Developmental Biology

Rtf1-dependent transcriptional pausing regulates cardiogenesis

Adam D. Langenbacher, Fei Lu, Luna Tsang, Zi Yi Stephanie Huang, Benjamin Keer, Zhiyu Tian, Alette Eide, Matteo Pellegrini, Haruko Nakano, Atsushi Nakano, Jau-Nian Chen

Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

https://en.wikipedia.org/wiki/Open_access

Copyright information

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint version 1
February 2, 2024 (this version)

Sent for peer review
November 28, 2023

Posted to preprint server
October 15, 2023

Abstract

During heart development, a well-characterized network of transcription factors initiates cardiac gene expression and defines the precise timing and location of cardiac progenitor specification. However, our understanding of the post-initiation transcriptional events that regulate cardiac gene expression is still incomplete. The PAF1C component Rtf1 is a transcription regulatory protein that modulates pausing and elongation of RNA Pol II, as well as cotranscriptional histone modifications. Here we report that Rtf1 is essential for cardiogenesis in fish and mammals, and that in the absence of Rtf1 activity, cardiac progenitors arrest in an immature state. We found that Rtf1's Plus3 domain, which confers interaction with the transcriptional pausing and elongation regulator Spt5, was necessary for cardiac progenitor formation. ChIP-seq analysis further revealed changes in the occupancy of RNA Pol II around the transcription start site (TSS) of cardiac genes in *rtf1* morphants reflecting a reduction in transcriptional pausing. Intriguingly, inhibition of pause release in *rtf1* morphants and mutants restored the formation of cardiac cells and improved Pol II occupancy at the TSS of key cardiac genes. Our findings highlight the crucial role that transcriptional pausing plays in promoting normal gene expression levels in a cardiac developmental context.

eLife assessment

This **important** study conducts genetic analyses utilizing zebrafish, mouse, and mouse embryonic stem cell models to elucidate the role of Rtf1, a component of the PAF1 complex, in early cardiac development. By combining marker gene expression analysis, single-cell transcriptomics, ChIP-seq, and chemical inhibition, the study provides **convincing** evidence that Rtf1-mediated RNAPII (Pol2) transcriptional pausing is required for early cardiac development and that attenuation of pause release by pharmacological inhibition of Cdk9, a component of the PTEF-b complex that regulates the transition between the pausing and elongation phases of transcription, can partially restore transcriptional pausing and cardiogenesis in zebrafish rtf1 mutants. The work will be of broad interest to developmental biologists.

Introduction

Embryonic heart development is directed by an evolutionarily conserved network of transcription factors. Notably, members of the GATA, NK-2 homeobox, TBX, HAND, and MEF2 transcription factor families play crucial, conserved, and well characterized roles in specifying cardiac cells from the mesoderm[1–4]. Through a careful balance of positive and negative regulatory signals, the cardiac transcription regulatory network coordinates with other transcriptional networks to ensure that the appropriate number of cardiac progenitors are specified. In zebrafish, this activity partitions the anterior lateral plate mesoderm (LPM) into multiple fates including the cardiac mesoderm, which is bounded anteriorly by blood and endothelial progenitors and posteriorly by pectoral fin progenitors[5–8].

Chromatin remodeling and epigenetic modifications provide an additional level of transcriptional control at cardiac genes by regulating the binding of transcription factors and the binding and elongation rate of RNA Polymerase II (RNA Pol II). For example, the BAF (Brg1/Brm-associated factor) complex, an ATP-dependent chromatin remodeling complex that actively positions nucleosomes, is critical for normal vertebrate cardiogenesis[9,10]. The BAF complex directly interacts with cardiac transcription factors including Tbx5, Nkx2-5, and Gata4[11] and controls temporal changes in chromatin accessibility during cardiac differentiation[9]. Intriguingly, transplanted cells overexpressing the BAF component Baf60c/Smarca3b and Gata5 preferentially contribute to the embryonic zebrafish heart[12]. Similarly, forced expression of Baf60c along with Tbx5/Nkx2-5/Gata4 can induce cardiac differentiation from non-cardiac mesoderm[13], suggesting that nucleosomal positioning and chromatin accessibility are key mechanisms regulating the transcription of cardiac genes.

Accumulating evidence suggests that cardiac gene expression is also highly sensitive to the activity of transcriptional pausing and elongation factors. Spt6 (Suppressor of Ty 6) is a histone chaperone that promotes RNA Polymerase II elongation and co-transcriptionally modulates chromatin structure[14,15]. In zebrafish, loss of function of Spt6 disrupts cardiomyocyte differentiation and interacts genetically with the transcriptional pausing/elongation factor Spt5 (Suppressor of Ty 5) to support Nkx2.5+ cardiac progenitor specification[16]. While both Spt6 and Spt5 promote transcription elongation, they likely do so via distinct mechanisms. Spt5 directly interacts with Spt4 to form the DSIF (DRB Sensitivity Inducing Factor) complex, which has both inhibitory and stimulatory roles in transcription[17]. Unphosphorylated Spt5 fosters promoter-proximal pausing of RNA Polymerase II, but phosphorylation of Spt5's C-terminal repeat region by Cdk9 (cyclin dependent kinase 9) produces a binding site for the Polymerase Associated Factor 1 complex (PAF1C) component Rtf1, promotes transcriptional pause release, and converts DSIF into an elongation-promoting factor[18–20]. How the transcription regulatory activities of these elongation and pausing factors facilitate cardiac development is not yet understood.

We have previously shown that Rtf1 and other PAF1C members differentially regulate cardiogenesis[21]. The PAF1C consists of five core proteins, Paf1, Ctr9, Cdc73, Rtf1, and Leo1, and plays wide-ranging transcription regulatory roles including promoting co-transcriptional epigenetic modification of histones, supporting both pausing and elongation of RNA Polymerase II, influencing mRNA 3' end formation, and serving as a bridge between specific transcription factors and RNA Polymerase II[22–24]. In zebrafish, Leo1 is dispensable for the specification of cardiomyocytes, but is required for differentiation of the atrioventricular boundary and maturation of the cardiac chambers[25]. Embryos lacking Ctr9, Cdc73, or Paf1, on the other hand, exhibit a deficiency of cardiomyocytes and only form small, dysmorphic hearts. Intriguingly, loss of the Rtf1 subunit of the PAF1C results in the most dramatic defects in cardiogenesis. Rtf1

deficient embryos have severe reduction in *nkx2.5+* cardiac progenitors and fail to generate any cardiomyocytes[21]. Given the numerous roles played by the PAF1C in transcription regulation, it remains an open question how PAF1C members mechanistically regulate cardiogenesis.

In this study, we investigated the molecular mechanisms of Rtf1-dependent cardiac progenitor formation. Using CRISPR mutagenesis we showed that zygotic Rtf1 mutants, like *rtf1* morphants, lack expression of most cardiac progenitor marker genes and produce no cardiomyocytes. In addition, knockout of Rtf1 in the cardiogenic mesoderm of mouse embryos and knockdown of Rtf1 in mouse ESCs reduced the formation of cardiac tissue, indicating that the role of Rtf1 in specification of the cardiac lineage is conserved among vertebrates. Using FACS and single cell sequencing-based approaches, we show that differentiation of the LPM into cardiac precursors requires Rtf1 activity. We investigated which of Rtf1's functional domains participated in cardiac development and found that its Plus3 domain is critical for cardiac progenitor formation, while its HMD (Histone modification domain) is dispensable. Consistent with the Plus3 domain's function as an interaction point between Rtf1 and the DSIF component Spt5, Rtf1 deficient embryos display a genome-wide decrease in promoter-proximal pausing of RNA Polymerase II. Excitingly, chemical and morpholino inhibition of pause release improved pausing at cardiac genes and restored cardiac progenitor formation to Rtf1-deficient embryos. Together, our findings suggest that during development, Rtf1's primary function in cardiac gene regulation is to support promoter proximal pausing, and that normal levels of pausing facilitate expression of many genes, including key members of the cardiac transcription program.

Results

Rtf1 is essential for heart development in zebrafish

Morpholino knockdown of *rtf1* abolishes cardiac progenitor formation in zebrafish, indicating an essential role for Rtf1 in cardiogenesis[21]. To genetically corroborate this finding, we used CRISPR/Cas9 to generate stable *rtf1* mutant lines. We identified two independent lines (*rtf1*^{LA2678} and *rtf1*^{LA2679}) with small deletions in exon 3 of the *rtf1* gene that are predicted to result in frameshifts and early stop codons and which nearly eliminate the production of Rtf1 protein (**Fig 1A-C**). Like *rtf1* morphants, both *rtf1*^{LA2678} and *rtf1*^{LA2679} mutants display no cardiac tissue at 24 hours post fertilization (hpf) and lack the *nkx2.5+* cardiac progenitor population (**Figure 2A,B,E,F**). Rtf1 deficient embryos also fail to express normal levels of other cardiac transcription factors including *mef2ca* and *tbx5a* (**Fig 2C,D,G,H**), and exhibit reduced expression of *tbx20*, a key regulator of cardiac cell number that demarcates the anterior lateral plate mesoderm (ALPM) at the 8-somite stage (**Fig 2I,J**)[26]. Knockdown and knockout of Rtf1 activity produced comparable changes in cardiac gene expression (**Fig 2A-J**), demonstrating that morpholino knockdown faithfully recapitulates the *rtf1* mutant phenotype. Together, these data show that Rtf1 activity is required for the expression of multiple cardiac transcription factors and suggest that the absence of cardiomyocytes in Rtf1 deficient embryos is likely a result of this dramatic gene expression defect.

Conserved role of Rtf1 in mammalian cardiac development

The sequence of the Rtf1 protein is highly conserved among vertebrates (83% identity at the amino acid level between fish and humans) indicating that it is likely playing a fundamental role in biology. To examine whether Rtf1 activity is also needed for cardiogenesis in mammals, we expressed Cre-recombinase under the control of the Mesp1 promoter in Rtf1-flox mouse embryos (**Fig 3A**)[27]. The Mesp1 promoter drives expression of Cre in the early cardiogenic mesoderm (E6.25)[28], allowing for knockout of Rtf1 activity prior to the onset of heart development. Rtf1 Mesp1-Cre knockout embryos die prior to birth, suggesting that Rtf1 function in the cardiogenic mesoderm is essential for survival. We next examined the developing hearts of Rtf1 Mesp1-Cre

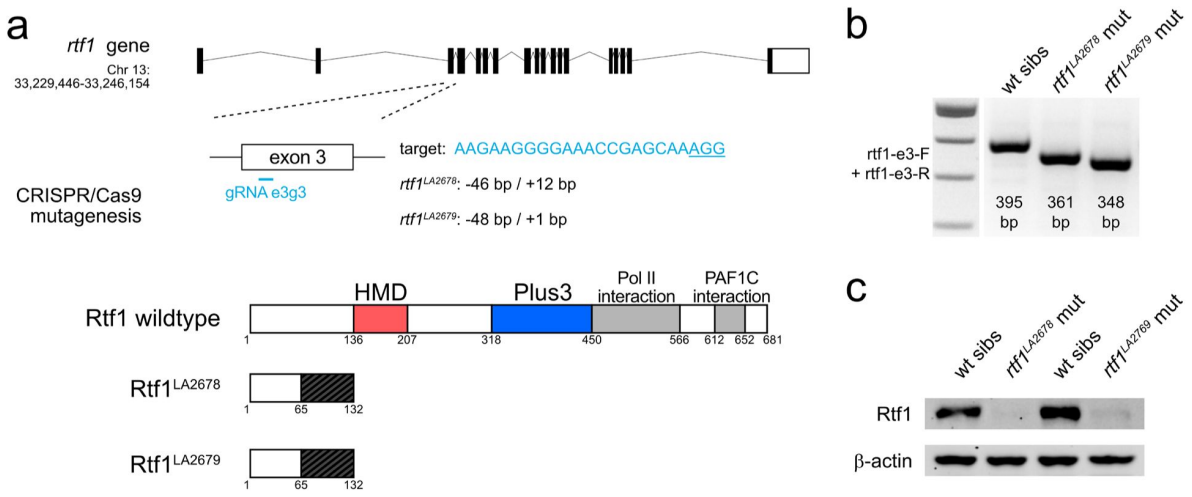

Figure 1.

Generation of zebrafish *rtf1* null mutants with CRISPR/Cas9.

