Title:

Rtf1-dependent transcriptional pausing regulates cardiogenesis

Authors:

Adam D. Langenbacher[1], Fei Lu[1], Luna Tsang[1], Zi Yi Stephanie Huang[1], Benjamin Keer[1], Zhiyu Tian[1], Alette Eide[1], Matteo Pellegrini[1], Haruko Nakano[1], Atsushi Nakano[1], Jau-Nian Chen[1]*

Author Affiliations:

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

Correspondence and Materials Requests:

Jau-Nian Chen

Department of Molecular, Cell and Developmental Biology, UCLA, Los Angeles, CA 90095

chenjn@mcdb.ucla.edu

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.

Keywords:

cardiogenesis, lateral plate mesoderm, transcription regulation, promoter-proximal pausing, Rtf1, PAF1 complex, PAF1C

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

53 elongation rate of RNA Polymerase II (RNA Pol II). For example, the BAF (Brg1/Brm-associated
54 factor) complex, an ATP-dependent chromatin remodeling complex that actively positions
55 nucleosomes, is critical for normal vertebrate cardiogenesis[9,10]. The BAF complex directly
56 interacts with cardiac transcription factors including Tbx5, Nkx2-5, and Gata4[11] and controls
57 temporal changes in chromatin accessibility during cardiac differentiation[9]. Intriguingly,
58 transplanted cells overexpressing the BAF component Baf60c/Smarca3b and Gata5
59 preferentially contribute to the embryonic zebrafish heart[12]. Similarly, forced expression of
60 Baf60c along with Tbx5/Nkx2-5/Gata4 can induce cardiac differentiation from non-cardiac
61 mesoderm[13], suggesting that nucleosomal positioning and chromatin accessibility are key
62 mechanisms regulating the transcription of cardiac genes.
63
64 Accumulating evidence suggests that cardiac gene expression is also highly sensitive to the
65 activity of transcriptional pausing and elongation factors. Spt6 (Suppressor of Ty 6) is a histone
66 chaperone that promotes RNA Polymerase II elongation and co-transcriptionally modulates
67 chromatin structure[14,15]. In zebrafish, loss of function of Spt6 disrupts cardiomyocyte
68 differentiation and interacts genetically with the transcriptional pausing/elongation factor Spt5
69 (Suppressor of Ty 5) to support Nkx2.5+ cardiac progenitor specification[16]. While both Spt6 and
70 Spt5 promote transcription elongation, they likely do so via distinct mechanisms. Spt5 directly
71 interacts with Spt4 to form the DSIF (DRB Sensitivity Inducing Factor) complex, which has both
72 inhibitory and stimulatory roles in transcription[17]. Unphosphorylated Spt5 fosters promoter-
73 proximal pausing of RNA Polymerase II, but phosphorylation of Spt5's C-terminal repeat region
74 by Cdk9 (cyclin dependent kinase 9) produces a binding site for the Polymerase Associated
75 Factor 1 complex (PAF1C) component Rtf1, promotes transcriptional pause release, and
76 converts DSIF into an elongation-promoting factor[18-20]. How the transcription regulatory
77 activities of these elongation and pausing factors facilitate cardiac development is not yet
78 understood.

79

80 We have previously shown that Rtf1 and other PAF1C members differentially regulate
81 cardiogenesis[21]. The PAF1C consists of five core proteins, Paf1, Ctr9, Cdc73, Rtf1, and Leo1,
82 and plays wide-ranging transcription regulatory roles including promoting co-transcriptional
83 epigenetic modification of histones, supporting both pausing and elongation of RNA Polymerase
84 II, influencing mRNA 3' end formation, and serving as a bridge between specific transcription
85 factors and RNA Polymerase II[22-24]. In zebrafish, Leo1 is dispensable for the specification of
86 cardiomyocytes, but is required for differentiation of the atrioventricular boundary and
87 maturation of the cardiac chambers[25]. Embryos lacking Ctr9, Cdc73, or Paf1, on the other
88 hand, exhibit a deficiency of cardiomyocytes and only form small, dysmorphic hearts.
89 Intriguingly, loss of the Rtf1 subunit of the PAF1C results in the most dramatic defects in
90 cardiogenesis. Rtf1 deficient embryos have severe reduction in *nkx2.5*+ cardiac progenitors
91 and fail to generate any cardiomyocytes[21]. Given the numerous roles played by the PAF1C in
92 transcription regulation, it remains an open question how PAF1C members mechanistically
93 regulate cardiogenesis.

94

95 In this study, we investigated the molecular mechanisms of Rtf1-dependent cardiac progenitor
96 formation. Using CRISPR mutagenesis we showed that zygotic Rtf1 mutants, like *rtf1*
97 morphants, lack expression of most cardiac progenitor marker genes and produce no
98 cardiomyocytes. In addition, knockout of Rtf1 in the cardiogenic mesoderm of mouse embryos
99 and knockdown of Rtf1 in mouse ESCs reduced the formation of cardiac tissue, indicating that
100 the role of Rtf1 in specification of the cardiac lineage is conserved among vertebrates. Using
101 FACS and single cell sequencing-based approaches, we show that differentiation of the LPM
102 into cardiac precursors requires Rtf1 activity. We investigated which of Rtf1's functional
103 domains participated in cardiac development and found that its Plus3 domain is critical for
104 cardiac progenitor formation, while its HMD (Histone modification domain) is dispensable.

105 Consistent with the Plus3 domain's function as an interaction point between Rtf1 and the DSIF
106 component Spt5, Rtf1 deficient embryos display a genome-wide decrease in promoter-proximal
107 pausing of RNA Polymerase II. Excitingly, chemical and morpholino inhibition of pause release
108 improved pausing at cardiac genes and restored cardiac progenitor formation to Rtf1-deficient
109 embryos. Together, our findings suggest that during development, Rtf1's primary function in
110 cardiac gene regulation is to support promoter proximal pausing, and that normal levels of
111 pausing facilitate expression of many genes, including key members of the cardiac transcription
112 program.
113
114 **Results**:
115
116 *Rtf1 is essential for heart development in zebrafish*
117
118 Morpholino knockdown of *rtf1* abolishes cardiac progenitor formation in zebrafish, indicating an
119 essential role for Rtf1 in cardiogenesis[21]. To genetically corroborate this finding, we used
120 CRISPR/Cas9 to generate stable *rtf1* mutant lines. We identified two independent lines
121 (*rtf1*^{LA2678} and *rtf1*^{LA2679}) with small deletions in exon 3 of the *rtf1* gene that are predicted to result
122 in frameshifts and early stop codons and which nearly eliminate the production of Rtf1 protein
123 (Fig 1A-C). Like *rtf1* morphants, both *rtf1*^{LA2678} and *rtf1*^{LA2679} mutants display no cardiac tissue at
124 24 hours post fertilization (hpf) and lack the *nkx2.5*+ cardiac progenitor population (Figure
125 2A,B,E,F). Rtf1 deficient embryos also fail to express normal levels of other cardiac
126 transcription factors including *mef2ca* and *tbx5a* (Fig 2C,D,G,H), and exhibit reduced expression
127 of *tbx20*, a key regulator of cardiac cell number that demarcates the anterior lateral plate
128 mesoderm (ALPM) at the 8-somite stage (Fig 2I,J)[26]. Knockdown and knockout of Rtf1 activity
129 produced comparable changes in cardiac gene expression (Fig 2A-J), demonstrating that
130 morpholino knockdown faithfully recapitulates the *rtf1* mutant phenotype. Together, these data

131 show that Rtf1 activity is required for the expression of multiple cardiac transcription factors and
132 suggest that the absence of cardiomyocytes in Rtf1 deficient embryos is likely a result of this
133 dramatic gene expression defect.
134
135 *Conserved role of Rtf1 in mammalian cardiac development*
136
137 The sequence of the Rtf1 protein is highly conserved among vertebrates (83% identity at the
138 amino acid level between fish and humans) indicating that it is likely playing a fundamental role
139 in biology. To examine whether Rtf1 activity is also needed for cardiogenesis in mammals, we
140 expressed Cre-recombinase under the control of the Mesp1 promoter in Rtf1-flox mouse
141 embryos (Fig 3A)[27]. The Mesp1 promoter drives expression of Cre in the early cardiogenic
142 mesoderm (E6.25)[28], allowing for knockout of Rtf1 activity prior to the onset of heart
143 development. Rtf1 Mesp1-Cre knockout embryos die prior to birth, suggesting that Rtf1 function
144 in the cardiogenic mesoderm is essential for survival. We next examined the developing hearts
145 of Rtf1 Mesp1-Cre knockouts and their wild type siblings by in situ hybridization using *Nkx2.5*
146 and *Tbx20* as cardiac markers and found that while Cre+;Rtf1flox/+ and Cre- embryos display a
147 heart tube undergoing looping at E8.5, Rtf1 Mesp1-Cre knockouts have little or no expression of
148 these markers (Fig 3B). Cardiac muscle, marked by expression of *MLC2v* at E9.5, was also
149 dramatically reduced in Mesp1-Cre Rtf1 knockouts (Fig 3B). These data suggest a conservation
150 of the critical role of Rtf1 in differentiation of the cardiac lineage from the mesoderm.
151
152 To further investigate Rtf1's role in cardiac differentiation, we knocked down Rtf1 activity in
153 mouse embryonic stem cells (mESCs) using short hairpin RNA (Fig 4A,B). Unmanipulated
154 embryoid bodies (EBs) derived from mESCs are capable of differentiating into lineages
155 representing all three germ layers, as evidenced by temporally dynamic expression of
156 *Brachyury* (mesoderm), *Afp* (endoderm), and *Ncam1* (ectoderm) (Fig 4E-G). Similarly, we found

