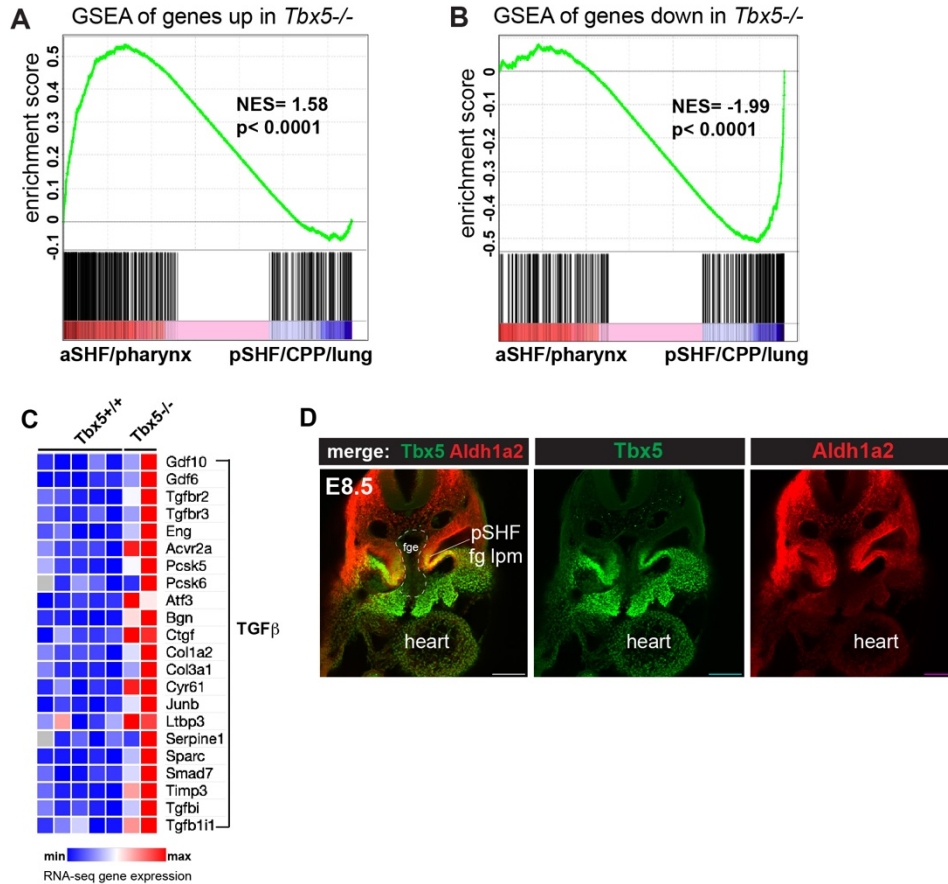


Figure 7. Model of the Tbx5-dependent GRN coordinating SHF patterning and pulmonary induction.

Tbx5 directly maintains *Aldh1a2* expression and a RA-Tbx5 positive feedback loop exists in the pSHF, which is necessary for *Hh* ligand expression, Wnt2/2b-dependent pulmonary fate induction, and SHF heart field patterning. Blue arrows in the model indicate relationships tested and demonstrated in this study. Tbx5/*Aldh1a2* dependent RA signaling restricts FGF/Cyp activity to the aSHF, promotes pSHF identity, and drives expression of *shh* in pulmonary foregut endoderm. Gene models indicate the locations of the *aldh1a2* enh1 enhancer regulated by Tbx5 and the *shh* MACS1 enhancer regulated by RA/RXR/RAR.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure S1