(a) Schematic of CRISPR/Cas9 mutagenesis of the *rtf1* gene. Two mutant alleles, *rtf1*^{LA2678} and *rtf1*^{LA2679}, were recovered after targeting of *rtf1* exon 3. Both alleles are predicted to disrupt translation of the Rtf1 protein and eliminate the HMD, Plus3, Pol II interaction, and PAF1C interaction domains. (b) Agarose gel electrophoresis results of genotyping *rtf1* mutants by PCR. Deletions in *rtf1*^{LA2678} and *rtf1*^{LA2679} alleles can be distinguished from wild type allele using primers rtf1-e3-F and rtf1-e3-R. (c) Western blot detecting Rtf1 and β-actin (loading control) proteins in lysates of wild type and *rtf1* mutant embryos. Image is representative of two independent experiments.

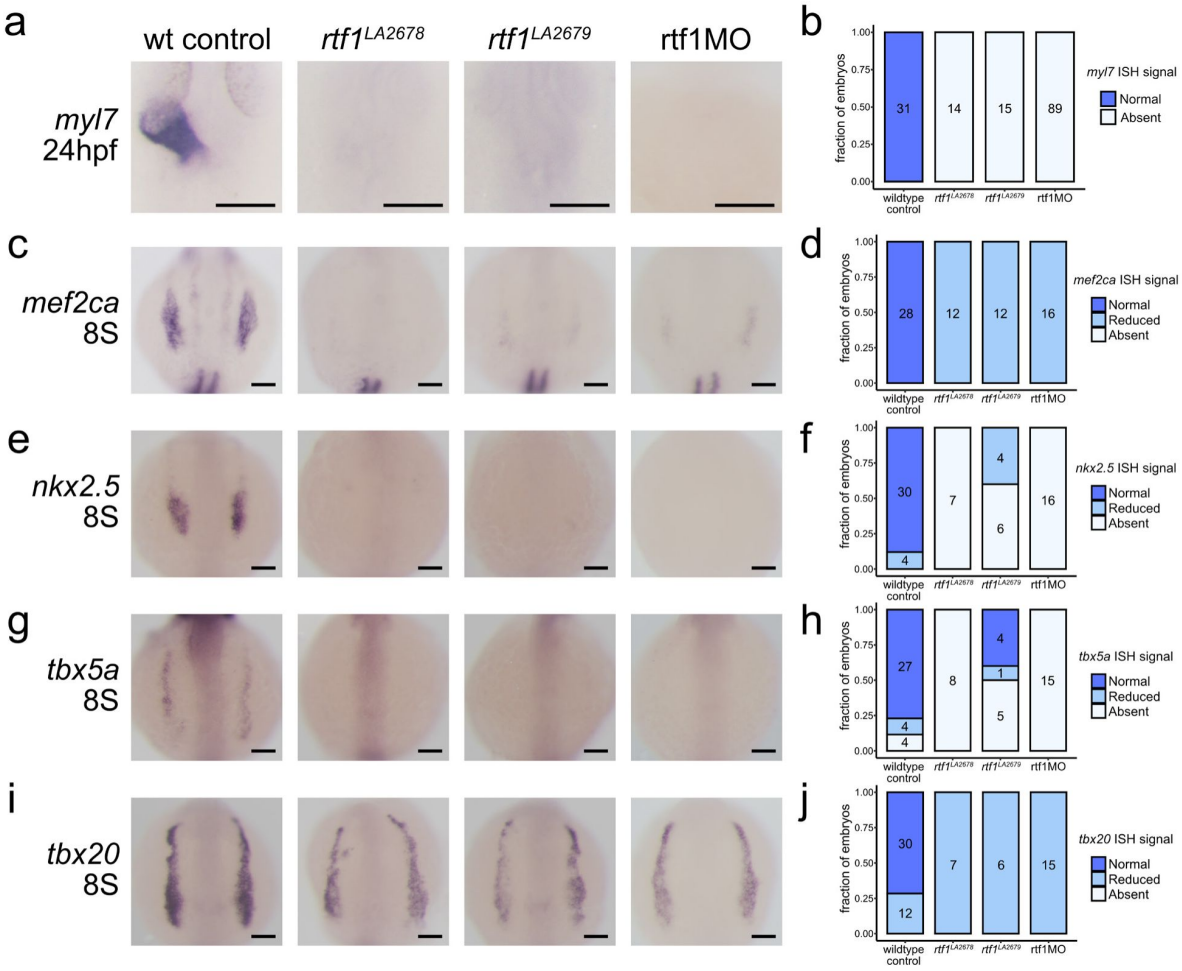

Figure 2.

Characterization of early cardiac development in Rtf1 deficient embryos.

(a) Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf zebrafish embryos. (b) Quantification of *myl7* signal intensity in control and Rtf1 deficient embryos at 24 hpf. Numbers on bars indicate the number of embryos analyzed. (c) Representative images of RNA in situ hybridization detecting *mef2ca* expression in 8 somite stage zebrafish embryos. (d) Quantification of *mef2ca* signal intensity in control and Rtf1 deficient embryos at the 8 somite stage. Numbers on bars indicate the number of embryos analyzed. (e) Representative images of RNA in situ hybridization detecting *nkx2.5* expression in 8 somite stage zebrafish embryos. (f) Quantification of *nkx2.5* signal intensity in control and Rtf1 deficient embryos at the 8 somite stage. Numbers on bars indicate the number of embryos analyzed. (g) Representative images of RNA in situ hybridization detecting *tbx5a* expression in 8 somite stage zebrafish embryos. (h) Quantification of *tbx5a* signal intensity in control and Rtf1 deficient embryos at the 8 somite stage. Numbers on bars indicate the number of embryos analyzed. (i) Representative images of RNA in situ hybridization detecting *tbx20* expression in 8 somite stage zebrafish embryos. (j) Quantification of *tbx20* signal intensity in control and Rtf1 deficient embryos at the 8 somite stage. Numbers on bars indicate the number of embryos analyzed. Scale bars in a, c, e, g, and i represent 0.1 mm.

knockouts and their wild type siblings by in situ hybridization using *Nkx2.5* and *Tbx20* as cardiac markers and found that while Cre+;Rtf1flox/+ and Cre-embryos display a heart tube undergoing looping at E8.5, Rtf1 Mesp1-Cre knockouts have little or no expression of these markers (**Fig 3B**). Cardiac muscle, marked by expression of *MLC2v* at E9.5, was also dramatically reduced in Mesp1-Cre Rtf1 knockouts (**Fig 3B**). These data suggest a conservation of the critical role of Rtf1 in differentiation of the cardiac lineage from the mesoderm.

To further investigate Rtf1's role in cardiac differentiation, we knocked down Rtf1 activity in mouse embryonic stem cells (mESCs) using short hairpin RNA (**Fig 4A,B**). Unmanipulated embryoid bodies (EBs) derived from mESCs are capable of differentiating into lineages representing all three germ layers, as evidenced by temporally dynamic expression of *Brachyury* (mesoderm), *Afp* (endoderm), and *Ncam1* (ectoderm) (**Fig 4E-G**). Similarly, we found that Rtf1 shRNA mESCs induce expression of *Brachyury*, *Afp*, and *Ncam1* in a temporally appropriate manner, indicating that Rtf1 activity is dispensable for germ layer formation (**Fig 4E-G**). However, EBs derived from Rtf1-deficient mESCs produced significantly fewer beating patches (~67% reduction) and also displayed a dramatic reduction in cardiac genes like *Myh6*, *Nkx2.5*, and *ANF*/*Nppa* (**Figure 4C,D**). Together, these data demonstrate that Rtf1 is essential for differentiation of the cardiac lineage from mesoderm in mammals, and reveal a conservation of Rtf1's function in cardiogenesis between fish and mammals.

Rtf1 is required for expression of genes associated with the progression of cardiac precursor formation

We previously found that loss of Rtf1 activity does not disrupt expression of the LPM marker *hand2*[21]. To better understand how Rtf1 directs cardiac differentiation from multipotent mesoderm cells, we took advantage of a hand2:GFP zebrafish transgenic reporter line to isolate LPM cells from wild type and Rtf1 deficient embryos for transcriptomic analysis (**Fig 5A-C**, Supplementary Figure 1). Compared to wild type control samples, the *rtf1* morphant LPM displayed 1420 significantly downregulated and 1185 significantly upregulated genes (adjusted *p*-value < 0.05) at the 10-12 somite stage when cardiac progenitor formation is underway in the LPM (Supplementary Data File 1). We examined the most significantly differentially expressed genes and observed that many of the significantly downregulated genes, including *nkx2.5*, *rbfox1l*, *atp2a2a*, *tnnt2a*, *isl2b*, *mef2ca*, *mef2cb* and *wnt2bb*, are cardiac-lineage markers or have critical known roles in cardiac development (**Fig 5D**). Indeed, the most significantly downregulated genes were highly enriched for gene ontologies related to cardiac development, embryonic heart morphogenesis, and muscle cell development (**Fig 5E**). These data show that Rtf1 activity is required for expression of a broad cardiac gene program in the LPM that is likely under the control of multiple cardiac transcription factors.

To further analyze the deployment of the cardiac gene program during cardiac progenitor formation from the ALPM, we performed Single Cell Multiome sequencing of control and *rtf1* morphant embryos at the 11-12 somite stage. Clustering of control and *rtf1* morphant cells based on gene expression and chromatin accessibility resolved 39 cell types corresponding to expected major cell lineages (Supplementary Data File 2, Supplementary Figures 2 & 3). Interestingly, while *rtf1* morphant cells of each cell type were present, the frequencies of cell types differed significantly between controls and *rtf1* morphants (Supplementary Figure 4). Notably, we also observed that while *rtf1* morphant cells expressing the precardiac mesoderm marker *mef2cb* were present, they failed to express other cardiac marker genes including *ryr2b* and *tnni1b* (Supplementary Figure 3).

To examine developmental trajectories within the ALPM, we reclustered ALPM cells (*sema3e*+/*gata5*+ LPM) and ALPM derivatives (putative precardiac mesoderm, endothelial precursors, rostral blood precursors) (**Figure 6A,B**). We next calculated pseudotime trajectories

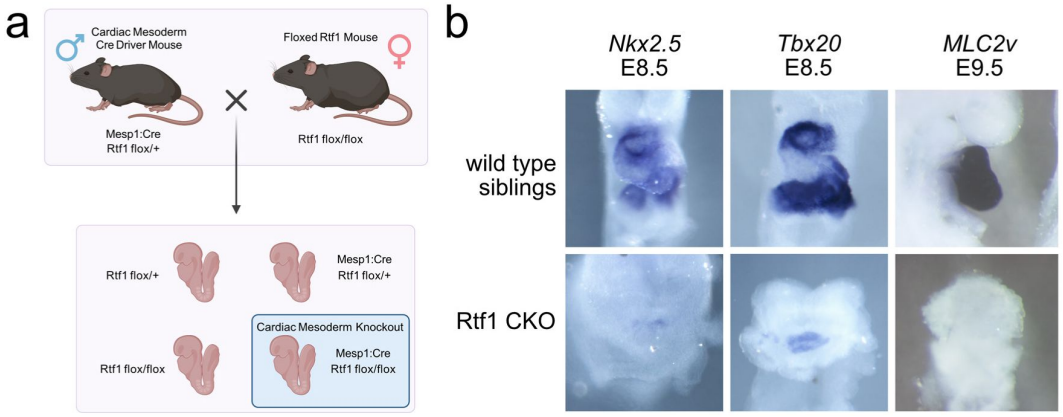

Figure 3.

Failure of cardiac differentiation upon loss of Rtf1 in the mammalian cardiac mesoderm.

(a) Diagram of generation of cardiac mesoderm-specific Rtf1 knockout mouse embryos. A Mesp1: Cre insertion allele was bred into an Rtf1 exon 3 floxed background. Heterozygous Rtf1 floxed male mice with Mesp1:Cre were bred to homozygous Rtf1 floxed females, resulting in one quarter of embryos lacking Rtf1 activity in the cardiac mesoderm. (b) Representative images of RNA in situ hybridization detecting *Nkx2.5* (wild type n = 7, Rtf1 CKO n = 5), *Tbx20* (wild type n = 16, Rtf1 CKO n = 4), and *MLC2v* (wild type n = 1, Rtf1 CKO n = 1) gene expression in Cre-negative (wild type siblings) and Cre-positive homozygous Rtf1 floxed (Rtf1 CKO) embryos at E8.5 or E9.5.