157 that Rtf1 shRNA mESCs induce expression of *Brachyury*, *Afp*, and *Ncam1* in a temporally
158 appropriate manner, indicating that Rtf1 activity is dispensable for germ layer formation (Fig 4E-
159 G). However, EBs derived from Rtf1-deficient mESCs produced significantly fewer beating
160 patches (~67% reduction) and also displayed a dramatic reduction in cardiac genes like *Myh6*,
161 *Nkx2.5*, and *ANF*/*Nppa* (Figure 4C,D). Together, these data demonstrate that Rtf1 is essential
162 for differentiation of the cardiac lineage from mesoderm in mammals, and reveal a conservation
163 of Rtf1's function in cardiogenesis between fish and mammals.
164
165 *Rtf1 is required for expression of genes associated with the progression of cardiac precursor*
166 *formation*
167
168 We previously found that loss of Rtf1 activity does not disrupt expression of the LPM marker
169 *hand2*[21]. To better understand how Rtf1 directs cardiac differentiation from multipotent
170 mesoderm cells, we took advantage of a hand2:GFP zebrafish transgenic reporter line to isolate
171 LPM cells from wild type and Rtf1 deficient embryos for transcriptomic analysis (Fig 5A-C,
172 Supplementary Figure 1). Compared to wild type control samples, the *rtf1* morphant LPM
173 displayed 1420 significantly downregulated and 1185 significantly upregulated genes (adjusted
174 p-value < 0.05) at the 10-12 somite stage when cardiac progenitor formation is underway in the
175 LPM (Supplementary Data File 1). We examined the most significantly differentially expressed
176 genes and observed that many of the significantly downregulated genes, including *nkx2.5*,
177 *rbfox1l*, *atp2a2a*, *tnnt2a*, *isl2b*, *mef2ca*, *mef2cb* and *wnt2bb*, are cardiac-lineage markers or
178 have critical known roles in cardiac development (Fig 5D). Indeed, the most significantly
179 downregulated genes were highly enriched for gene ontologies related to cardiac development,
180 embryonic heart morphogenesis, and muscle cell development (Fig 5E). These data show that
181 Rtf1 activity is required for expression of a broad cardiac gene program in the LPM that is likely
182 under the control of multiple cardiac transcription factors.

183

184 To further analyze the deployment of the cardiac gene program during cardiac progenitor
185 formation from the ALPM, we performed Single Cell Multiome sequencing of control and *rtf1*
186 morphant embryos at the 11-12 somite stage. Clustering of control and *rtf1* morphant cells
187 based on gene expression and chromatin accessibility resolved 39 cell types corresponding to
188 expected major cell lineages (Supplementary Data File 2, Supplementary Figures 2 & 3).
189 Interestingly, while *rtf1* morphant cells of each cell type were present, the frequencies of cell
190 types differed significantly between controls and *rtf1* morphants (Supplementary Figure 4).
191 Notably, we also observed that while *rtf1* morphant cells expressing the precardiac mesoderm
192 marker *mef2cb* were present, they failed to express other cardiac marker genes including *ryr2b*
193 and *tnni1b* (Supplementary Figure 3).
194
195 To examine developmental trajectories within the ALPM, we reclustered ALPM cells
196 (*sema3e*+/*gata5*+ LPM) and ALPM derivatives (putative precardiac mesoderm, endothelial
197 precursors, rostral blood precursors) (Figure 6A,B). We next calculated pseudotime trajectories
198 based on the assumption that undifferentiated ALPM, intermediates, and more highly
199 differentiated derivatives are present at this developmental timepoint (Figure 6C). In support of
200 this notion, we observed a gradual increase in the expression levels of cardiac markers
201 including *ttn.2*, *mef2cb*, *tnnt2a*, *ryr2b*, and *myh7bb* along the ALPM-cardiac developmental
202 trajectory in control embryos (Fig 6D,E). Interestingly, Rtf1 deficient embryos properly
203 downregulate expression of the ALPM marker *sema3e*, maintain the expression of *gata5* in
204 cardiac precursors, and initiate expression of the early cardiac lineage genes *bmp6*, *ttn.2*, and
205 *mef2cb* (Fig 6D-G). However, loss of Rtf1 activity prevented the activation of cardiac lineage-
206 restricted markers that are normally expressed later in the ALPM-cardiac trajectory including
207 *tnnt2a*, *ryr2b*, and *myh7bb* (Fig 6D-G). Taken together, these data suggest that Rtf1 activity is
208 required for gene expression changes associated with the progression of cardiac precursors

209 into a more mature state, and that Rtf1-deficient cardiac precursors are arresting in an immature
210 state rather than transfating to a different cell lineage.
211
212 *Rtf1's Plus3 functional domain is necessary for cardiac progenitor formation*
213
214 The Rtf1 protein is a multifunctional platform for transcription regulation: it consists of several
215 defined functional domains that mediate its interaction with other transcription regulatory
216 proteins and serve as additional points of contact for these proteins with the transcription
217 apparatus (Figures 1A, 7A)[29,30]. The N-terminal half of Rtf1 notably contains its Histone
218 Modification Domain (HMD), which mediates its interaction with E2 ubiquitin-conjugating
219 enzymes and supports the ubiquitylation of histone H2B K120[31-33]. Rtf1's C-terminus contains
220 the Plus3 domain, which interacts with the C-terminal repeat region of the pausing and
221 elongation factor Spt5[18,19,34], a Pol II-interaction domain that allosterically stimulates
222 transcription elongation[30], and a PAF1C interaction domain.
223
224 We found that injection of *rtf1* mRNA with 8 silent mutations that render it morpholino resistant
225 (Rtf1 wt) into *rtf1* morphant embryos was able to robustly rescue the knockdown phenotype and
226 support normal formation of the embryonic heart tube (Figure 7C). This knockdown-rescue
227 assay provides an in vivo platform to test whether Rtf1's HMD and Plus3 domains are required
228 for the formation of cardiac progenitors during development. We generated expression
229 constructs for mutant versions of Rtf1 lacking these functional domains and tested whether they
230 could support cardiogenesis in an *rtf1* morphant background. Mutant FLAG-tagged Rtf1
231 proteins lacking the HMD or Plus3 domain were expressed at levels comparable to Rtf1 wt and
232 properly localized to the nucleus demonstrating that the activity of these proteins was not
233 compromised by instability or mislocalization (Fig 7B). Interestingly, we found that Rtf1's HMD
234 was dispensable for its role in cardiac progenitor formation (100% of Rtf1 wt injected *rtf1*

235 morphants with *nkx2.5* signal, n=25 vs. 100% of Rtf1 ΔHMD *rtf1* morphants with *nkx2.5* signal,
236 n=25) (Figure 7C). On the other hand, Rtf1's Plus3 domain was essential for supporting the
237 formation of cardiac progenitors (0% of Rtf1 ΔPlus3 *rtf1* morphants with *nkx2.5* signal, n=25)
238 (Figure 7C). These data suggest that Rtf1's domains function in a modular fashion for
239 regulating cardiac gene expression and point to potential specific effects on transcriptional
240 pausing and elongation in cardiogenesis.
241
242 *Rtf1 deficiency disrupts promoter-proximal transcriptional pausing*
243
244 Given the critical role of Rtf1's Plus3 domain in cardiac progenitor formation, we hypothesized
245 that Rtf1 deficiency may impact the transcriptional activity of RNA Pol II at cardiac genes. We
246 then examined the occupancy of RNA Pol II by performing ChIP-seq on embryos at the 10-12
247 somite stage, when cardiac progenitors are arising from the lateral plate mesoderm. We
248 compared the occupancy of Pol II between uninjected control and *rtf1* morphant embryos and
249 found that Pol II was substantially diminished at the transcription start site (TSS) of many genes
250 compared to the remainder of the gene body (Figure 8D-I and data not shown). Pausing of
251 RNA Pol II at the TSS is a normal aspect of transcription in vertebrates and has been
252 hypothesized to regulate both the timing and expression level of genes[35,36]. To quantify
253 transcriptional pausing at each gene, we calculated the Pause Release Ratio (PRR), a ratio of
254 Pol II occupancy in the promoter proximal region of a gene (-300 to +300) to the occupancy in
255 the downstream region (+301 to +1301). A higher PRR indicates increased pausing of Pol II
256 near the transcription start site (TSS), while a lower PRR signifies that a gene is in a less
257 paused state. We found that on a global scale, pausing is diminished upon loss of Rtf1 activity
258 during embryonic development (Figure 8A,B). Out of 6,078 genes with substantial Pol II ChIP
259 signal that were analyzed, 5,799 displayed a significantly decreased PRR (Benjamini–Hochberg
260 false discovery rate of 10%) in Rtf1-deficient embryos compared to 37 genes that displayed a