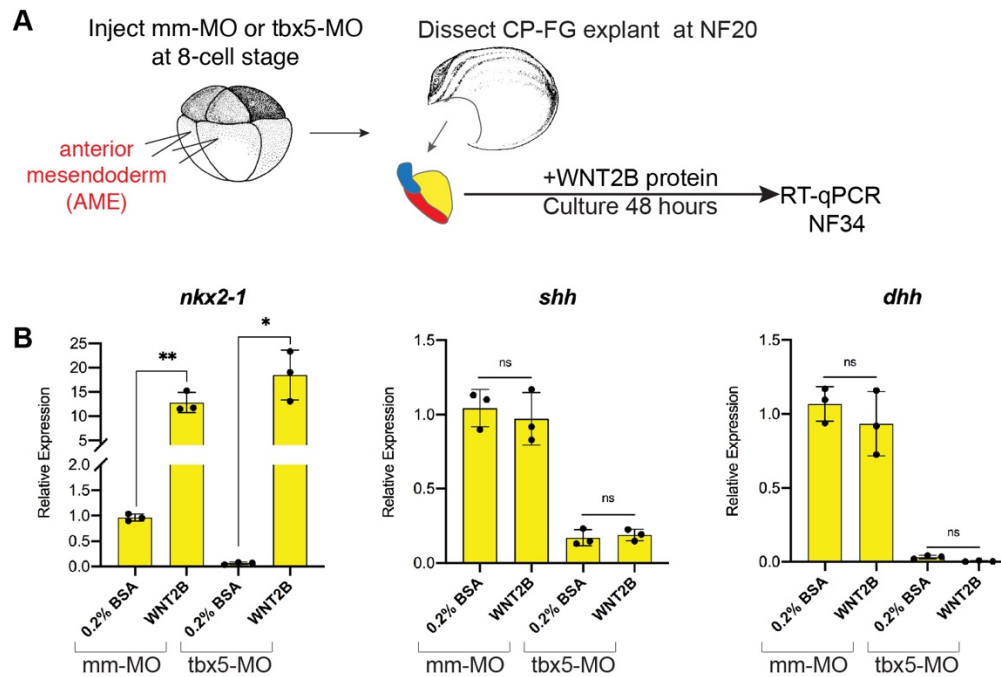


Supplemental Figure S1 – relevant to Fig.1. Analysis of the *Tbx5* regulated transcriptome in cardiopulmonary tissue.

(A-B) Gene set enrichment analysis (GSEA) comparing *Tbx5*-regulated transcriptome in CP tissue ($|\geq 1.5|$ fold change, 5%FDR; Steimle et al 2018, GSE75077) with gene sets from single cell RNA-seq studies defining aSHF versus pSHF (de Soysa et al 2019; GSE126128) and pharynx versus CPP+lung progenitor cells (Han et al 2020; GSE136689) (these gene sets are also listed in **supplemental tables S2,S3**). Genes up regulated in *Tbx5*^{-/-} CP tissue are enriched in aSHF and pharynx and depleted in pSHF/CP/lung enriched markers (A). In contrast transcripts that were downregulated in *Tbx5*^{-/-} CP tissue were significantly enriched pSHF/CP/lung marks and depleted

for aSHF/pharynx markers (B). NES= normalized enrichment score. (C) Components and target genes of TGFB signaling are upregulated in *Tbx5*^{-/-} cardiopulmonary tissue. Heatmap of RNA-seq gene expression. Genes listed in the heatmap are known TGFB ligands, receptors, processing enzymes/signaling transducers, and TGFB target genes identified by manual curation of published literature including: Chen et al 2007; Li et al 2010; Ma et al 2016. (D) Immunostaining of e8.5 mouse embryo foregut region shows co-expression of *Tbx5* (green) and *Aldh1a2* (red) protein in the foregut lateral plate mesoderm (fg lpm) / posterior second heart field (pSHF).