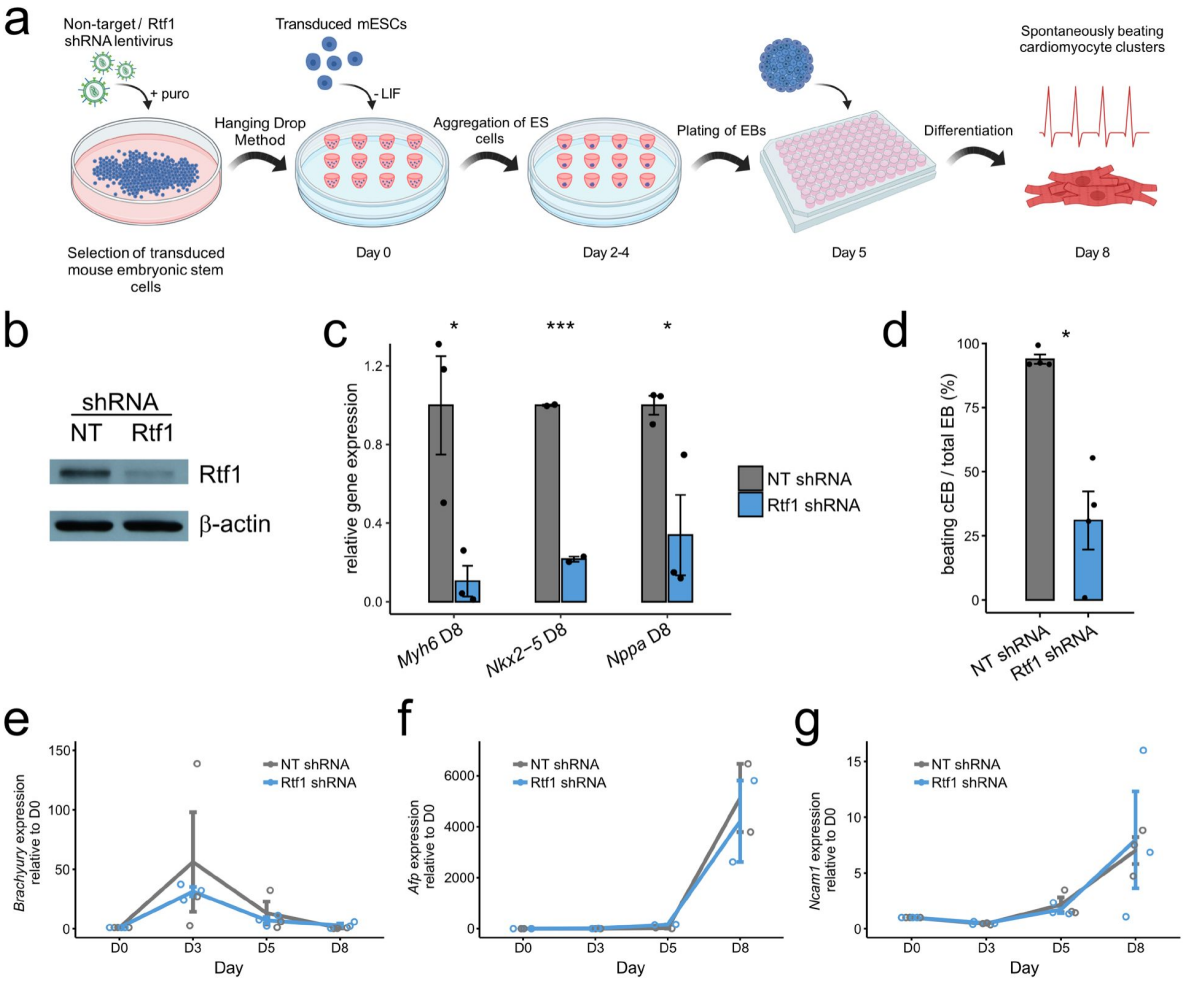

Figure 4.

Failure of cardiac differentiation from Rtf1 knockdown mouse embryonic stem cells.

(a) Diagram of shRNA knockdown of Rtf1 in mouse embryonic stem cells (mESCs) and differentiation of plated EBs. mESCs were transduced with non-target control or Rtf1 shRNA lentivirus and selected with puromycin. Transduced mESCs were grown into embryoid bodies (EBs) using the hanging drop method which were then plated to examine differentiation into cell types including beating cardiomyocyte clusters. (b) Western blot verifying reduction of Rtf1 protein in Rtf1 shRNA mESCs (~70% reduced based on densitometry) compared to unchanged level of the loading control protein β-actin. Image is representative of three independent experiments. (c) qPCR analysis of *Myh6*, *Nkx2-5*, and *Nppa* expression in non-target control (NT) and Rtf1 shRNA plated EBs. Data are normalized to the mean expression level in NT shRNA samples and error bars indicate the standard error of the mean. (d) Mean percentage (± standard error) of plated EBs exhibiting beating cardiomyocyte differentiation (cEBs) in NT and Rtf1 shRNA plated EBs. (e) qPCR analysis of *Brachyury* expression in non-target control (NT) and Rtf1 shRNA plated EBs. (f) qPCR analysis of *Afp* expression in non-target control (NT) and Rtf1 shRNA plated EBs. (g) qPCR analysis of *Ncam1* expression in non-target control (NT) and Rtf1 shRNA plated EBs. Data in e, f, and g are normalized to the mean expression level in NT shRNA samples at day 0 (D0) and error bars indicate the standard error of the mean. *: $p < 0.05$, ***: $p < 0.001$.

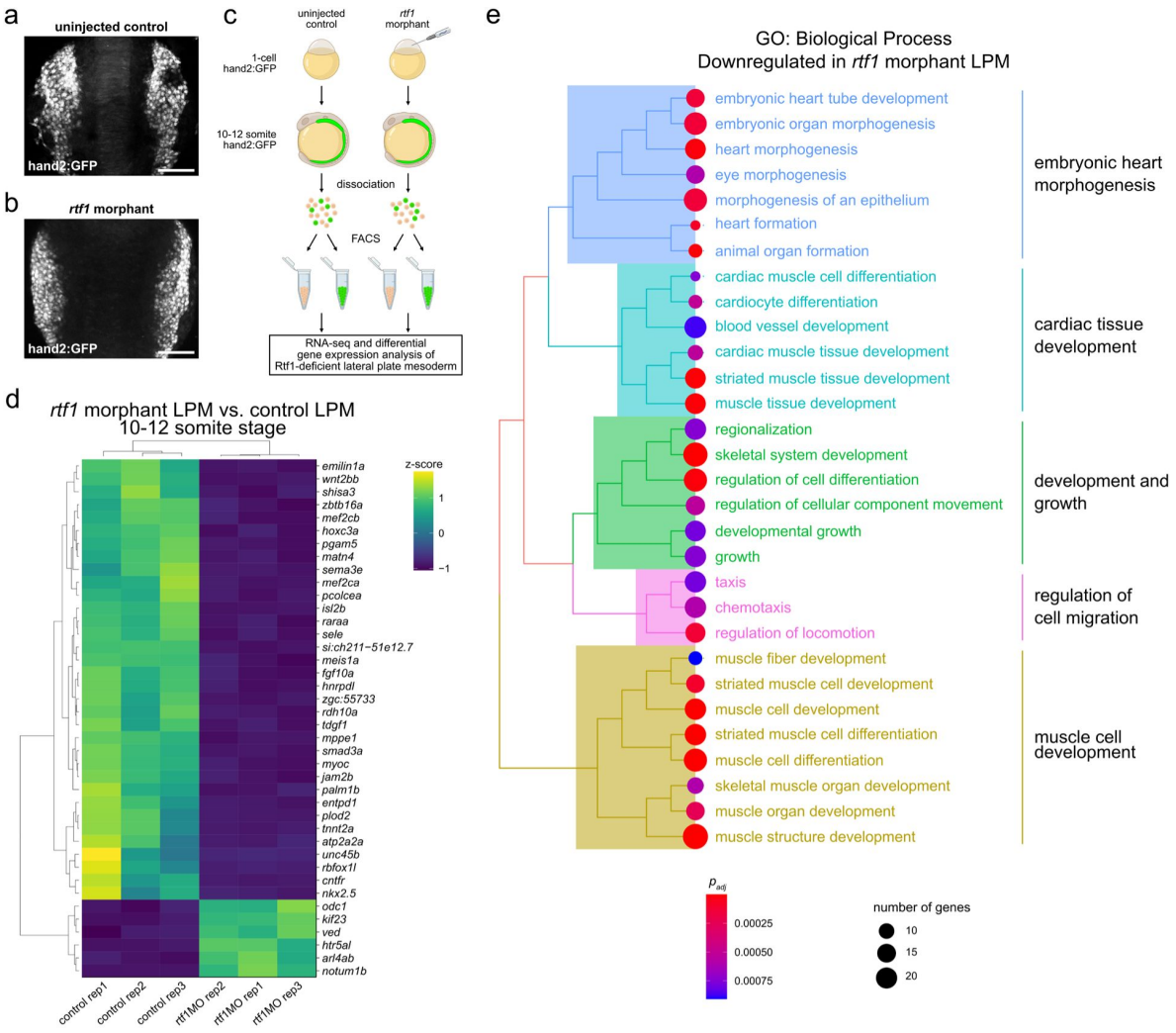

Figure 5.

Dysregulation of cardiac gene expression in the Rtf1 deficient lateral plate mesoderm.

(a,b) Projections of confocal z-stacks of hand2:GFP signal in 11 somite stage uninjected control (a) and *rtf1* morphant (b) embryos. Scale bars represent 100 µm. (c) Diagram of lateral plate mesoderm (LPM) FACS RNA-seq experiment. Transgenic hand2:GFP embryos were injected at the 1-cell stage with *rtf1* morpholino and grown to the 10-12 somite stage prior to dissociation and sorting based on GFP expression. RNA isolated from GFP-positive LPM cells was subjected to RNA-seq and compared to RNA from uninjected embryo LPM cells. (d) Heatmap of scaled expression levels of the top 40 genes most significantly differentially expressed (based on *p*-value) between uninjected control and *rtf1* morphant hand2:GFP-positive LPM. High and low z-scores represent high and low gene expression, respectively. (e) Hierarchical clustering of Biological Process gene ontology (GO) terms based on semantic similarity for the genes most significantly downregulated in the *rtf1* morphant hand2:GFP-positive LPM.

based on the assumption that undifferentiated ALPM, intermediates, and more highly differentiated derivatives are present at this developmental timepoint (**Figure 6C**). In support of this notion, we observed a gradual increase in the expression levels of cardiac markers including *ttn.2*, *mef2cb*, *tnnt2a*, *ryr2b*, and *myh7bb* along the ALPM-cardiac developmental trajectory in control embryos (**Fig 6D,E**). Interestingly, Rtf1 deficient embryos properly downregulate expression of the ALPM marker *sema3e*, maintain the expression of *gata5* in cardiac precursors, and initiate expression of the early cardiac lineage genes *bmp6*, *ttn.2*, and *mef2cb* (**Fig 6D-G**). However, loss of Rtf1 activity prevented the activation of cardiac lineage-restricted markers that are normally expressed later in the ALPM-cardiac trajectory including *tnnt2a*, *ryr2b*, and *myh7bb* (**Fig 6D-G**). Taken together, these data suggest that Rtf1 activity is required for gene expression changes associated with the progression of cardiac precursors into a more mature state, and that Rtf1-deficient cardiac precursors are arresting in an immature state rather than transfating to a different cell lineage.

Rtf1's Plus3 functional domain is necessary for cardiac progenitor formation

The Rtf1 protein is a multifunctional platform for transcription regulation: it consists of several defined functional domains that mediate its interaction with other transcription regulatory proteins and serve as additional points of contact for these proteins with the transcription apparatus (**Figures 1A**, **7A**)[29,30]. The N-terminal half of Rtf1 notably contains its Histone Modification Domain (HMD), which mediates its interaction with E2 ubiquitin-conjugating enzymes and supports the ubiquitylation of histone H2B K120[31–33]. Rtf1's C-terminus contains the Plus3 domain, which interacts with the C-terminal repeat region of the pausing and elongation factor Spt5[18,19,34], a Pol II-interaction domain that allosterically stimulates transcription elongation[30], and a PAF1C interaction domain.

We found that injection of *rtf1* mRNA with 8 silent mutations that render it morpholino resistant (Rtf1 wt) into *rtf1* morphant embryos was able to robustly rescue the knockdown phenotype and support normal formation of the embryonic heart tube (**Figure 7C**). This knockdown-rescue assay provides an in vivo platform to test whether Rtf1's HMD and Plus3 domains are required for the formation of cardiac progenitors during development. We generated expression constructs for mutant versions of Rtf1 lacking these functional domains and tested whether they could support cardiogenesis in an *rtf1* morphant background. Mutant FLAG-tagged Rtf1 proteins lacking the HMD or Plus3 domain were expressed at levels comparable to Rtf1 wt and properly localized to the nucleus demonstrating that the activity of these proteins was not compromised by instability or mislocalization (**Fig 7B**). Interestingly, we found that Rtf1's HMD was dispensable for its role in cardiac progenitor formation (100% of Rtf1 wt injected *rtf1* morphants with *nkx2.5* signal, n=25 vs. 100% of Rtf1 ΔHMD *rtf1* morphants with *nkx2.5* signal, n=25) (**Figure 7C**). On the other hand, Rtf1's Plus3 domain was essential for supporting the formation of cardiac progenitors (0% of Rtf1 ΔPlus3 *rtf1* morphants with *nkx2.5* signal, n=25) (**Figure 7C**). These data suggest that Rtf1's domains function in a modular fashion for regulating cardiac gene expression and point to potential specific effects on transcriptional pausing and elongation in cardiogenesis.