261 significantly increased PRR (Figure 8B). The median PRR in control embryos was 3.34, while
262 95% (5,793/6,078) of genes had a PRR of less than 3.34 in Rtf1-deficient embryos (Figure 8A),
263 reflecting a widespread decrease in promoter proximally paused Pol II. These changes in PRR
264 suggest that Rtf1 is critical for normal levels of pausing at transcriptionally active genes during
265 the developmental window when cardiac progenitor cells arise from the lateral plate mesoderm.
266
267 *Attenuation of pause release permits cardiogenesis in an Rtf1-deficient background*
268
269 The transition between the pausing and elongation phases of transcription is regulated by
270 PTEF-b, a complex consisting of a cyclin (T1, T2a, or T2b) and the kinase Cdk9[37-
271 39]. Phosphorylation of the negative elongation factor (NELF) complex, the pausing/elongation
272 factor Spt5, and the C-terminal tail of RNA Pol II by Cdk9 promotes changes in the factors
273 associated with RNA Pol II and triggers a switch to processive elongation[40]. Pharmacological
274 inhibition of Cdk9 with the small molecule flavopiridol can prevent or attenuate pause release,
275 leading to genome-wide increases in pausing at most genes[41]. We explored the effects of Cdk9
276 inhibition on Rtf1-dependent transcription by performing ChIP-seq for RNA Pol II in flavopiridol-
277 treated *rtf1* morphant embryos at the 10-12 somite stage. Flavopiridol treatment caused a small
278 but highly significant increase in promoter-proximal pausing in an Rtf1 deficient background (Fig
279 8A), with 87.1% of genes with significantly altered PRR values displaying an increase in PRR in
280 flavopiridol-treated *rtf1* morphants compared to untreated *rtf1* morphants (Figure 8C).
281 Furthermore, we found that critical LPM and cardiac genes including *hand2*, *gata5*, *aplnrb*,
282 *bmp4*, *nkx2.7*, and *rbfox1l* exhibited reduced promoter-proximal pausing in Rtf1 deficient
283 embryos that was partially rescued by treatment with flavopiridol (Figure 8D-I).
284
285 If diminished promoter-proximal pausing at genes crucial for cardiogenesis is responsible for the
286 absence of cardiac progenitor formation in Rtf1-deficient embryos, then we would predict that

287 reducing pause release may restore cardiogenesis in an Rtf1-deficient background. We
288 examined expression of the cardiomyocyte marker *myl7* in wild type, *rtf1* mutant, and
289 flavopiridol-treated *rtf1* mutant embryos and found that partial attenuation of Cdk9 activity (5
290 ug/ml dosage) potently restored cardiomyocyte formation in Rtf1-deficient embryos (Figure
291 8J,K). Similarly, we found that knockdown of Cdk9 by injection of antisense morpholino
292 oligonucleotides[42,43] also restored the formation of *myl7*+ cardiomyocytes in an Rtf1-deficient
293 background (Figure 8L,M), confirming that decreased activity of the Cdk9 kinase is responsible
294 for rescuing the *rtf1* cardiac phenotype. These dramatic results suggest that the primary role of
295 Rtf1 in cardiac progenitor formation is regulating promoter-proximal pausing of critical members
296 of the cardiac gene program. Altogether, our data support a model in which loss of Rtf1 causes
297 excessive pause release, a dysregulation of the cardiac transcription program, and an arrest in
298 the developmental trajectory of cardiac progenitors prior to the formation of mature
299 cardiomyocytes.
300
301 **Discussion**:
302
303 In this study we investigated the molecular mechanism of the PAF1C member Rtf1's role in
304 cardiac progenitor formation. We found that Rtf1 plays a critical and conserved role in
305 cardiogenesis and cardiac gene expression in vertebrate and mammalian models. Zebrafish
306 *rtf1* mutants and mouse Rtf1 cardiac mesoderm *Rtf1* conditional knockouts do not properly
307 initiate expression of cardiac transcription factors in the LPM and fail to generate mature
308 cardiomyocytes. Using single cell sequencing, we found that Rtf1 deficient embryos do
309 generate immature cardiac progenitors, but that these progenitors fail to express genes
310 associated with progression to a more mature state. Excitingly, we also showed that the Rtf1
311 protein can function in a modular fashion and that its Plus3 domain, a site of interaction with the
312 pausing/elongation factor Spt5, is essential for supporting cardiac development. In line with this

313 finding, we showed that loss of Rtf1 diminishes promoter-proximal pausing of RNA Polymerase
314 II, and that partially restoring normal pausing by chemical inhibition of pause release can
315 support cardiogenesis in an Rtf1-deficient background.
316
317 Despite having a nearly genome-wide decrease in promoter-proximal pausing of RNA
318 Polymerase II, Rtf1-deficient embryos display many upregulated and downregulated genes.
319 How exactly pausing regulates gene expression is not fully understood and likely involves both
320 general and gene/locus-specific mechanisms[35,44]. By limiting the release of RNA Polymerase II
321 into the elongation phase of transcription, pausing can restrain gene expression. This inhibitory
322 role of pausing on transcription allows for another layer of transcription regulation and has been
323 shown to modulate the expression of heat shock and immune stimulus-induced genes[45-47].
324 Pausing-dependent restraint of gene expression is also important for the timing of gene
325 expression in a developmental context. Supporting this notion, Spt5 and NELF-deficient
326 zebrafish embryos exhibit aberrant upregulation of TGFβ signaling and a resulting inhibition of
327 hematopoiesis[43]. Intriguingly, the PAF1C has also been shown in mammalian cells to restrain
328 transcription from a subset of promoters by inhibiting nearby enhancer activation[48].
329
330 On the other hand, pausing also positively regulates gene expression in some contexts. At
331 genes where nucleosome formation is favored in promoter regions, pausing of RNA Polymerase
332 II can promote a chromatin architecture that is permissive to transcription by competing with
333 nucleosomes for promoter occupancy[49,50]. This mechanism has been suggested to support
334 expression of the IFN-gamma pathway during zebrafish development[43]. Pausing may also
335 positively regulate gene expression during neural crest development. Loss of Paf1 activity
336 disrupts neural crest differentiation, but this defect can be partially rescued by inactivation of
337 Cdk9. Interestingly, zygotic loss of Cdk9 alone is sufficient to expand the neural crest

338 population in zebrafish embryos, suggesting that promoter-proximal pausing promotes the
339 neural crest gene program[51].
340
341 Recent work using a rapid protein degradation system in human cell culture has shed light on
342 the mechanisms by which the PAF1C regulates transcriptional pausing and gene expression
343 outputs[52]. In DLD-1 cells, acute depletion of PAF1 destabilizes the PAF1C and stimulates
344 pause release. Loss of PAF1 also decreases chromatin recruitment of INTS11, a component of
345 the Integrator-PP2A complex (INTAC) which interacts with PAF1C and counteracts the activity
346 of P-TEFb. This disruption in the balance of INTAC and P-TEFb activities at transcribed genes
347 results in hyperphosphorylation of SPT5 and likely explains the increase in pause release
348 caused by PAF1 depletion. Similar to our finding that Rtf1-deficiency positively and negatively
349 regulates gene expression, depletion of PAF1 resulted in both upregulation and downregulation
350 of many genes. Intriguingly, the gene expression changes caused by PAF1 depletion correlated
351 with the release frequency and processivity of RNA Polymerase II at a given gene, with genes
352 exhibiting low levels of release frequency tending to display decreases in their expression level.
353 These observations suggest that the combined regulatory functions of pausing and elongation
354 determine the transcriptional output of PAF1C-regulated genes.
355
356 In conclusion, our study reveals that the cardiac gene program is highly sensitive to loss of
357 promoter-proximal pausing. Future studies focused on the chromatin architecture of cardiac
358 gene promoters and other regulatory regions, and the dynamics of pause release and
359 elongation at cardiac genes are needed to provide a complete picture of the manner by which
360 promoter-proximal pausing fosters their expression.
361
362 **Materials and Methods**:
363

364 *Zebrafish*

365

366 Zebrafish colonies were cared for and bred under standard conditions[53]. Developmental stages
367 of zebrafish embryos were determined based on somite number or hours of development post
368 fertilization at 28.5 °C [53]. Fish husbandry and experiments were performed according to the
369 Institutional Approval for Appropriate Care and Use of Laboratory Animals by the UCLA
370 Institutional Animal Care and Use Committee (Protocols ARC-2000-051-TR-001 and BUA-2018-
371 195-002-CR).

372

373 *CRISPR mutagenesis*

374

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

384

385 *Mice*

386

387 All mice were maintained according to the Guide for the Care and Use of Laboratory Animals
388 published by the US National Institute of Health. Housing and experiments were performed
389 according to the Institutional Approval for Appropriate Care and Use of Laboratory Animals by

390 the UCLA Institutional Animal Care and Use Committee (Protocols ARC-2020-045-TR-001 and
391 ARC-2020-043-TR-001). ES cells were obtained from KOMP and used to rederive the
392 Rtf1$^{tm1a(KOMP)Wtsi}$ mouse line. This line was then crossed to a ROSA26::FLPe strain[54] (The
393 Jackson Laboratory, JAX stock #009086) to create an Rtf1 conditional knockout allele where
394 exon 3 of Rtf1 is flanked by two loxP sites (Rtf1^{flox}). These mice were crossed into the
395 Mesp1$^{tm2(cre)Ysa}$ background[28] to produce a strain capable of generating Mesp-Cre^{+};Rtf1$^{flox/flox}$
396 offspring.
397
398 *Whole mount in situ hybridization*
399
400 Whole-mount in situ hybridization was performed as described previously[55]. Embryos for in situ
401 hybridization were fixed in 4% paraformaldehyde in 1xPBS. After a brief proteinase K digestion,
402 embryos were incubated with digoxigenin-labeled antisense RNA probes overnight at 65-70°C.
403 Probes were detected using an anti-digoxigenin antibody conjugated with alkaline phosphatase
404 (Roche 11093274910). The zebrafish antisense RNA probes used in this study include *myl7*,
405 *nkx2.5*, *mef2ca*, *tbx20*, and *tbx5a*. The mouse probes used include *MLC2v*, *Tbx20*, and *Nkx2.5*.
406
407 Genotypes of zebrafish embryos were confirmed by PCR with primers:
408 rtf1-e3-F: 5'-TACAGTGACATCCACCATGCTG-3'
409 rtf1-e3-R: 5'-GTAAATAGGGGAGAAACTAATTGAAG-3'.
410
411 Genotypes of mouse embryos were confirmed by PCR with primers:
412 Rtf1-loxP-F: 5'-GACTGAGAAGGGAAATCTTTGAAAC-3'
413 Rtf1-loxP-R: 5'-CTAGTCCTTGAACAAGTTCAGCT-3'
414 Cre-F: 5'-GTAGCTGATGATCCGAATAACTAC-3'
415 Cre-R: 5'-ATCCAGGTTACGGATATAGT-3'.