Supplemental Fig S2

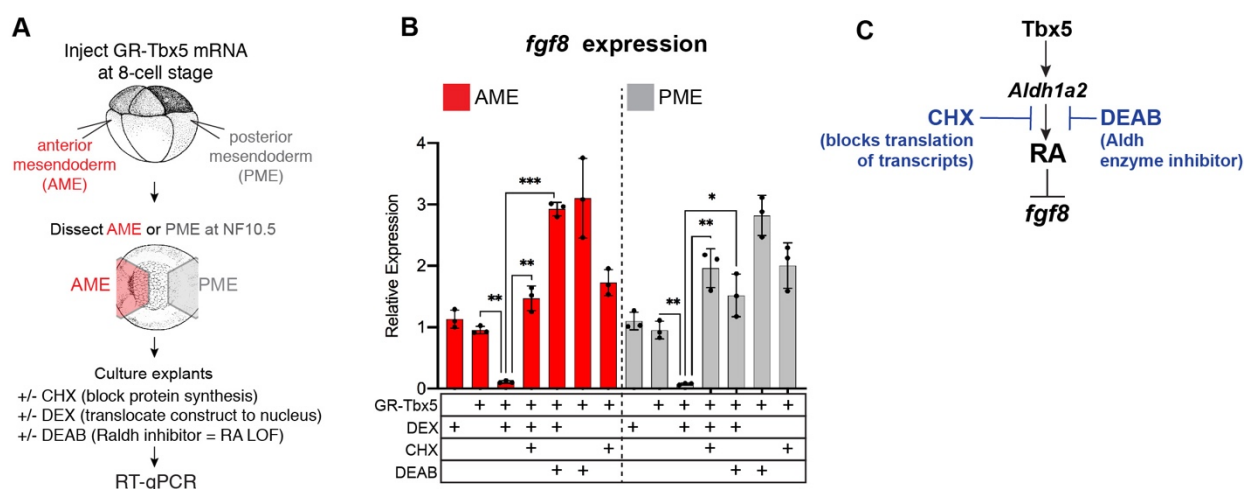


Supplemental Figure S2 – relevant to Fig. 3. WNT2B protein rescues *nkx2-1*+ pulmonary fate in *Tbx5*-depleted foregut explants.

(A-B) WNT2B protein treatment rescues pulmonary *nkx2-1*+ fate, but not *shh* or *dhh*, in *Tbx5*-depleted cardiopulmonary foregut explants. (A) Experimental schematic; (B)

100ng/mL recombinant human WNT2B rescues *nkx2-1* but not *Hh* ligand expression in Tbx5-depleted fg explants. Graphs show mean relative expression \pm standard deviation from N=3 biological replicates (4 explants/replicate). Each black dot in the graphs represents a biological replicate (pool of n=4 explants). *p<0.05, **p<0.01, parametric two-tailed paired T-test relative to uninjected, untreated explants; ns= not significant.

Supplemental Figure S3

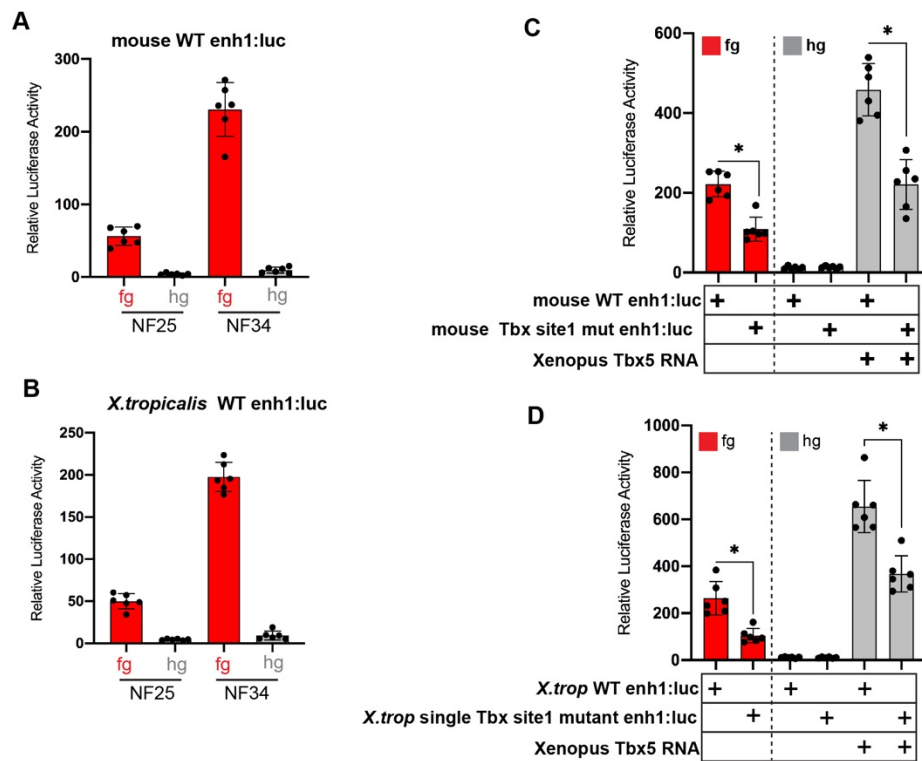


Supplemental Figure S3 – relevant to Fig. 4. GR-Tbx5 indirectly suppresses *fgf8* in *Xenopus* via RA.