Rtf1 deficiency disrupts promoter-proximal transcriptional pausing

Given the critical role of Rtf1's Plus3 domain in cardiac progenitor formation, we hypothesized that Rtf1 deficiency may impact the transcriptional activity of RNA Pol II at cardiac genes. We then examined the occupancy of RNA Pol II by performing ChIP-seq on embryos at the 10-12 somite stage, when cardiac progenitors are arising from the lateral plate mesoderm. We compared the occupancy of Pol II between uninjected control and *rtf1* morphant embryos and found that Pol II was substantially diminished at the transcription start site (TSS) of many genes compared to the

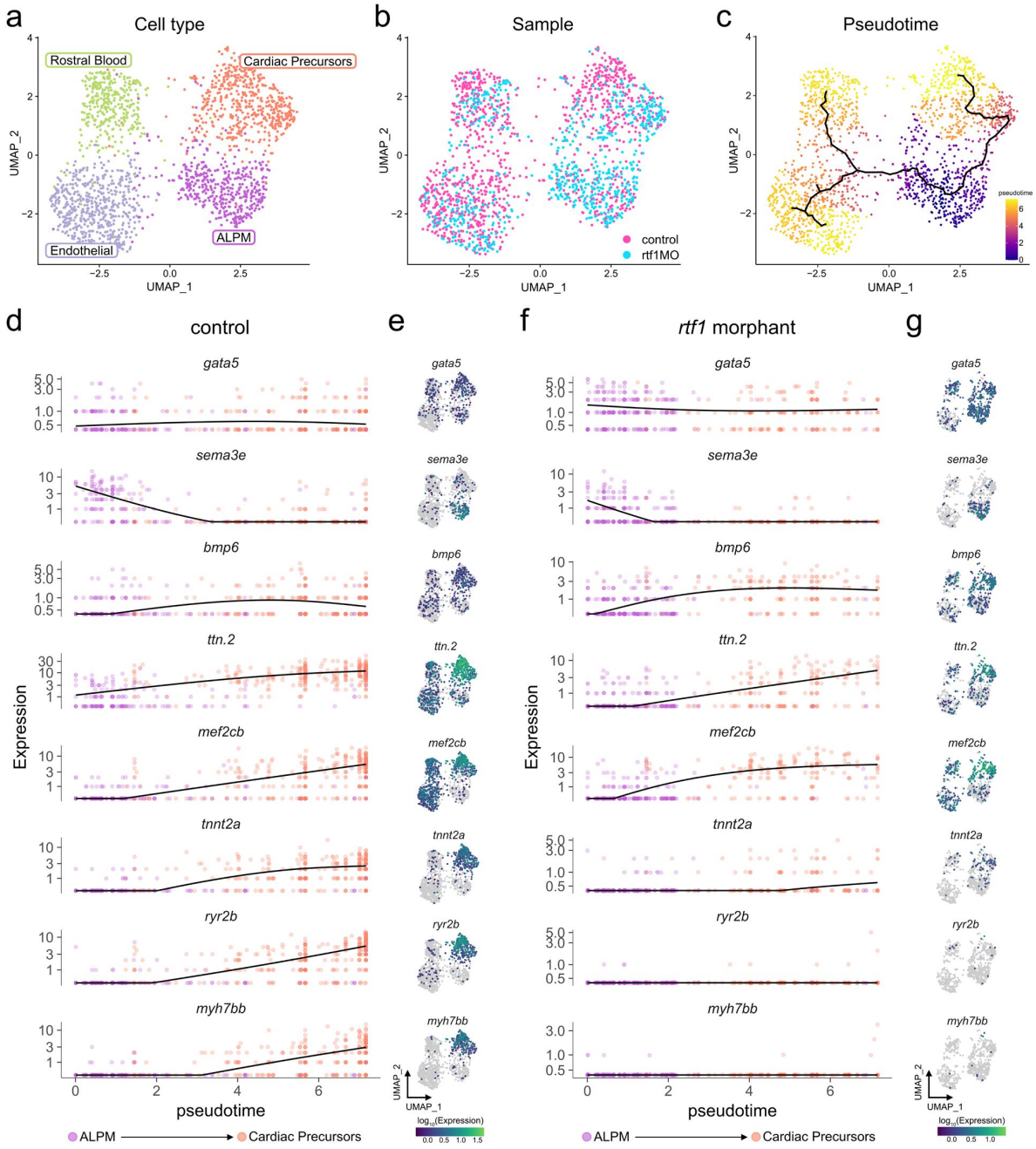

Figure 6.

Rtf1 is required for transcriptional progression of cardiac precursors to a more mature state.

(a-c) UMAP plots of anterior lateral plate mesoderm (ALPM) and derivatives from merged and integrated single cell RNA-seq datasets of 11-12 somite stage uninjected control and *rtf1* morphant embryos. Cells in plots are colored by cell type (a), sample (b), and pseudotime (c). (d,f) Expression dynamics plots displaying expression levels (y-axis) for selected genes in uninjected control (d) and *rtf1* morphant (f) cells over the pseudotime trajectory (x-axis) from ALPM (root) to cardiac precursor state shown in c. (e,g) UMAP plots of uninjected control (e) and *rtf1* morphant (g) cells colored by gene expression levels.

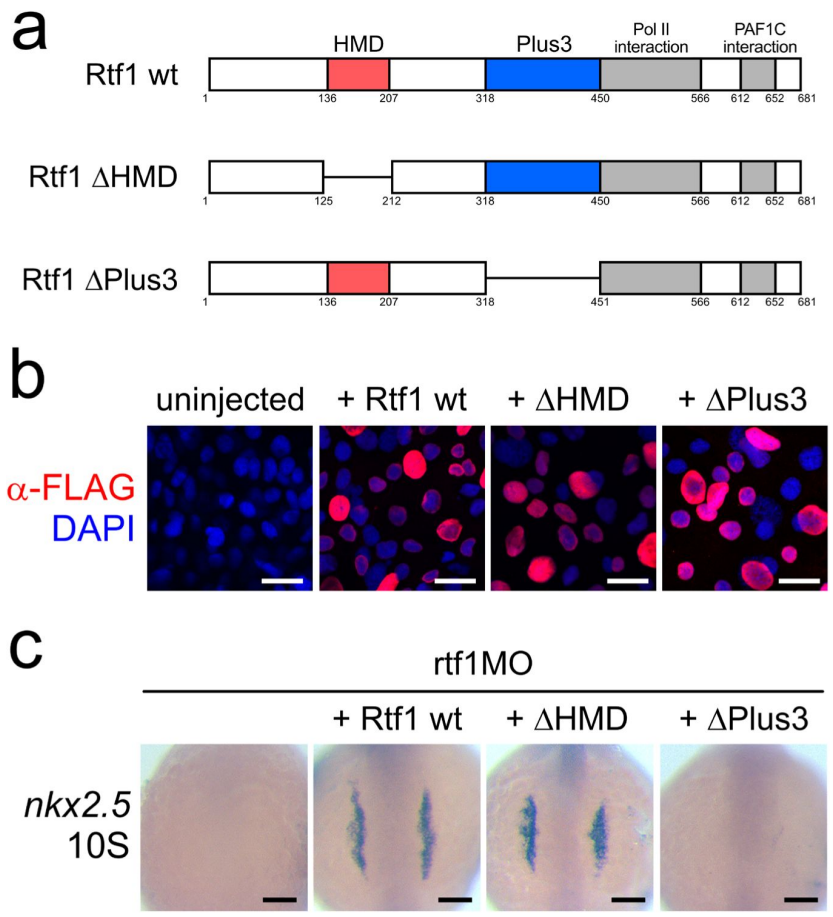

Figure 7.

Rtf1's Plus3 domain is necessary for cardiac progenitor formation.

(a) Schematics of Rtf1 wildtype (wt) and Rtf1 mutant constructs (ΔHMD and ΔPlus3). Domains are indicated by colored boxes (HMD: red, Plus3: blue, PolII and PAF1C interaction: grey). Deleted regions are represented by lines. Numbers refer to the amino acid positions in wildtype Rtf1. (b) Projections of confocal z-stacks of whole mount immunostaining detecting N-terminal FLAG-tagged Rtf1 constructs (red) expressed in 75% epiboly zebrafish embryos. Nuclei are labeled by DAPI staining (blue). Scale bars represent 20 μm. (c) Representative images of RNA in situ hybridization detecting *nkx2.5* expression in 10 somite stage (10S) zebrafish embryos coinjected with *rtf1* morpholino and mRNA encoding Rtf1 wildtype or mutant proteins. Scale bars represent 0.1 mm.

remainder of the gene body (**Figure 8D-I**) and data not shown). Pausing of RNA Pol II at the TSS is a normal aspect of transcription in vertebrates and has been hypothesized to regulate both the timing and expression level of genes[35,36]. To quantify transcriptional pausing at each gene, we calculated the Pause Release Ratio (PRR), a ratio of Pol II occupancy in the promoter proximal region of a gene (−300 to +300) to the occupancy in the downstream region (+301 to +1301). A higher PRR indicates increased pausing of Pol II near the transcription start site (TSS), while a lower PRR signifies that a gene is in a less paused state. We found that on a global scale, pausing is diminished upon loss of Rtf1 activity during embryonic development (**Figure 8A,B**). Out of 6,078 genes with substantial Pol II ChIP signal that were analyzed, 5,799 displayed a significantly decreased PRR (Benjamini–Hochberg false discovery rate of 10%) in Rtf1-deficient embryos compared to 37 genes that displayed a significantly increased PRR (**Figure 8B**). The median PRR in control embryos was 3.34, while 95% (5,793/6,078) of genes had a PRR of less than 3.34 in Rtf1-deficient embryos (**Figure 8A**), reflecting a widespread decrease in promoter proximally paused Pol II. These changes in PRR suggest that Rtf1 is critical for normal levels of pausing at transcriptionally active genes during the developmental window when cardiac progenitor cells arise from the lateral plate mesoderm.

Attenuation of pause release permits cardiogenesis in an Rtf1-deficient background

The transition between the pausing and elongation phases of transcription is regulated by PTEF-b, a complex consisting of a cyclin (T1, T2a, or T2b) and the kinase Cdk9[37–39]. Phosphorylation of the negative elongation factor (NELF) complex, the pausing/elongation factor Spt5, and the C-terminal tail of RNA Pol II by Cdk9 promotes changes in the factors associated with RNA Pol II and triggers a switch to processive elongation[40]. Pharmacological inhibition of Cdk9 with the small molecule flavopiridol can prevent or attenuate pause release, leading to genome-wide increases in pausing at most genes[41]. We explored the effects of Cdk9 inhibition on Rtf1-dependent transcription by performing ChIP-seq for RNA Pol II in flavopiridol-treated *rtf1* morphant embryos at the 10-12 somite stage. Flavopiridol treatment caused a small but highly significant increase in promoter-proximal pausing in an Rtf1 deficient background (**Fig 8A**), with 87.1% of genes with significantly altered PRR values displaying an increase in PRR in flavopiridol-treated *rtf1* morphants compared to untreated *rtf1* morphants (**Figure 8C**). Furthermore, we found that critical LPM and cardiac genes including *hand2*, *gata5*, *aplnrb*, *bmp4*, *nkx2.7*, and *rbfox1l* exhibited reduced promoter-proximal pausing in Rtf1 deficient embryos that was partially rescued by treatment with flavopiridol (**Figure 8D-I**).