416

417 *Constructs*

418

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

427

428 *Microinjections*

429

430 0.5 ng of *rtf1* morpholino (rtf1MO; 5'-CTTTCCGTTTCTTTACATTCACCAT-3') was injected at
431 the 1-cell stage[21,56] along with 1 ng of *p53* morpholino (5'-GCGCCATTGCTTTGCAAGAATTG-
432 3') to prevent p53-dependent apoptosis[57]. For *cdk9* knockdown rescue experiments, 5 ng of
433 cdk9 morpholino (cdk9MO; 5'-ACACACAAACATCAAATACTCACCC-3')[42,43] was injected into an
434 incross of *rtf1*^{LA2679} heterozygotes at the 1-cell stage. 150-200 pg of *rtf1* or *rtf1* deletion mutant
435 mRNAs were co-injected with *rtf1* and *p53* morpholinos at the 1-cell stage for mRNA rescue
436 experiments.

437

438 *Mouse ESC culture and differentiation*

439

440 E14TG2a mESCs (ATCC) were cultured on 0.1% gelatin covered plates without feeder cells.
441 The feeder-free mESCs were maintained in Dulbecco's modified Eagle's medium (Invitrogen)

442 supplemented with 15% fetal bovine serum (FBS; Invitrogen, USA), 1% non-essential amino
443 acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 1% (v/v) penicillin-streptomycin-
444 Glutamine (ThermoFisher Scientific), and 10 μg/μl recombinant mouse leukemia inhibitory factor
445 (LIF; Sigma-Aldrich) in a 5% CO2 incubator at 37°C. The ESC culture medium was changed
446 every two days. mESCs were subcultured every 3 days using 0.25% trypsin-EDTA (Invitrogen).
447 mESCs were differentiated in hanging drops (600 cells/20 μL) without LIF under non-adherent
448 conditions. Day 5 embryoid bodies (EBs) were plated onto 0.1% gelatin covered plates. Media
449 was changed every two days during adherent culture.
450
451 *Lentivirus transduction*
452
453 To knockdown Rtf1 activity in mESCs, Rtf1 or NT (non-target) shRNA lentivirus harboring
454 puromycin resistance were purchased from Sigma-Aldrich. mESCs were transduced with the
455 indicated virus at a multiplicity of infection of 1 and selected in 1.2 μg/ml puromycin. Rtf1 shRNA
456 knockdown efficacy was tested by Western blotting analysis that detected Rtf1 protein level.
457 Antibodies used were anti-Rtf1 (Bethyl Labs A301-329A) and anti-β-actin (Sigma A1978).
458 Relative protein levels were determined by densitometry measurements using Adobe
459 Photoshop.
460
461 *Quantitative RT-PCR*
462
463 Total RNA from NT and shRNA transduced EBs was extracted using TRIzol RNA isolation
464 reagents (ThermoFisher Scientific). cDNA was synthesized using iScript cDNA Synthesis Kit
465 from Bio-Rad. Real-time quantitative PCR was carried out using the Roche LightCycler 480
466 Real-Time PCR System. Primers for qPCR are listed in Supplementary Table 1.
467

468 *Antibody staining*

469

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

476

477 *Western Blotting*

478

479 Embryos were lysed in Rubinfeld's Lysis Buffer at 1 day post fertilization as previously
480 described[21]. Two embryo equivalents were loaded per lane of a 8% polyacrylamide gel.
481 Following electrophoresis, proteins were transferred to a nitrocellulose membrane and detected
482 with antibodies against Rtf1 (Bethyl Labs A301-329A, 1:4000) and beta-actin (Sigma A1978,
483 1:5000). ImageJ was used to perform densitometry comparing the levels of Rtf1 protein
484 between samples using beta-actin as the control.

485

486 *Chemical treatment*

487

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

492

493 *FACS*

494

495 10-12 somite stage TgBAC(hand2:EGFP)^{pd24} uninjected and *rtf1* morpholino-injected zebrafish
496 embryos[58] were nutated in 10 mg/ml protease (Protease from *Bacillus licheniformis*, Sigma
497 P5380) in DMEM/F12 media for 30 minutes at 4°C and pipetted to dissociate cells. Debris was
498 removed by filtering through a 40 um cell strainer and cells were then pelleted by centrifugation
499 at 300 rcf for 5 minutes at 4°C. Cells were resuspended in cold FACS buffer (1xPBS, 2% FBS,
500 1mM EDTA) and refiltered immediately prior to sorting on a BD FACSAria II using a 100 um
501 nozzle. Dead cells were excluded by staining with 0.1 µg/ml DAPI. Cells from batches of 100
502 embryos were sorted directly into a mixture of 350 ul Buffer RA1 and 3.5 ul 2-mercaptoethanol
503 from the NucleoSpin RNA kit (Machery-Nagel) for RNA isolation.

504

505 *mRNA-seq*

506

507 RNA was purified from FACS-sorted cells with the NucleoSpin RNA kit (Machery Nagel).
508 Approximately 5-10 ng of input RNA was used to produce libraries with the NEBNext Single
509 Cell/Low Input RNA Library Prep Kit for Illumina (NEB) which were sequenced on an Illumina
510 HiSeq 3000 to produce 50 bp single-end reads. Reads from triplicate samples were mapped to
511 the danRer7 genome using Tophat v2.1.1[59] using default parameters. Mapped reads were
512 counted with FeatureCounts v2.0.0[60] and differential expression analysis was carried out with
513 DESeq2 v1.32.0[61]. Gene ontology enrichment was examined using the R package
514 clusterProfiler[62] in R v4.1.0. For gene ontology analysis, the most significantly downregulated
515 genes in *rtf1* morphant LPM were selected as those genes that were downregulated among the
516 300 most significantly differentially expressed genes based on adjusted *p*-value.

517

518 *ChIP-seq*

519

520 Embryos for ChIP were dissociated with 4 mg/ml Collagenase Type IV and 0.25% Trypsin at
521 room temperature with intermittent pipetting. Dissociation was stopped with 10% FBS in DMEM
522 and cells were washed once with PBS. Fixation and preparation of chromatin was carried out
523 using the Covaris truChIP Chromatin Shearing Kit and 7 minutes of sonication in a Covaris
524 E220 ultrasonicator using the recommended settings. Immunoprecipitations were carried out as
525 previously described[63] on approximately 5 ug of chromatin (minimum 2.5 ug, maximum 8.8 ug)
526 using 5 ul of total RNA Pol II antibody (Cell Signaling D8L4Y). Libraries were prepared using
527 the NEBNext Ultra II DNA Library Kit for Illumina (NEB) and sequenced with an Illumina
528 Novaseq 6000 system to produce 50 bp paired-end reads. Reads from triplicate samples were
529 mapped to the danRer11 genome using bowtie2 v2.4.2[64] (options: --local -X 1000) and read
530 densities of UCSC RefSeq gene regions were quantified using the package csaw v1.24.3[65,66] in
531 R v4.0.3. For our pausing analyses, we defined the transcription start site (TSS) region as -300
532 bp to +300 bp surrounding the TSS. RNA Pol II ChIP-seq read densities for the TSS regions of
533 genes displayed a bimodal distribution which was assumed to be two partially overlapping
534 distributions of genes with high and low RNA Pol II TSS signal. We examined those genes with
535 a substantial level of RNA Pol II signal in the TSS region of control samples, defined as those
536 genes present in the upper mode of the distribution of log-transformed TSS read densities
537 based on a cutoff for type-I error of 10% using the R cutoff package v0.1.0
538 (https://github.com/choisy/cutoff). The read density in the TSS region was divided by the read
539 density in the region extending from +301 to +1301 bp with respect to the TSS to calculate the
540 pause release ratio (PRR). ChIP-seq data was visualized in Python 3.10.0 using the plotting
541 tool SparK v2.6.2[67].
542
543 *Nuclei isolation for single cell sequencing*
544

545 Nuclei were isolated from 200 11-12 somite stage zebrafish embryos for 10x Genomics
546 Multiome sequencing using recommended procedures with slight modifications. Embryos were
547 deyolked in Deyolking buffer (55 mM NaCl, 1.8 mM KCl, 1.25 mM NaHCO3) and then
548 embryonic cells were partially dissociated by pipetting in DMEM/F12 media (Gibco 21041025).
549 Cells were pelleted and then lysed on ice for 5 minutes in 100 µl of chilled 0.1X Lysis buffer (10
550 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.01% Tween-20, 0.01% IGEPAL CA-630, 0.001%
551 digitonin, 1% BSA, 1 mM DTT). 1 ml of Wash buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM
552 $MgCl_2$, 0.1% Tween-20, 1% BSA, 1 mM DTT, 0.1 U/µl Sigma Protector RNase Inhibitor) was
553 then added to the cell lysis and the nuclei were pelleted by centrifuging at 500xg for 5 minutes
554 at 4°C. Two more washes with 1 ml of Wash buffer were performed. Nuclei were then
555 resuspended in 50 µl of Nuclei Dilution Buffer (1X 10x Genomics Single Cell Multiome Nuclei
556 Buffer, 1 mM DTT, 1 U/µl Sigma Protector RNase Inhibitor) and filtered with a 40 µm Flowmi cell
557 strainer. Approximately 10,000 nuclei per sample were targeted for capture with the 10x
558 Chromium Single Cell Multiome ATAC + Gene Expression system by the UCLA Technology
559 Center for Genomics & Bioinformatics.
560
561 *Single cell sequencing analysis*
562
563 10x Chromium Single Cell Multiome ATAC + Gene Expression libraries were prepared by the
564 UCLA Technology Center for Genomics & Bioinformatics following the manufacturer's
565 recommended protocols and sequenced with an Illumina Novaseq 6000 system. Reads from
566 RNA-seq and ATAC-seq assays were aligned to the zebrafish GRCz11 genome with the
567 Lawson laboratory's improved transcriptome annotation[68] using 10x Genomics' Cell Ranger
568 ARC software v2.0.0 with default parameters. Mapped reads from single cell RNA-seq and
569 ATAC-seq assays were analyzed in R v4.2.2 using Seurat v4.3.0[69]. Good quality cells were
570 retained by filtering using Seurat with the following thresholds: 15000 > RNA UMI > 2000,