(A-C) Experimental schematic of GR-Tbx5 direct/indirect target gene assay in *Xenopus* gastrula explants. (B) GR-Tbx5 suppresses *fgf8* in both anterior and posterior tissue in a CHX sensitive manner demonstrating indirect suppression; the ability of GR-Tbx5 to suppress *fgf8* requires is DEAB sensitive and thus is dependent on RA production. Graphs show mean relative expression \pm standard deviation from N=3 biological replicates (4 explants/replicate). Each black dot in the graphs represents a biological replicate (pool of n=4 explants). *p<0.05, **p<0.01, p<0.001, parametric two-tailed

paired T-test relative to uninjected, untreated explants; ns= not significant. (C) Model of the indirect regulation of *fgf8* by Tbx5, which these experiments show is CHX and DEAB sensitive.

Supplemental Figure S4



Supplemental Figure S4 – relevant to Fig. 4. Additional analyses of the mouse and *X.tropicalis* *Aldh1a2* enhancers in luciferase reporter assays.

(A-B) Both the mouse and *X.tropicalis* (*X.trop*) *Aldh1a2* *enh1* enhancers drive reporter activity in NF25 and NF34 foregut but not hindgut tissue, demonstrating spatial specificity. Graphs show mean relative luciferase activity \pm standard deviation. N=6 biological replicates/time point, each containing n=5 embryos/replicate. *p<0.05,

parametric two-tailed paired T-test. (C-D) Analysis of wild-type enh1 or enh1 with the single, perfectly conserved T-box/Tbx5 motif mutated. (C) Mutation of the single T-box/Tbx5 motif in the mouse (C) or *X.trop* (D) *Aldh1a2* enh1 enhancers significantly reduces enhancer ability to drive expression in the foregut as well as to be activated by exogenous Tbx5 in hindgut gain-of-function injections. Graphs show mean relative luciferase activity \pm standard deviation. N=6 biological replicates/time point, each containing n=5 embryos/replicate. *p<0.05, parametric two-tailed paired T-test.

Supplemental Figure S5

CLUSTAL O(1.2.4) multiple sequence alignment for *Aldh1a2* "enh1" enhancer
 Predicted Tbx5 sites CIS-BP: **score >10**, **score >8**

A



Supplemental Figure S5 – relevant to Fig. 4. Multiple species sequence alignment of the *Aldh1a2* enh1 enhancers.

(A) Clustal DNA sequence alignment; putative T-box/Tbx5 motifs predicted by the CisBP tool (Weirauch et al 2013) are shaded in aqua blue (CisBP score >10) and in yellow (CisBP score >8). Asterisks below the nucleotide alignment indicated conserved

bases amongst all four species. Tbx motifs mutated and tested in this study are indicated in gray. Genomic co-ordinates of the *Aldh1a2* enh1 enhancers used were:

X.tropicalis, >xenTro9.1_dna range=chr3:89631924-89632943;

Chicken, >galGal6_dna range=chr10:7578306-7579190;

Mouse, >mm10_dna range=chr9:71241739-71242765

Human, >hg38_dna range=chr15:58038775-58039647

Supplemental Fig S6

Shh MACS1 enhancer CLUSTAL O(1.2.4) multiple sequence alignment
RAR/RXR sites CIS-BP: **score >10**, **score >8**

A



Supplemental Figure S6 – relevant to Fig. 6. Multiple species alignment of the

Shh MACS1 enhancers.

(A) Clustal DNA sequence alignment; putative RAR/RXR nuclear receptor motifs predicted by the CisBP tool (Weirauch et al 2013) are shaded in aqua blue (CisBP score >10) and in yellow (CisBP score >8). Asterisks below the nucleotide alignment

indicated conserved bases amongst all four species. RAR/RXR motifs mutated and tested in this study are indicated and boxed in red. Genomic co-ordinates of the *Shh* MACS1 (mammalian-amphibian-conserved sequence 1; Sagai et al 2009) enhancers used in the multiple alignment were:

X.tropicalis: v9.1/xenTro9 genome build, chr6:9535614-9536245

Chick: galGal3 genome build, chr2:8370257-8370909

Human: hg19 genome build, chr7:156459384-156460049

Mouse: mm9 genome build, chr5:29538631-29539277