If diminished promoter-proximal pausing at genes crucial for cardiogenesis is responsible for the absence of cardiac progenitor formation in Rtf1-deficient embryos, then we would predict that reducing pause release may restore cardiogenesis in an Rtf1-deficient background. We examined expression of the cardiomyocyte marker *myl7* in wild type, *rtf1* mutant, and flavopiridol-treated *rtf1* mutant embryos and found that partial attenuation of Cdk9 activity (5 ug/ml dosage) potently restored cardiomyocyte formation in Rtf1-deficient embryos (**Figure 8J,K**). Similarly, we found that knockdown of Cdk9 by injection of antisense morpholino oligonucleotides[42,43] also restored the formation of *myl7*+ cardiomyocytes in an Rtf1-deficient background (**Figure 8L,M**), confirming that decreased activity of the Cdk9 kinase is responsible for rescuing the *rtf1* cardiac phenotype. These dramatic results suggest that the primary role of Rtf1 in cardiac progenitor formation is regulating promoter-proximal pausing of critical members of the cardiac gene program. Altogether, our data support a model in which loss of Rtf1 causes excessive pause release, a dysregulation of the cardiac transcription program, and an arrest in the developmental trajectory of cardiac progenitors prior to the formation of mature cardiomyocytes.

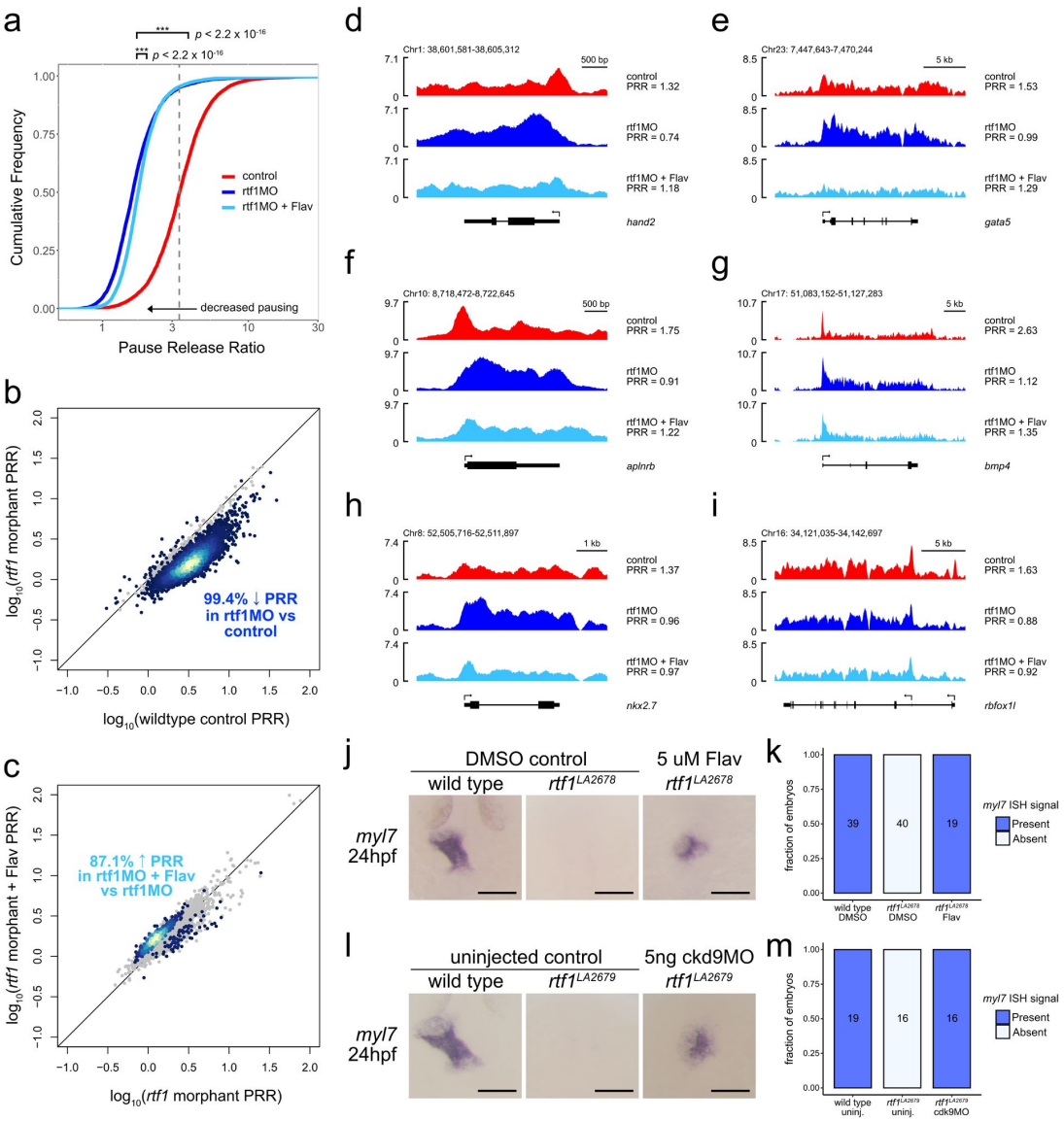

Figure 8.

Rtf1-dependent transcriptional pausing regulates cardiogenesis.

(a) Cumulative frequency plot of pause release ratios (PRRs) of 6,078 genes with substantial RNA Pol II signal. PRRs in control (red) and flavopiridol treated *rtf1* morphant (light blue) samples differed significantly from untreated *rtf1* morphant PRRs (blue). ***: $p < 2.2 \times 10^{-16}$; Welch's paired two-tailed *t*-test. The median PRR of control samples is indicated by a vertical grey dashed line. (b) Dot plot comparing PRRs in controls and *rtf1* morphants. Each dot represents the PRR values for a single gene that differ significantly (colored point) or are not significantly different (grey point) between controls and *rtf1* morphants. (c) Dot plot comparing PRRs in *rtf1* morphants and flavopiridol-treated *rtf1* morphants. Each dot represents the PRR values for a single gene that differ significantly (colored point) or are not significantly different (grey point) between *rtf1* morphants and flavopiridol-treated *rtf1* morphants. Dot colors in b and c are based on the density of points, with lighter colors indicating more dense points. (d-i) Representative ChIP-seq tracks displaying RNA Pol II read densities (y-axis) for control (red), *rtf1* morphant (blue), and flavopiridol-treated *rtf1* morphant (light blue) embryos at cardiac mesoderm-related genes including *hand2* (d), *gata5* (e), *aplnrb* (f), *bmp4* (g), *nkx2.7* (h), and *rbfox1l* (i). (j) Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf control and Rtf1 deficient (± flavopiridol) zebrafish embryos. (k) Quantification of *myl7* signal intensity in control and Rtf1 deficient (± flavopiridol) embryos at 24 hpf. Numbers on bars indicate the number of embryos analyzed. (l) Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf control and Rtf1 deficient (± *cdk9* morpholino) zebrafish embryos. (m) Quantification of *myl7* signal intensity in control and Rtf1 deficient (± *cdk9* morpholino) embryos at 24 hpf. Numbers on bars indicate the number of embryos analyzed.

Discussion

In this study we investigated the molecular mechanism of the PAF1C member Rtf1's role in cardiac progenitor formation. We found that Rtf1 plays a critical and conserved role in cardiogenesis and cardiac gene expression in vertebrate and mammalian models. Zebrafish *rtf1* mutants and mouse Rtf1 cardiac mesoderm *Rtf1* conditional knockouts do not properly initiate expression of cardiac transcription factors in the LPM and fail to generate mature cardiomyocytes. Using single cell sequencing, we found that Rtf1 deficient embryos do generate immature cardiac progenitors, but that these progenitors fail to express genes associated with progression to a more mature state. Excitingly, we also showed that the Rtf1 protein can function in a modular fashion and that its Plus3 domain, a site of interaction with the pausing/elongation factor Spt5, is essential for supporting cardiac development. In line with this finding, we showed that loss of Rtf1 diminishes promoter-proximal pausing of RNA Polymerase II, and that partially restoring normal pausing by chemical inhibition of pause release can support cardiogenesis in an Rtf1-deficient background.

Despite having a nearly genome-wide decrease in promoter-proximal pausing of RNA Polymerase II, Rtf1-deficient embryos display many upregulated and downregulated genes. How exactly pausing regulates gene expression is not fully understood and likely involves both general and gene/locus-specific mechanisms[35,44]. By limiting the release of RNA Polymerase II into the elongation phase of transcription, pausing can restrain gene expression. This inhibitory role of pausing on transcription allows for another layer of transcription regulation and has been shown to modulate the expression of heat shock and immune stimulus-induced genes[45–47].

Pausing-dependent restraint of gene expression is also important for the timing of gene expression in a developmental context. Supporting this notion, Spt5 and NELF-deficient zebrafish embryos exhibit aberrant upregulation of TGFβ signaling and a resulting inhibition of hematopoiesis[43]. Intriguingly, the PAF1C has also been shown in mammalian cells to restrain transcription from a subset of promoters by inhibiting nearby enhancer activation[48].

On the other hand, pausing also positively regulates gene expression in some contexts. At genes where nucleosome formation is favored in promoter regions, pausing of RNA Polymerase II can promote a chromatin architecture that is permissive to transcription by competing with nucleosomes for promoter occupancy[49,50]. This mechanism has been suggested to support expression of the IFN-gamma pathway during zebrafish development[43]. Pausing may also positively regulate gene expression during neural crest development. Loss of Paf1 activity disrupts neural crest differentiation, but this defect can be partially rescued by inactivation of Cdk9. Interestingly, zygotic loss of Cdk9 alone is sufficient to expand the neural crest population in zebrafish embryos, suggesting that promoter-proximal pausing promotes the neural crest gene program[51].

Recent work using a rapid protein degradation system in human cell culture has shed light on the mechanisms by which the PAF1C regulates transcriptional pausing and gene expression outputs[52]. In DLD-1 cells, acute depletion of PAF1 destabilizes the PAF1C and stimulates pause release. Loss of PAF1 also decreases chromatin recruitment of INTS11, a component of the Integrator-PP2A complex (INTAC) which interacts with PAF1C and counteracts the activity of P-TEFb. This disruption in the balance of INTAC and P-TEFb activities at transcribed genes results in hyperphosphorylation of SPT5 and likely explains the increase in pause release caused by PAF1 depletion. Similar to our finding that Rtf1-deficiency positively and negatively regulates gene expression, depletion of PAF1 resulted in both upregulation and downregulation of many genes. Intriguingly, the gene expression changes caused by PAF1 depletion correlated with the release frequency and processivity of RNA Polymerase II at a given gene, with genes exhibiting low levels of release frequency tending to display decreases in their expression level. These observations suggest that the combined regulatory functions of pausing and elongation determine the transcriptional output of PAF1C-regulated genes.

In conclusion, our study reveals that the cardiac gene program is highly sensitive to loss of promoter-proximal pausing. Future studies focused on the chromatin architecture of cardiac gene promoters and other regulatory regions, and the dynamics of pause release and elongation at cardiac genes are needed to provide a complete picture of the manner by which promoter-proximal pausing fosters their expression.

Materials and Methods

Zebrafish

Zebrafish colonies were cared for and bred under standard conditions[53]. Developmental stages of zebrafish embryos were determined based on somite number or hours of development post fertilization at 28.5 °C [53]. Fish husbandry and experiments were performed according to the Institutional Approval for Appropriate Care and Use of Laboratory Animals by the UCLA Institutional Animal Care and Use Committee (Protocols ARC-2000-051-TR-001 and BUA-2018-195-002-CR).

CRISPR mutagenesis

Template DNA for *rtf1* exon 3 guideRNA (rtf1-e3 gRNA) synthesis was amplified with KOD polymerase (MilliporeSigma) using the pXTW-gRNA plasmid and the primers rtf1-e3-F (5'-TAATACGACTCACTATAGGAAGAAGGGGAAACCGAGCAAGTTTTAGAGCTAGAAATAGC-3') and gRNA-R (5'-AAAAAAAGCACCGACTCGGTGCCAC-3'). rtf1-e3 gRNA (target sequence: 5'-GGAAGAAGGGGAAACCGAGCAA-3') was synthesized using the MEGAshortscript T7 Transcription Kit (ThermoFisher Scientific) and purified using the RNeasy Mini Kit (Qiagen). Zebrafish embryos were injected at the 1-cell stage with 300 pg of Cas9 mRNA and 200 pg of rtf1-e3 gRNA. Injected embryos were raised to adulthood and then outcrossed to wild type AB fish to establish lines carrying specific *rtf1* mutations.

Mice

All mice were maintained according to the Guide for the Care and Use of Laboratory Animals published by the US National Institute of Health. Housing and experiments were performed according to the Institutional Approval for Appropriate Care and Use of Laboratory Animals by the UCLA Institutional Animal Care and Use Committee (Protocols ARC-2020-045-TR-001 and ARC-2020-043-TR-001). ES cells were obtained from KOMP and used to rederive the $Rtf1^{tm1a(KOMP)Wtsi}$ mouse line. This line was then crossed to a ROSA26::FLPe strain[54] (The Jackson Laboratory, JAX stock #009086) to create an Rtf1 conditional knockout allele where exon 3 of Rtf1 is flanked by two loxP sites ($Rtf1^{flox}$). These mice were crossed into the $Mesp1^{tm2(cre)Ysa}$ background[28] to produce a strain capable of generating Mesp-Cre^{+};$Rtf1^{flox/flox}$ offspring.