571 Genes > 1500, 20000 > ATAC UMI > 5000, TSS enrichment score > 3.5. Filtered control and
572 *rtf1* morphant cell datasets were integrated using Harmony v0.1.1[70], and preliminary
573 dimensional reduction, cell clustering, and marker gene identification were carried out in Seurat.
574 Clusters of cells that exhibited enriched expression of mitochondrial and ribosomal transcript
575 expression levels but lacked enrichment for other genes were considered abnormal cells and
576 were manually removed from the dataset. Following manual filtering, a final dataset of 14340
577 control and 13839 *rtf1* morphant cells was again integrated with Harmony and multimodal
578 dimensional reduction (weighted-nearest neighbor method with RNA and ATAC assays), cell
579 clustering, and marker gene identification were carried out in Seurat. Cluster cell number
580 proportions and gene expression levels were visualized with Seurat and with published code[71].
581 Based on marker gene expression, cells that were members of clusters corresponding to
582 anterior lateral plate mesoderm (*sema3e*+), endothelium (*flt4*+), cardiac precursor (*mef2cb*+),
583 and rostral blood (*spi1b*+) were extracted, independently integrated with Harmony, and
584 dimensional reduction (RNA assay only), cell clustering, and marker gene identification were
585 carried out in Seurat. Pseudotime analysis of Harmony/Seurat cell embeddings and plotting of
586 gene expression over pseudotime was performed with Monocle v3.0[72].
587
588 *Statistical analysis*
589
590 Statistical significance of gene expression changes in qPCR analyses was determined based on
591 2 (for *Nkx2-5* and *Afp* expression) or 3 (for *Myh6*, *Nppa*, *Brachyury*, and *Ncam1* expression)
592 independent experiments using a two-tailed Student's *t*-test with a *p*-value of less than 0.05
593 considered to be significant. Statistical significance of the difference between beating
594 cardiomyocyte cluster formation in control and Rtf1 shRNA mESCs was determined based on 4
595 independent experiments using a two-tailed Student's *t*-test with a *p*-value of less than 0.05
596 considered to be significant. To compare PRR means between uninjected, *rtf1* morphant, and

597 flavopiridol-treated *rtf1* morphant embryos, mean PRR values were first log-transformed (log10)
598 and then a Welch's paired two-tailed *t*-test was carried out with a *p*-value of less than 0.05
599 considered to be significant. Significantly changed PRRs for individual genes were determined
600 based on 3 independent replicates per condition using the rowttests function in the R package
601 genefilter v1.72.1 to perform two-tailed *t*-tests. The p.adjust function in the R stats package
602 v4.0.3 using a Benjamini and Hochberg (false discovery rate) correction for multiple
603 comparisons was then applied with adjusted *p*-values of less than 0.1 considered to be
604 significant.
605
606 **Acknowledgements**
607
608 We thank the UCLA Technology Center for Genomics & Bioinformatics for technical assistance
609 with Next Generation Sequencing and the Broad Stem Cell Research Center Flow Cytometry
610 Core Facility for technical assistance with cell sorting. Figures 3A, 4A, and 5C were created
611 with BioRender.com. This work was supported by funding from the NIH/NHLBI to J.-N.C.
612 (R01HL155905 and R01HL140472).
613
614 **Author Contributions**
615
616 A.D.L.: study design, data acquisition, analysis, and interpretation, manuscript writing. F.L.:
617 study design, data acquisition, analysis, and interpretation. L.T.: data acquisition. Z.Y.S.H.: data
618 acquisition. B.K.: data acquisition. Z.T.: data acquisition. A.E.: data acquisition. M.P.: data
619 analysis. H.N.: study design and data interpretation. A.N.: study design and data interpretation.
620 J.-N.C.: study design, data interpretation, manuscript writing. All authors approved the final
621 manuscript.
622

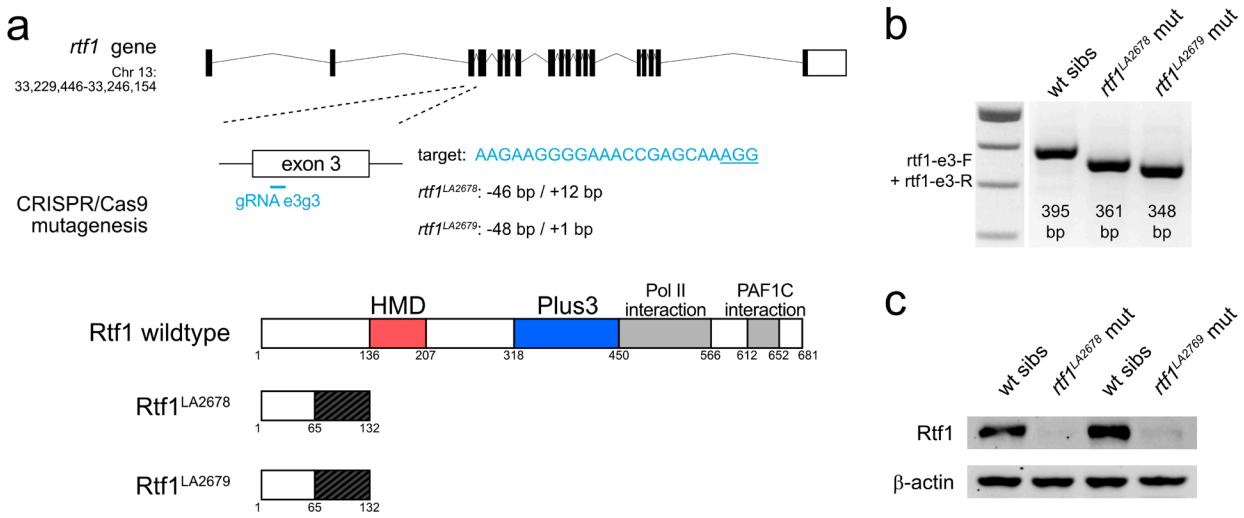
Competing Interests

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).

649



650

651 **Figure 1. Generation of zebrafish *rtf1* null mutants with CRISPR/Cas9.** (a) Schematic of
652 CRISPR/Cas9 mutagenesis of the *rtf1* gene. Two mutant alleles, *rtf1*^{LA2678} and *rtf1*^{LA2679}, were
653 recovered after targeting of *rtf1* exon 3. Both alleles are predicted to disrupt translation of the
654 Rtf1 protein and eliminate the HMD, Plus3, Pol II interaction, and PAF1C interaction domains.
655 (b) Agarose gel electrophoresis results of genotyping *rtf1* mutants by PCR. Deletions in
656 *rtf1*^{LA2678} and *rtf1*^{LA2679} alleles can be distinguished from wild type allele using primers rtf1-e3-F
657 and rtf1-e3-R. (c) Western blot detecting Rtf1 and β-actin (loading control) proteins in lysates of
658 wild type and *rtf1* mutant embryos. Image is representative of two independent experiments.