Whole mount in situ hybridization

Whole-mount in situ hybridization was performed as described previously[55]. Embryos for in situ hybridization were fixed in 4% paraformaldehyde in 1xPBS. After a brief proteinase K digestion, embryos were incubated with digoxigenin-labeled antisense RNA probes overnight at 65-70°C. Probes were detected using an anti-digoxigenin antibody conjugated with alkaline phosphatase (Roche 11093274910). The zebrafish antisense RNA probes used in this study include *myl7*, *nkx2.5*, *mef2ca*, *tbx20*, and *tbx5a*. The mouse probes used include *MLC2v*, *Tbx20*, and *Nkx2.5*.

Genotypes of zebrafish embryos were confirmed by PCR with primers:

rtf1-e3-F: 5'-TACAGTGACATCCACCATGCTG-3'

rtf1-e3-R: 5'-GTAAATAGGGGAGAAACTAATTGAAG-3'.

Genotypes of mouse embryos were confirmed by PCR with primers:

Rtf1-loxP-F: 5'-GACTGAGAAGGGAAATCTTTGAAAC-3'

Rtf1-loxP-R: 5'-CTAGTCCTTGAACAAGTTCAGCT-3'

Cre-F: 5'-GTAGCTGATGATCCGAATAACTAC-3'

Cre-R: 5'-ATCCAGGTTACGGATATAGT-3'.

Constructs

To prevent complementarity between *rtf1* morpholino and mRNA sequences, PCR was used to introduce 8 silent mutations to the morpholino target sequence using pCS2/3XFLAG-rtf1 as a template[21]. The morpholino-resistant wild type *rtf1* sequence was cloned into pCS2/3XFLAG to produce an N-terminally FLAG-tagged morpholino-resistant rtf1 construct (pCS2/3XFLAG-rtf1-mr). Rtf1 deletion mutant sequences were amplified from pCS2/3XFLAG-rtf1-mr by SOE-PCR to produce Rtf1ΔHMD (missing a.a. 126-211 of Rtf1) and Rtf1ΔPlus3 (missing a.a. 319-450 of Rtf1), which were subsequently cloned into pCS2/3XFLAG to produce N-terminally tagged constructs.

Microinjections

0.5 ng of *rtf1* morpholino (rtf1MO; 5'-CTTTCCGTTTCTTTACATTCACCAT-3') was injected at the 1-cell stage[21,56] along with 1 ng of *p53* morpholino (5'-GCGCCATTGCTTTGCAAGAATTG-3') to prevent p53-dependent apoptosis[57]. For *cdk9* knockdown rescue experiments, 5 ng of cdk9 morpholino (cdk9MO; 5'-ACACACAAACATCAAATACTCACCC-3')[42,43] was injected into an incross of $rtf1^{LA2679}$ heterozygotes at the 1-cell stage. 150-200 pg of *rtf1* or *rtf1* deletion mutant mRNAs were co-injected with *rtf1* and *p53* morpholinos at the 1-cell stage for mRNA rescue experiments.

Mouse ESC culture and differentiation

E14TG2a mESCs (ATCC) were cultured on 0.1% gelatin covered plates without feeder cells. The feeder-free mESCs were maintained in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 15% fetal bovine serum (FBS; Invitrogen, USA), 1% non-essential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 1% (v/v) penicillin-streptomycin-Glutamine (ThermoFisher Scientific), and 10 μg/μl recombinant mouse leukemia inhibitory factor (LIF; Sigma-Aldrich) in a 5% CO2 incubator at 37°C. The ESC culture medium was changed every two days. mESCs were subcultured every 3 days using 0.25% trypsin-EDTA (Invitrogen). mESCs were differentiated in hanging drops (600 cells/20 μL) without LIF under non-adherent conditions. Day 5 embryoid bodies (EBs) were plated onto 0.1% gelatin covered plates. Media was changed every two days during adherent culture.

Lentivirus transduction

To knockdown Rtf1 activity in mESCs, Rtf1 or NT (non-target) shRNA lentivirus harboring puromycin resistance were purchased from Sigma-Aldrich. mESCs were transduced with the indicated virus at a multiplicity of infection of 1 and selected in 1.2 μg/ml puromycin. Rtf1 shRNA knockdown efficacy was tested by Western blotting analysis that detected Rtf1 protein level. Antibodies used were anti-Rtf1 (Bethyl Labs A301-329A) and anti-β-actin (Sigma A1978). Relative protein levels were determined by densitometry measurements using Adobe Photoshop.

Quantitative RT-PCR

Total RNA from NT and shRNA transduced EBs was extracted using TRIzol RNA isolation reagents (ThermoFisher Scientific). cDNA was synthesized using iScript cDNA Synthesis Kit from Bio-Rad. Real-time quantitative PCR was carried out using the Roche LightCycler 480 Real-Time PCR System. Primers for qPCR are listed in Supplementary Table 1.

Antibody staining

Embryos injected with mRNA encoding FLAG-tagged zebrafish Rtf1 constructs were fixed in 4% PFA in PBS at 75% epiboly. The fixed embryos were incubated in primary antibody (1:50 mouse anti-FLAG M2, F1804, Sigma) in blocking solution (10% goat serum in PBT) for 2 hours at room temperature followed by detection with fluorescent secondary antibody (1:200 anti-mouse IgG-Alexa Fluor 555, A-31570). Nuclei were stained with DAPI and embryos were embedded in 1% low-melt agarose and imaged on a Zeiss LSM800 confocal microscope.

Western Blotting

Embryos were lysed in Rubinfeld's Lysis Buffer at 1 day post fertilization as previously described[21]. Two embryo equivalents were loaded per lane of a 8% polyacrylamide gel. Following electrophoresis, proteins were transferred to a nitrocellulose membrane and detected with antibodies against Rtf1 (Bethyl Labs A301-329A, 1:4000) and beta-actin (Sigma A1978, 1:5000). ImageJ was used to perform densitometry comparing the levels of Rtf1 protein between samples using beta-actin as the control.

Chemical treatment

Flavopiridol (Cayman Chemicals) was dissolved in DMSO at a stock concentration of 10 mg/ml. Zebrafish embryos were treated with 1.375 – 5 ug/ml flavopiridol diluted in E3 buffer beginning at 60-70% epiboly. For *rtf1* mutant rescue experiments, genotypes were confirmed using the primers rtf1-e3-F and rtf1-e3-R.

Facs

10-12 somite stage TgBAC(hand2:EGFP)^{pd24}[24] uninjected and *rtf1* morpholino-injected zebrafish embryos[58] were nutated in 10 mg/ml protease (Protease from *Bacillus licheniformis*, Sigma P5380) in DMEM/F12 media for 30 minutes at 4°C and pipetted to dissociate cells. Debris was removed by filtering through a 40 um cell strainer and cells were then pelleted by centrifugation at 300 rcf for 5 minutes at 4°C. Cells were resuspended in cold FACS buffer (1xPBS, 2% FBS, 1mM EDTA) and refiltered immediately prior to sorting on a BD FACSAria II using a 100 um nozzle. Dead cells were excluded by staining with 0.1 µg/ml DAPI. Cells from batches of 100 embryos were sorted directly into a mixture of 350 ul Buffer RA1 and 3.5 ul 2-mercaptoethanol from the NucleoSpin RNA kit (Machery-Nagel) for RNA isolation.

mRNA-seq

RNA was purified from FACS-sorted cells with the NucleoSpin RNA kit (Machery Nagel). Approximately 5-10 ng of input RNA was used to produce libraries with the NEBNext Single Cell/Low Input RNA Library Prep Kit for Illumina (NEB) which were sequenced on an Illumina HiSeq 3000 to produce 50 bp single-end reads. Reads from triplicate samples were mapped to the danRer7 genome using Tophat v2.1.1[59] using default parameters. Mapped reads were counted with FeatureCounts v2.0.0[60] and differential expression analysis was carried out with DESeq2 v1.32.0[61]. Gene ontology enrichment was examined using the R package clusterProfiler[62] in R v4.1.0. For gene ontology analysis, the most significantly downregulated genes in *rtf1* morphant LPM were selected as those genes that were downregulated among the 300 most significantly differentially expressed genes based on adjusted *p*-value.

ChIP-seq

Embryos for ChIP were dissociated with 4 mg/ml Collagenase Type IV and 0.25% Trypsin at room temperature with intermittent pipetting. Dissociation was stopped with 10% FBS in DMEM and cells were washed once with PBS. Fixation and preparation of chromatin was carried out using the Covaris truChIP Chromatin Shearing Kit and 7 minutes of sonication in a Covaris E220 ultrasonicator using the recommended settings. Immunoprecipitations were carried out as previously described[63] on approximately 5 ug of chromatin (minimum 2.5 ug, maximum 8.8 ug) using 5 ul of total RNA Pol II antibody (Cell Signaling D8L4Y). Libraries were prepared using the NEBNext Ultra II DNA Library Kit for Illumina (NEB) and sequenced with an Illumina Novaseq 6000 system to produce 50 bp paired-end reads. Reads from triplicate samples were mapped to the danRer11 genome using bowtie2 v2.4.2[64] (options: --local -X 1000) and read densities of UCSC RefSeq gene regions were quantified using the package csaw v1.24.3[65,66] in R v4.0.3. For our pausing analyses, we defined the transcription start site (TSS) region as −300 bp to +300 bp surrounding the TSS. RNA Pol II ChIP-seq read densities for the TSS regions of genes displayed a bimodal distribution which was assumed to be two partially overlapping distributions of genes with high and low RNA Pol II TSS signal. We examined those genes with a substantial level of RNA Pol II signal in the TSS region of control samples, defined as those genes present in the upper mode of the distribution of log-transformed TSS read densities based on a cutoff for type-I error of 10% using the R cutoff package v0.1.0 (*https://github.com/choisy/cutoff*). The read density in the TSS region was divided by the read density in the region extending from +301 to +1301 bp with respect to the TSS to calculate the pause release ratio (PRR). ChIP-seq data was visualized in Python 3.10.0 using the plotting tool SparK v2.6.2[67].

Nuclei isolation for single cell sequencing

Nuclei were isolated from 200 11-12 somite stage zebrafish embryos for 10x Genomics Multiome sequencing using recommended procedures with slight modifications. Embryos were deyolked in Deyolking buffer (55 mM NaCl, 1.8 mM KCl, 1.25 mM NaHCO3) and then embryonic cells were partially dissociated by pipetting in DMEM/F12 media (Gibco 21041025). Cells were pelleted and then lysed on ice for 5 minutes in 100 µl of chilled 0.1X Lysis buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.01% Tween-20, 0.01% IGEPAL CA-630, 0.001% digitonin, 1% BSA, 1 mM DTT). 1 ml of Wash buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween-20, 1% BSA, 1 mM DTT, 0.1 U/µl Sigma Protector RNase Inhibitor) was then added to the cell lysis and the nuclei were pelleted by centrifuging at 500xg for 5 minutes at 4°C. Two more washes with 1 ml of Wash buffer were performed. Nuclei were then resuspended in 50 µl of Nuclei Dilution Buffer (1X 10x Genomics Single Cell Multiome Nuclei Buffer, 1 mM DTT, 1 U/µl Sigma Protector RNase Inhibitor) and filtered with a 40 µm Flowmi cell strainer. Approximately 10,000 nuclei per sample were targeted for capture with the 10x Chromium Single Cell Multiome ATAC + Gene Expression system by the UCLA Technology Center for Genomics & Bioinformatics.

Single cell sequencing analysis

10x Chromium Single Cell Multiome ATAC + Gene Expression libraries were prepared by the UCLA Technology Center for Genomics & Bioinformatics following the manufacturer's recommended protocols and sequenced with an Illumina Novaseq 6000 system. Reads from RNA-seq and ATAC-seq assays were aligned to the zebrafish GRCz11 genome with the Lawson laboratory's improved transcriptome annotation[68] using 10x Genomics' Cell Ranger ARC software v2.0.0 with default parameters. Mapped reads from single cell RNA-seq and ATAC-seq assays were analyzed in R v4.2.2 using Seurat v4.3.0[69]. Good quality cells were retained by filtering using Seurat with the following thresholds: 15000 > RNA UMI > 2000, Genes > 1500, 20000 > ATAC UMI > 5000, TSS enrichment score > 3.5. Filtered control and *rtf1* morphant cell datasets were integrated using Harmony v0.1.1[70], and preliminary dimensional reduction, cell clustering, and marker gene identification were carried out in Seurat. Clusters of cells that exhibited enriched expression of mitochondrial and ribosomal transcript expression levels but lacked enrichment for other genes were considered abnormal cells and were manually removed from the dataset. Following manual

filtering, a final dataset of 14340 control and 13839 *rtf1* morphant cells was again integrated with Harmony and multimodal dimensional reduction (weighted-nearest neighbor method with RNA and ATAC assays), cell clustering, and marker gene identification were carried out in Seurat. Cluster cell number proportions and gene expression levels were visualized with Seurat and with published code[71]. Based on marker gene expression, cells that were members of clusters corresponding to anterior lateral plate mesoderm (*sema3e+*), endothelium (*flt4+*), cardiac precursor (*mef2cb+*), and rostral blood (*spi1b+*) were extracted, independently integrated with Harmony, and dimensional reduction (RNA assay only), cell clustering, and marker gene identification were carried out in Seurat. Pseudotime analysis of Harmony/Seurat cell embeddings and plotting of gene expression over pseudotime was performed with Monocle v3.0[72].

Statistical analysis

Statistical significance of gene expression changes in qPCR analyses was determined based on 2 (for *Nkx2-5* and *Afp* expression) or 3 (for *Myh6*, *Nppa*, *Brachyury*, and *Ncam1* expression) independent experiments using a two-tailed Student's *t*-test with a *p*-value of less than 0.05 considered to be significant. Statistical significance of the difference between beating cardiomyocyte cluster formation in control and Rtf1 shRNA mESCs was determined based on 4 independent experiments using a two-tailed Student's *t*-test with a *p*-value of less than 0.05 considered to be significant. To compare PRR means between uninjected, *rtf1* morphant, and flavopiridol-treated *rtf1* morphant embryos, mean PRR values were first log-transformed (log10) and then a Welch's paired two-tailed *t*-test was carried out with a *p*-value of less than 0.05 considered to be significant. Significantly changed PRRs for individual genes were determined based on 3 independent replicates per condition using the rowttests function in the R package genefilter v1.72.1 to perform two-tailed *t*-tests. The p.adjust function in the R stats package v4.0.3 using a Benjamini and Hochberg (false discovery rate) correction for multiple comparisons was then applied with adjusted *p*-values of less than 0.1 considered to be significant.

Acknowledgements

We thank the UCLA Technology Center for Genomics & Bioinformatics for technical assistance with Next Generation Sequencing and the Broad Stem Cell Research Center Flow Cytometry Core Facility for technical assistance with cell sorting. **Figures 3A**, **4A**, and **5C** were created with BioRender.com. This work was supported by funding from the NIH/NHLBI to J.-N.C. (R01HL155905 and R01HL140472).

Author Contributions

A.D.L.: study design, data acquisition, analysis, and interpretation, manuscript writing. F.L.: study design, data acquisition, analysis, and interpretation. L.T.: data acquisition. Z.Y.S.H.: data acquisition. B.K.: data acquisition. Z.T.: data acquisition. A.E.: data acquisition. M.P.: data analysis. H.N.: study design and data interpretation. A.N.: study design and data interpretation. J.-N.C.: study design, data interpretation, manuscript writing. All authors approved the final manuscript.

Competing Interests

The authors declare no competing interests.

Data Availability Statement

The RNA-seq, ChIP-seq, and single cell-seq data presented in this study are available in the NCBI SRA (BioProject PRJNA1015262).

References

1 Bruneau B. G (2013) **Signaling and transcriptional networks in heart development and regeneration** *Cold Spring Harb Perspect Biol* **5** https://doi.org/10.1101/cshperspect.a008292

2 Paige S. L., Plonowska K., Xu A., Wu S. M (2015) **Molecular regulation of cardiomyocyte differentiation** *Circ Res* **116**:341–353 https://doi.org/10.1161/CIRCRESAHA.116.302752

3 Rana M. S., Christoffels V. M., Moorman A. F (2013) **A molecular and genetic outline of cardiac morphogenesis** *Acta Physiol (Oxf* **207**:588–615 https://doi.org/10.1111/apha.12061

4 Lu F., Langenbacher A. D., Chen J. N (2016) **Transcriptional Regulation of Heart Development in Zebrafish** *J Cardiovasc Dev Dis* **3** https://doi.org/10.3390/jcdd3020014

5 Schoenebeck J. J., Keegan B. R., Yelon D (2007) **Vessel and blood specification override cardiac potential in anterior mesoderm** *Dev Cell* **13**:254–267 https://doi.org/10.1016/j.devcel.2007.05.012

6 Mao L. M. F., Boyle Anderson E. A. T., Ho R. K (2021) **Anterior lateral plate mesoderm gives rise to multiple tissues and requires tbx5a function in left-right asymmetry, migration dynamics, and cell specification of late-addition cardiac cells** *Dev Biol* **472**:52–66 https://doi.org/10.1016/j.ydbio.2021.01.007

7 Duong T. B., Holowiecki A., Waxman J. S (2021) **Retinoic acid signaling restricts the size of the first heart field within the anterior lateral plate mesoderm** *Dev Biol* **473**:119–129 https://doi.org/10.1016/j.ydbio.2021.02.005

8 Sorrell M. R., Waxman J. S (2011) **Restraint of Fgf8 signaling by retinoic acid signaling is required for proper heart and forelimb formation** *Dev Biol* **358**:44–55 https://doi.org/10.1016/j.ydbio.2011.07.022

9 Hota S. K., et al. (2019) **Dynamic BAF chromatin remodeling complex subunit inclusion promotes temporally distinct gene expression programs in cardiogenesis** *Development* **146** https://doi.org/10.1242/dev.174086

10 Takeuchi J. K., et al. (2011) **Chromatin remodelling complex dosage modulates transcription factor function in heart development** *Nat Commun* **2** https://doi.org/10.1038/ncomms1187

11 Lickert H., et al. (2004) **Baf60c is essential for function of BAF chromatin remodelling complexes in heart development** *Nature* **432**:107–112 https://doi.org/10.1038/nature03071

12 Lou X., Deshwar A. R., Crump J. G., Scott I. C (2011) **Smarcd3b and Gata5 promote a cardiac progenitor fate in the zebrafish embryo** *Development* **138**:3113–3123 https://doi.org/10.1242/dev.064279

13 Takeuchi J. K., Bruneau B. G (2009) **Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors** *Nature* **459**:708–711 https://doi.org/10.1038/nature08039

14 Endoh M., et al. (2004) **Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro** *Mol Cell Biol* **24**:3324–3336 https://doi.org/10.1128/MCB.24.8.3324-3336.2004

15 Hartzog G. A., Wada T., Handa H., Winston F (1998) **Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae** *Genes Dev* **12**:357–369 https://doi.org/10.1101/gad.12.3.357

16 Keegan B. R., et al. (2002) **The elongation factors Pandora/Spt6 and Foggy/Spt5 promote transcription in the zebrafish embryo** *Development* **129**:1623–1632 https://doi.org/10.1242/dev.129.7.1623

17 Wada T., et al. (1998) **DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs** *Genes Dev* **12**:343–356 https://doi.org/10.1101/gad.12.3.343

18 Wier A. D., Mayekar M. K., Heroux A., Arndt K. M., VanDemark A. P (2013) **Structural basis for Spt5-mediated recruitment of the Paf1 complex to chromatin** *Proc Natl Acad Sci U S A* **110**:17290–17295 https://doi.org/10.1073/pnas.1314754110

19 Chen J. J., et al. (2020) **Spt5 Phosphorylation and the Rtf1 Plus3 Domain Promote Rtf1 Function through Distinct Mechanisms** *Mol Cell Biol* **40** https://doi.org/10.1128/MCB.00150-20

20 Hu S., et al. (2021) **SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains the enhancer landscape** *Mol Cell* **81**:4425–4439 https://doi.org/10.1016/j.molcel.2021.08.029

21 Langenbacher A. D., et al. (2011) **The PAF1 complex differentially regulates cardiomyocyte specification** *Dev Biol* **353**:19–28 https://doi.org/10.1016/j.ydbio.2011.02.011

22 Park J., Park S., Lee J. S (2023) **Role of the Paf1 complex in the maintenance of stem cell pluripotency and development** *FEBS J* **290**:951–961 https://doi.org/10.1111/febs.16582

23 Van Oss S. B., Cucinotta C. E., Arndt K. M (2017) **Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation** *Trends Biochem Sci* **42**:788–798 https://doi.org/10.1016/j.tibs.2017.08.003

24 Jaehning J. A (2010) **The Paf1 complex: platform or player in RNA polymerase II transcription?** *Biochim Biophys Acta* **1799**:379–388 https://doi.org/10.1016/j.bbagrm.2010.01.001

25 Nguyen C. T., Langenbacher A., Hsieh M., Chen J. N (2010) **The PAF1 complex component Leo1 is essential for cardiac and neural crest development in zebrafish** *Dev Biol* **341**:167–175 https://doi.org/10.1016/j.ydbio.2010.02.020

26 Lu F., Langenbacher A., Chen J. N (2017) **Tbx20 drives cardiac progenitor formation and cardiomyocyte proliferation in zebrafish** *Dev Biol* **421**:139–148 https://doi.org/10.1016/j.ydbio.2016.12.009

27 Langenbacher A. D., et al. (2023) **Rtf1 Transcriptionally Regulates Neonatal and Adult Cardiomyocyte Biology** *J Cardiovasc Dev Dis* **10** https://doi.org/10.3390/jcdd10050221

28 Saga Y., et al. (1999) **MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube** *Development* **126**:3437–3447 https://doi.org/10.1242/dev.126.15.3437

29 Warner M. H., Roinick K. L., Arndt K. M (2007) **Rtf1 is a multifunctional component of the Paf1 complex that regulates gene expression by directing cotranscriptional histone modification** *Mol Cell Biol* **27**:6103–6115 https://doi.org/10.1128/MCB.00772-07

30 Vos S. M., Farnung L., Linden A., Urlaub H., Cramer P (2020) **Structure of complete Pol II-DSIF-PAF-SPT6 transcription complex reveals RTF1 allosteric activation** *Nat Struct Mol Biol* **27**:668–677 https://doi.org/10.1038/s41594-020-0437-1

31 Cucinotta C. E., Hildreth A. E., McShane B. M., Shirra M. K., Arndt K. M (2019) **The nucleosome acidic patch directly interacts with subunits of the Paf1 and FACT complexes and controls chromatin architecture in vivo** *Nucleic Acids Res* **47**:8410–8423 https://doi.org/10.1093/nar/gkz549

32 Van Oss S. B., et al. (2016) **The Histone Modification Domain of Paf1 Complex Subunit Rtf1 Directly Stimulates H2B Ubiquitylation through an Interaction with Rad6** *Mol Cell* **64**:815–825 https://doi.org/10.1016/j.molcel.2016.10.008

33 Piro A. S., Mayekar M. K., Warner M. H., Davis C. P., Arndt K. M (2012) **Small region of Rtf1 protein can substitute for complete Paf1 complex in facilitating global histone H2B ubiquitylation in yeast** *Proc Natl Acad Sci U S A* **109**:10837–10842 https://doi.org/10.1073/pnas.1116994109

34 de Jong R. N., et al. (2008) **Structure and DNA binding of the human Rtf1 Plus3 domain** *Structure* **16**:149–159 https://doi.org/10.1016/j.str.2007.10.018

35 Core L., Adelman K (2019) **Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation** *Genes Dev* **33**:960–982 https://doi.org/10.1101/gad.325142.119

36 Adelman K., Lis J. T (2012) **Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans** *Nat Rev Genet* **13**:720–731 https://doi.org/10.1038/nrg3293

37 Peng J., Zhu Y., Milton J. T., Price D. H (1998) **Identification of multiple cyclin subunits of human P-TEFb** *Genes Dev* **12**:755–762 https://doi.org/10.1101/gad.12.5.755

38 Price D. H (2000) **P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II** *Mol Cell Biol* **20**:2629–2634 https://doi.org/10.1128/MCB.20.8.2629-2634.2000

39 Paparidis N. F., Durvale M. C., Canduri F (2017) **The emerging picture of CDK9/P-TEFb: more than 20 years of advances since PITALRE** *Mol Biosyst* **13**:246–276 https://doi.org/10.1039/c6mb00387g

40 Fujinaga K., Huang F., Peterlin B. M (2023) **P-TEFb: The master regulator of transcription elongation** *Mol Cell* **83**:393–403 https://doi.org/10.1016/j.molcel.2022.12.006