659

660

661

662

663

664

665

666

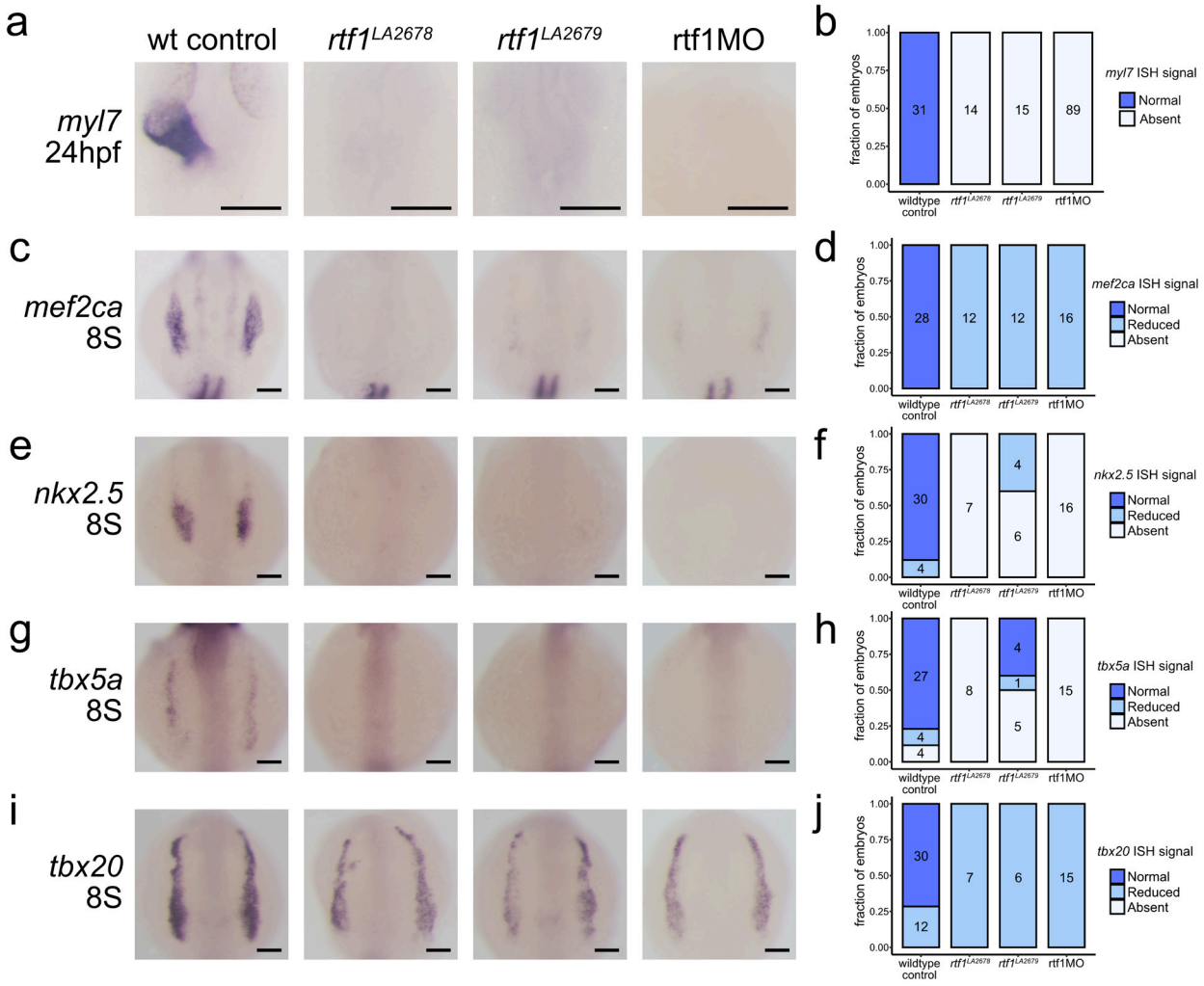


Figure 2. Characterization of early cardiac development in Rtf1 deficient embryos. (a) Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf zebrafish embryos. (b) Quantification of *myl7* signal intensity in control and Rtf1 deficient embryos at 24 hpf. Numbers on 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

678 analyzed. (g) Representative images of RNA in situ hybridization detecting *tbx5a* expression in
679 8 somite stage zebrafish embryos. (h) Quantification of *tbx5a* signal intensity in control and Rtf1
680 deficient embryos at the 8 somite stage. Numbers on bars indicate the number of embryos
681 analyzed. (i) Representative images of RNA in situ hybridization detecting *tbx20* expression in 8
682 somite stage zebrafish embryos. (j) Quantification of *tbx20* signal intensity in control and Rtf1
683 deficient embryos at the 8 somite stage. Numbers on bars indicate the number of embryos
684 analyzed. Scale bars in a, c, e, g, and i represent 0.1 mm.
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703

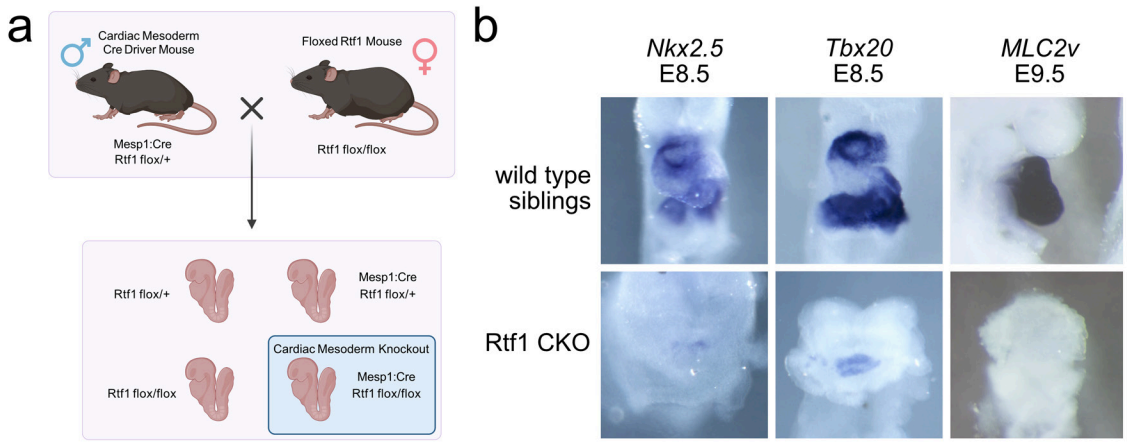
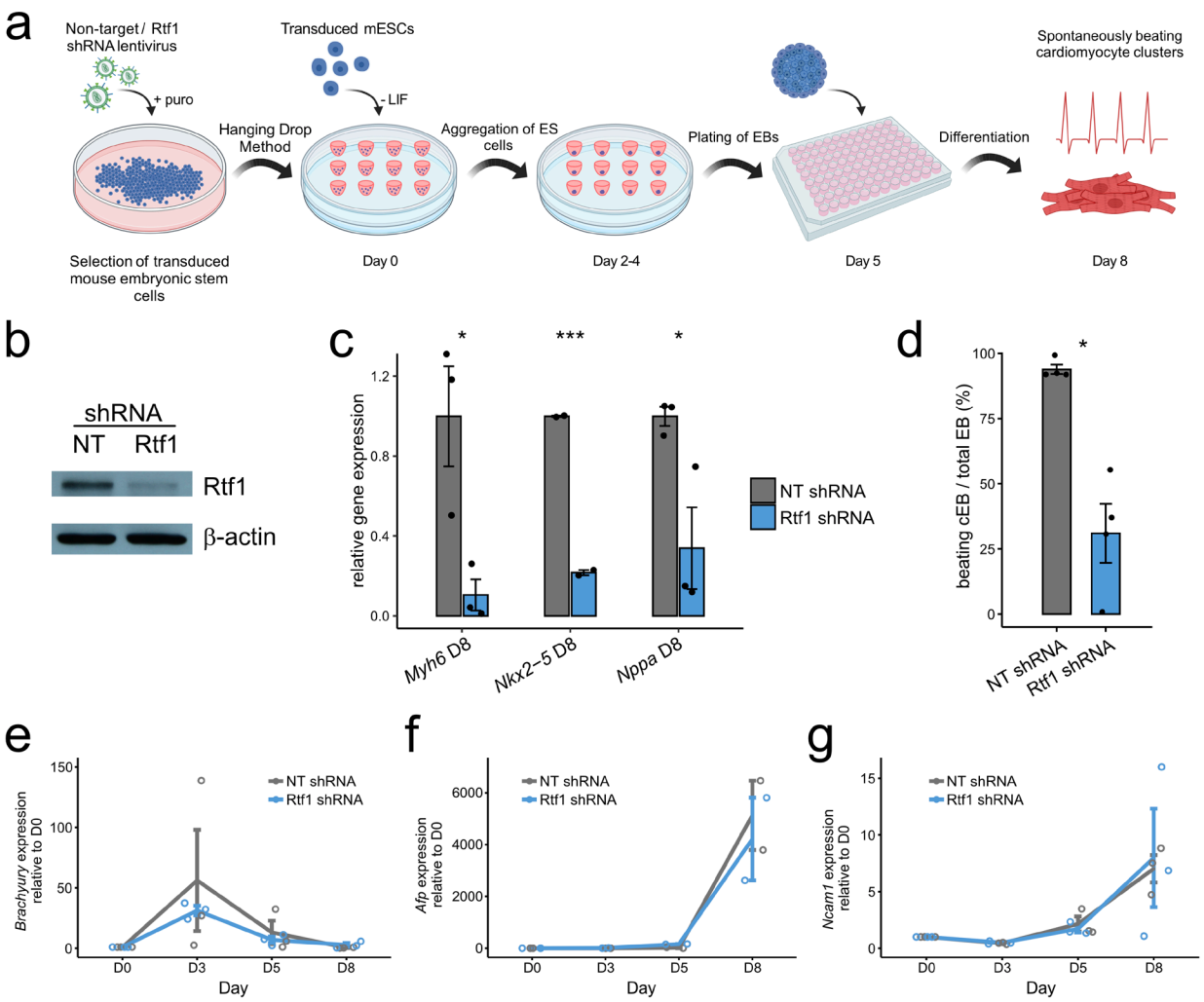


Figure 3. Failure of cardiac differentiation upon loss of Rtf1 in the mammalian cardiac mesoderm. (a) Diagram of generation of cardiac mesoderm-specific Rtf1 knockout mouse embryos. A Mesp1: Cre insertion allele was bred into an Rtf1 exon 3 floxed background. Heterozygous Rtf1 floxed male mice with Mesp1:Cre 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.



723

724 **Figure 4. Failure of cardiac differentiation from Rtf1 knockdown mouse embryonic stem**

725 **cells.** (a) Diagram of shRNA knockdown of Rtf1 in mouse embryonic stem cells (mESCs) and

726 differentiation of plated EBs. mESCs were transduced with non-target control or Rtf1 shRNA

727 lentivirus and selected with puromycin. Transduced mESCs were grown into embryoid bodies

728 (EBs) using the hanging drop method which were then plated to examine differentiation into cell

729 types including beating cardiomyocyte clusters. (b) Western blot verifying reduction of Rtf1

730 protein in Rtf1 shRNA mESCs (~70% reduced based on densitometry) compared to unchanged

731 level of the loading control protein β-actin. Image is representative of three independent

732 experiments. (c) qPCR analysis of *Myh6*, *Nkx2-5*, and *Nppa* expression in non-target control

733 (NT) and Rtf1 shRNA plated EBs. Data are normalized to the mean expression level in NT

734 shRNA samples and error bars indicate the standard error of the mean. (d) Mean percentage (±
735 standard error) of plated EBs exhibiting beating cardiomyocyte differentiation (cEBs) in NT and
736 Rtf1 shRNA plated EBs. (e) qPCR analysis of *Brachyury* expression in non-target control (NT)
737 and Rtf1 shRNA plated EBs. (f) qPCR analysis of *Afp* expression in non-target control (NT) and
738 Rtf1 shRNA plated EBs. (g) qPCR analysis of *Ncam1* expression in non-target control (NT) and
739 Rtf1 shRNA plated EBs. Data in e, f, and g are normalized to the mean expression level in NT
740 shRNA samples at day 0 (D0) and error bars indicate the standard error of the mean. *: $p <$
741 0.05, ***: $p < 0.001$.
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759

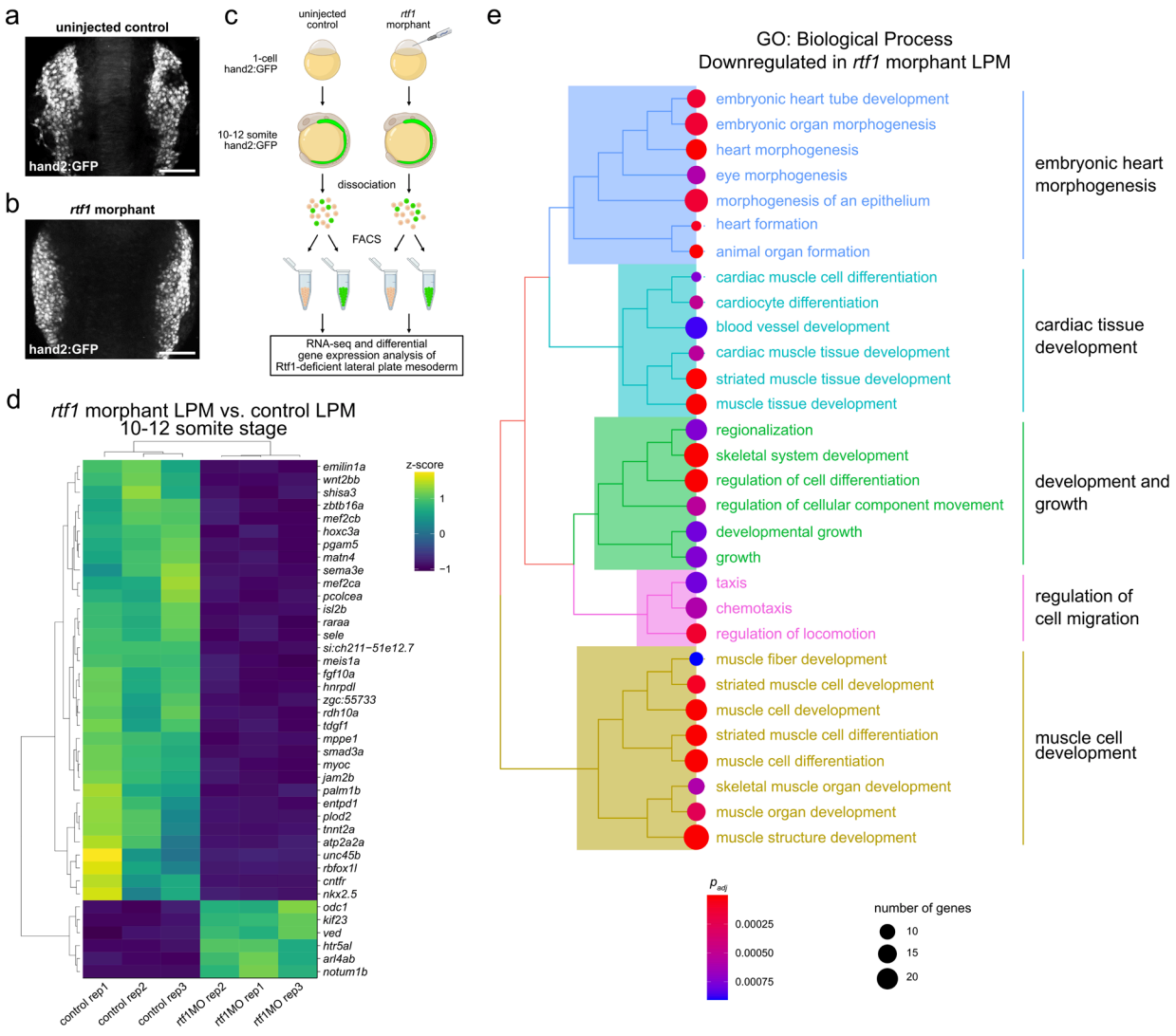


Figure 5. Dysregulation of cardiac gene expression in the Rtf1 deficient lateral plate mesoderm. (a,b) Projections of confocal z-stacks of hand2:GFP signal in 11 somite stage uninjected control (a) and *rtf1* morphant (b) embryos. Scale bars represent 100 μm. (c) Diagram of lateral plate mesoderm (LPM) 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

770 hand2:GFP-positive LPM. High and low z-scores represent high and low gene expression,
771 respectively. (e) Hierarchical clustering of Biological Process gene ontology (GO) terms based
772 on semantic similarity for the genes most significantly downregulated in the *rtf1* morphant
773 hand2:GFP-positive LPM.
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795

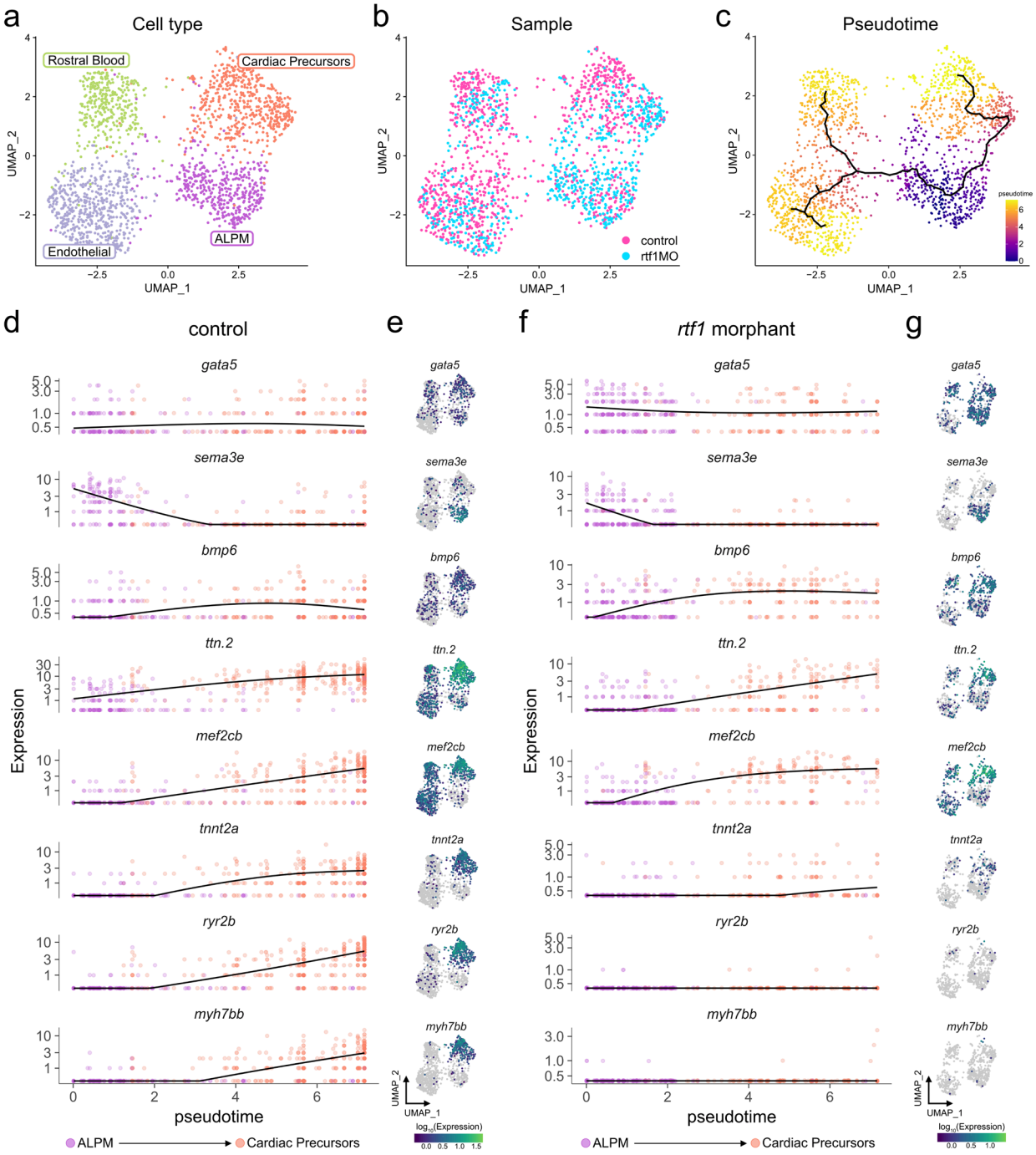


Figure 6. Rtf1 is required for transcriptional progression of cardiac precursors to a more mature state. (a-c) UMAP plots of anterior lateral plate mesoderm (ALPM) and derivatives from merged and integrated single cell RNA-seq datasets of 11-12 somite stage uninjected control and *rtf1* morphant embryos. Cells in plots are colored by cell type (a), sample (b), and