41 Jonkers I., Kwak H., Lis J. T (2014) **Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons** *Elife* **3** https://doi.org/10.7554/eLife.02407

42 Huang M., et al. (2022) **Negative Elongation Factor (NELF) Inhibits Premature Granulocytic Development in Zebrafish** *Int J Mol Sci* **23** https://doi.org/10.3390/ijms23073833

43 Yang Q., et al. (2016) **RNA polymerase II pausing modulates hematopoietic stem cell emergence in zebrafish** *Blood* **128**:1701–1710 https://doi.org/10.1182/blood-2016-02-697847

44 Gaertner B., Zeitlinger J (2014) **RNA polymerase II pausing during development** *Development* **141**:1179–1183 https://doi.org/10.1242/dev.088492

45 Gilchrist D. A., et al. (2012) **Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks** *Genes Dev* **26**:933–944 https://doi.org/10.1101/gad.187781.112

46 Adelman K., et al. (2009) **Immediate mediators of the inflammatory response are poised for gene activation through RNA polymerase II stalling** *Proc Natl Acad Sci U S A* **106**:18207–18212 https://doi.org/10.1073/pnas.0910177106

47 Rougvie A. E., Lis J. T (1988) **The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged** *Cell* **54**:795–804 https://doi.org/10.1016/s0092-8674(88)91087-2

48 Chen F. X., et al. (2017) **PAF1 regulation of promoter-proximal pause release via enhancer activation** *Science* **357**:1294–1298 https://doi.org/10.1126/science.aan3269

49 Gilchrist D. A., et al. (2010) **Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation** *Cell* **143**:540–551 https://doi.org/10.1016/j.cell.2010.10.004

50 Gilchrist D. A., et al. (2008) **NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly** *Genes Dev* **22**:1921–1933 https://doi.org/10.1101/gad.1643208

51 Jurynec M. J., et al. (2019) **The Paf1 complex and P-TEFb have reciprocal and antagonist roles in maintaining multipotent neural crest progenitors** *Development* **146** https://doi.org/10.1242/dev.180133

52 Wang Z., et al. (2022) **Coordinated regulation of RNA polymerase II pausing and elongation progression by PAF1** *Sci Adv* **8** https://doi.org/10.1126/sciadv.abm5504

53 Westerfield M. (2000) **The zebrafish book** *A guide for the laboratory use of zebrafish (Danio rerio)*

54 Farley F. W., Soriano P., Steffen L. S., Dymecki S. M (2000) **Widespread recombinase expression using FLPeR (flipper) mice** *Genesis* **28**:106–110

55 Chen J. N., Fishman M. C (1996) **Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation** *Development* **122**:3809–3816

56 Akanuma T., Koshida S., Kawamura A., Kishimoto Y., Takada S (2007) **Paf1 complex homologues are required for Notch-regulated transcription during somite segmentation** *EMBO Rep* **8**:858–863 https://doi.org/10.1038/sj.embor.7401045

57 Robu M. E., et al. (2007) **p53 activation by knockdown technologies** *PLoS Genet* **3** https://doi.org/10.1371/journal.pgen.0030078

58 Yin C., Kikuchi K., Hochgreb T., Poss K. D., Stainier D. Y (2010) **Hand2 regulates extracellular matrix remodeling essential for gut-looping morphogenesis in zebrafish** *Dev Cell* **18**:973–984 https://doi.org/10.1016/j.devcel.2010.05.009

59 Kim D., et al. (2013) **TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions** *Genome Biol* **14** https://doi.org/10.1186/gb-2013-14-4-r36

60 Liao Y., Smyth G. K., Shi W (2014) **featureCounts: an efficient general purpose program for assigning sequence reads to genomic features** *Bioinformatics* **30**:923–930 https://doi.org/10.1093/bioinformatics/btt656

61 Love M. I., Huber W., Anders S (2014) **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2** *Genome Biol* **15** https://doi.org/10.1186/s13059-014-0550-8

62 Yu G., Wang L. G., Han Y., He Q. Y (2012) **clusterProfiler: an R package for comparing biological themes among gene clusters** *OMICS* **16**:284–287 https://doi.org/10.1089/omi.2011.0118

63 Chicas A., et al. (2010) **Dissecting the unique role of the retinoblastoma tumor suppressor during cellular senescence** *Cancer Cell* **17**:376–387 https://doi.org/10.1016/j.ccr.2010.01.023

64 Langmead B., Salzberg S. L (2012) **Fast gapped-read alignment with Bowtie 2** *Nat Methods* **9**:357–359 https://doi.org/10.1038/nmeth.1923

65 Lun A. T., Smyth G (2016) **K. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows** *Nucleic Acids Res* **44** https://doi.org/10.1093/nar/gkv1191

66 Lun A. T., Smyth G. K (2014) **De novo detection of differentially bound regions for ChIP-seq data using peaks and windows: controlling error rates correctly** *Nucleic Acids Res* **42** https://doi.org/10.1093/nar/gku351

67 Kurtenbach S., Harbour J. W. (2019) **SparK: A Publication-quality NGS Visualization Tool** *bioRxiv* https://doi.org/10.1101/845529

68 Lawson N. D., et al. (2020) **An improved zebrafish transcriptome annotation for sensitive and comprehensive detection of cell type-specific genes** *Elife* **9** https://doi.org/10.7554/eLife.55792

69 Hao Y., et al. (2021) **Integrated analysis of multimodal single-cell data** *Cell* **184**:3573–3587 https://doi.org/10.1016/j.cell.2021.04.048

70 Korsunsky I., et al. (2019) **Fast, sensitive and accurate integration of single-cell data with Harmony** *Nat Methods* **16**:1289–1296 https://doi.org/10.1038/s41592-019-0619-0

71 Song M., et al. (2022) **GATA4/5/6 family transcription factors are conserved determinants of cardiac versus pharyngeal mesoderm fate** *Sci Adv* **8** https://doi.org/10.1126/sciadv.abg0834

72 Cao J., et al. (2019) **The single-cell transcriptional landscape of mammalian organogenesis** *Nature* **566**:496–502 https://doi.org/10.1038/s41586-019-0969-x

Article and author information

Adam D. Langenbacher
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America
ORCID iD: 0000-0002-3752-0208

Fei Lu
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Luna Tsang
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Zi Yi Stephanie Huang
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Benjamin Keer
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Zhiyu Tian
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Alette Eide
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Matteo Pellegrini
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Haruko Nakano
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Atsushi Nakano
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America

Jau-Nian Chen
Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California, United States of America
For correspondence: chenjn@mcdb.ucla.edu
ORCID iD: 0000-0001-8807-3607

Copyright

© 2024, Langenbacher et al.

This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Editors

Reviewing Editor
Jiandong Liu
University of North Carolina at Chapel Hill, Chapel Hill, United States of America

Senior Editor
Didier Stainier
Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany

Reviewer #1 (Public Review):

Summary:
The manuscript submitted by Langenbacher et al., entitled " Rtf1-dependent transcriptional pausing regulates cardiogenesis", describes very interesting and highly impactful observations about the function of Rtf-1 in cardiac development. Over the last few years, the Chen lab has published novel insights into the genes involved in cardiac morphogenesis. Here, they used the mouse model, the zebrafish model, cellular assays, single cell transcription, chemical inhibition, and pathway analysis to provide a comprehensive view of Rtf1 in RNAPII (Pol2) transcription pausing during cardiac development. They also conducted knockdown-rescue experiments to dissect the functions of Rtf1 domains.

Strengths:
The most interesting discovery is the connection between Rtf1 and CDK9 in regulating Pol2 pausing as an essential step in normal heart development. The design and execution of these experiments also demonstrate a thorough approach to revealing a previously underappreciated role of Pol2 transcription pausing in cardiac development. This study also highlights the potential amelioration of related cardiac deficiencies using small molecule inhibitors against cyclin dependent kinases, many of which are already clinically approved, while many other specific inhibitors are at various preclinical stages of development for the treatment of other human diseases. Thus, this work is impactful and highly significant.

https://doi.org/10.7554/eLife.94524.1.sa1

Reviewer #2 (Public Review):

Summary:

Langenbacher at el. examine the requirement of Rtf1, a component of the PAF1C, which regulates transcriptional pausing in cardiac development. The authors first confirm their previous morphant study with newly generated rtf1 mutant alleles, which recapitulate the defects in cardiac progenitor and differentiation gene expression observed previously in morphants. They then examine the conservation of Rtf1 in mouse embryos and embryonic stem cell-derived cardiomyocytes. Conditional loss of Rtf1 in mesodermal lineages and depletion in murine ESCs demonstrates a failure to turn on cardiac progenitor and differentiation marker genes, supporting conservation of Rtf1 in promoting cardiac

development. The authors subsequently employ bulk RNA-seq on flow-sorted hand2:GFP+ cells and multiomic single-cell RNA-seq on whole Rtf1-depleted embryos at the 10-12 stage. These experiments corroborate that genes associated with cardiac and muscle development are lost. Furthermore, the differentiation trajectories suggest that the expression of genes associated with cardiac maturation is not initiated. Structure-function analysis supports that the Plus3 domain is necessary for its function in promoting cardiac progenitor formation. ChIP-seq for RNA Pol II on 10-12 somite stage embryos suggests that Rtf1 is required for proper promoter pausing. This defect can partially be rescued through use of a pharmacological inhibitor for Cdk9, which inhibits elongation, can partially restore elongation in rtf1 mutants.

Strengths:

Many aspects of the data are strong, which support the basic conclusions of the authors that Rtf1 is required for transcriptional pausing and has a conserved requirement in vertebrate cardiac development. Areas of strength include the genetic data supporting the conserved requirement for Rtf1 in promoting cardiac development, the complementary bulk and single-cell RNA-sequencing approaches providing some insight into the gene expression changes of the cardiac progenitors, the structure-function analysis supporting the requirement of the Plus3 domain, and the pharmacological epistasis combined with the RNA Pol II ChIP-seq, supporting the mechanism implicating Cdk9 in the Rtf1 dependent mechanism of RNA Pol II pausing.

Weaknesses:

While most of the basic conclusions are supported by the data, there are a number of analyses that are confusing as to why they chose to perform the experiments the way they did and some places where the interpretations presently do not support the interpretations. One of the conclusions is that the phenotype affects the maturation of the cardiomyocytes and they are arresting in an immature state. However, this seems to be mostly derived from picking a few candidates from the single cell data in Fig. 6. If that were the case, wouldn't the expectation be to observe relatively normal expression of earlier marker genes required for specification, such as Nkx2.5 and Gata5/6? The in situ expression analysis from fish and mice (Fig. 2 and Fig. 3) and bulk RNA-seq (Fig. 5) seems to suggest that there are pretty early specification and differentiation defects. While some genes associated with cardiac development are not changed, many of these are not specific to cardiomyocyte progenitors and expressed broadly throughout the ALPM. Similarly, it is not clear why a consistent set of cardiac progenitor genes (for instance mef2ca, nkx2.5, and tbx20) was analyzed for all the experiments, in particular with the single cell analysis.

The point of the multiomic analysis is confusing. RNA- and ATAC-seq were apparently done at the same time. Yet, the focus of the analysis that is presented is on a small part of the RNA-seq data. This data set could have been more thoroughly analyzed, particularly in light of how chromatin changes may be associated with the transcriptional pausing. This seems to be a lost opportunity. Additionally, how the single cell data is covered in Supplemental Fig. 2 and 3 is confusing. There is no indication of what the different clusters are in the Figure or the legend.

While the effect of Rtf1 loss on cardiomyocyte markers is certainly dramatic, it is not clear how well the mutant fish have been analyzed and how specific the effect is to this population. It is interpreted that the effects on cardiomyocytes are not due to "transfating" of other cell fates, yet supplemental Fig. 4 shows numerous effects on potentially adjacent cell populations. Minimally, additional data needs to be provided showing the live fish at these stages and marker analysis to support these statements. In some images, it is not clear the embryos are the same stage (one can see pigmentation in the eyes of controls that is not in the mutants/morphants), causing some concern about developmental delay in the mutants.

With respect to the transcriptional pausing defects in the Rtf1 deficient embryos, it is not clear from the data how this effect relates to the expression of the cardiac markers. This could have been directly analyzed with some additional sequencing, such as PRO-seq, which would provide a direct analysis of transcriptional elongation.

Some additional minor issues include the rationale that sequence conservation suggests an important requirement of a gene (line 137), which there are many examples this isn't the case, referencing figures panels out of order in Figs. 4, 7, and 8) as described in the text, and using the morphants for some experiments, such as the rescue, that could have been done in a blinded manner with the mutants.

https://doi.org/10.7554/eLife.94524.1.sa0