801 pseudotime (c). (d,f) Expression dynamics plots displaying expression levels (y-axis) for
802 selected genes in uninjected control (d) and *rtf1* morphant (f) cells over the pseudotime
803 trajectory (x-axis) from ALPM (root) to cardiac precursor state shown in c. (e,g) UMAP plots of
804 uninjected control (e) and *rtf1* morphant (g) cells colored by gene expression levels.
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826

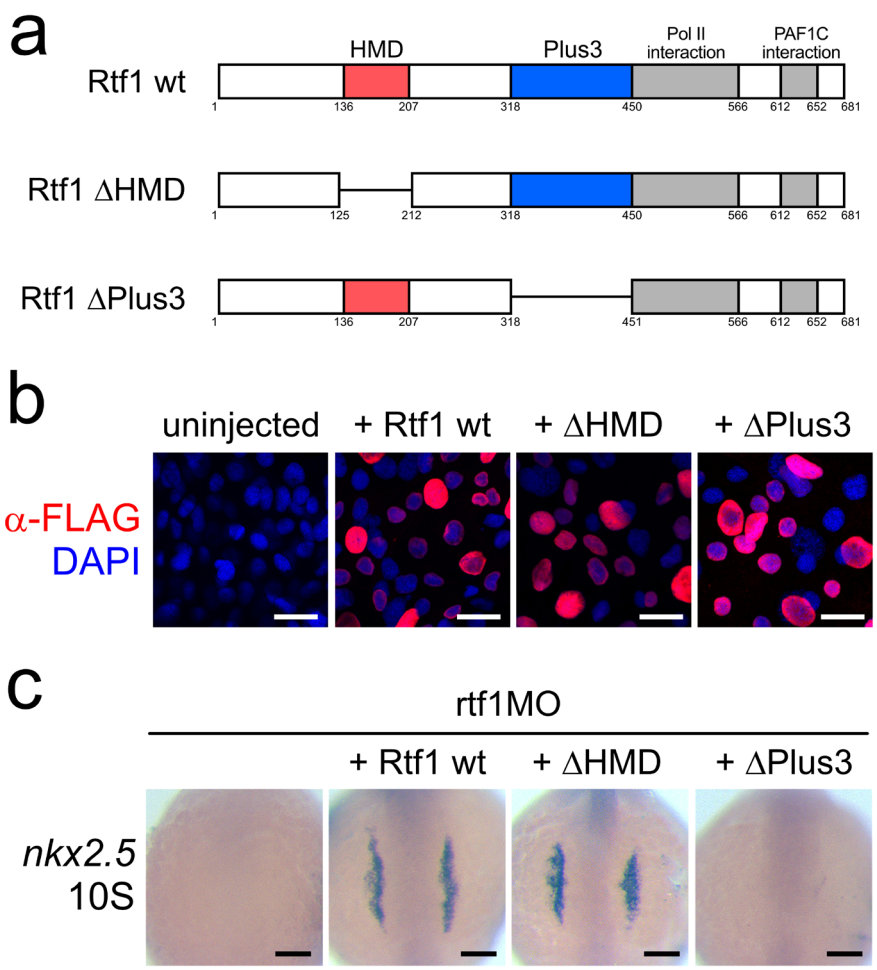


Figure 7. Rtf1's Plus3 domain is necessary for cardiac progenitor formation. (a) Schematics of Rtf1 wildtype (wt) and Rtf1 mutant constructs (ΔHMD and ΔPlus3). Domains are indicated by colored boxes (HMD: red, Plus3: blue, PolII and PAF1C interaction: grey). Deleted regions are represented by lines. 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.

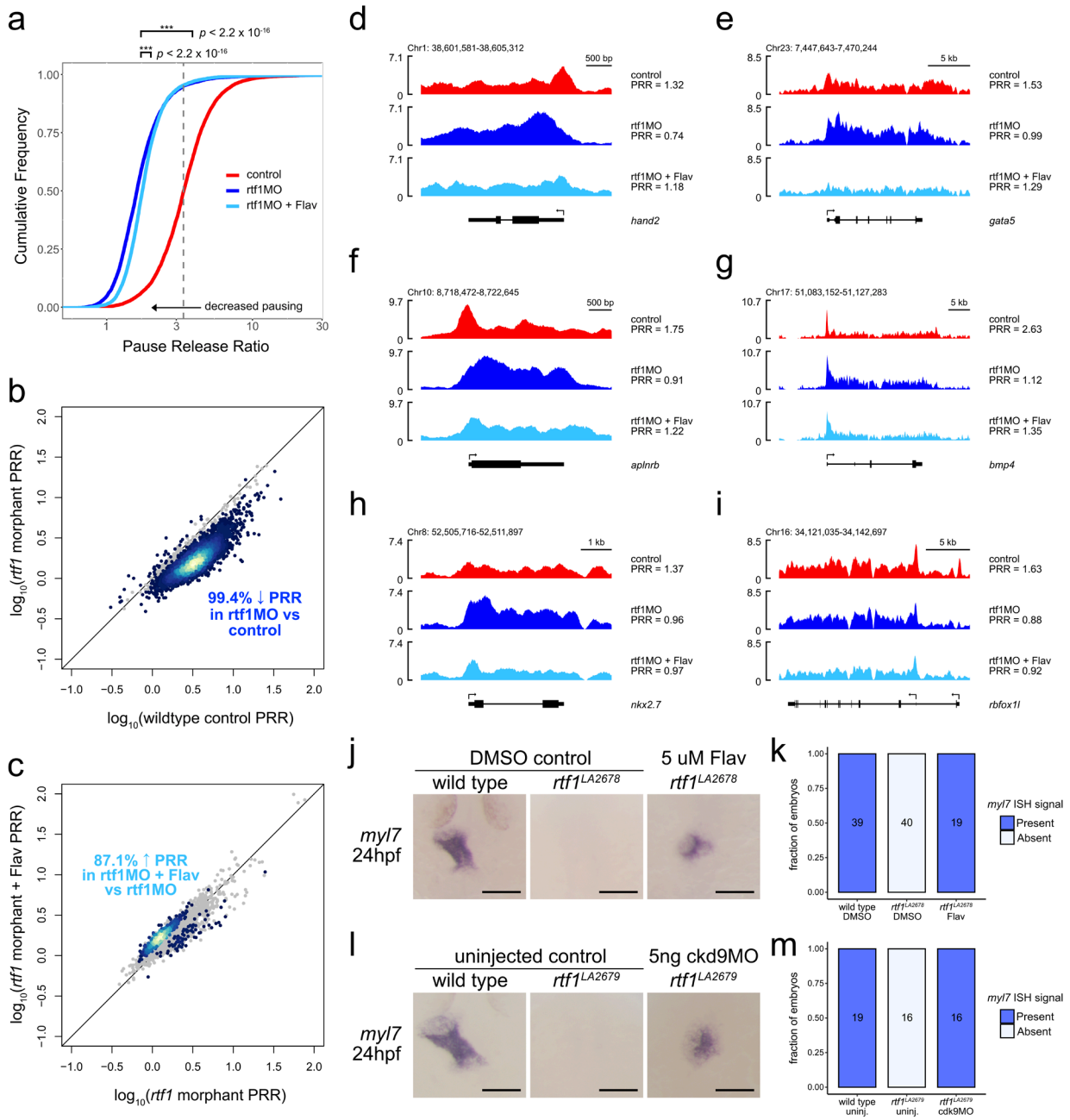


Figure 8. Rtf1-dependent transcriptional pausing regulates cardiogenesis. (a) Cumulative frequency plot of pause release ratios (PRRs) of 6,078 genes with substantial RNA Pol II signal. PRRs in control (red) and flavopiridol treated *rtf1* morphant (light blue) samples differed significantly from untreated *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

846 for a single gene that differ significantly (colored point) or are not significantly different (grey
847 point) between controls and *rtf1* morphants. (c) Dot plot comparing PRRs in *rtf1* morphants and
848 flavopiridol-treated *rtf1* morphants. Each dot represents the PRR values for a single gene that
849 differ significantly (colored point) or are not significantly different (grey point) between *rtf1*
850 morphants and flavopiridol-treated *rtf1* morphants. Dot colors in b and c are based on the
851 density of points, with lighter colors indicating more dense points. (d-i) Representative ChIP-seq
852 tracks displaying RNA Pol II read densities (y-axis) for control (red), *rtf1* morphant (blue), and
853 flavopiridol-treated *rtf1* morphant (light blue) embryos at cardiac mesoderm-related genes
854 including *hand2* (d), *gata5* (e), *aplnrb* (f), *bmp4* (g), *nkx2.7* (h), and *rbfox1l* (i). (j)
855 Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf control
856 and Rtf1 deficient (± flavopiridol) zebrafish embryos. (k) Quantification of *myl7* signal intensity in
857 control and Rtf1 deficient (± flavopiridol) embryos at 24 hpf. Numbers on bars indicate the
858 number of embryos analyzed. (l) Representative images of RNA in situ hybridization detecting
859 *myl7* expression in 24 hpf control and Rtf1 deficient (± *cdk9* morpholino) zebrafish embryos. (m)
860 Quantification of *myl7* signal intensity in control and Rtf1 deficient (± *cdk9* morpholino) embryos
861 at 24 hpf. Numbers on bars indicate the number of embryos analyzed.
862
863 **Supplementary Data**
864
865 **Supplementary Table 1.** Sequences of primers used for quantitative real-time PCR analysis of
866 gene expression in mouse embryonic stem cells.
867
868 **Supplementary Figure 1.** Gating strategy for 10-12 somite stage zebrafish hand2:GFP FACS.
869 Abnormal cells or multiplets were excluded by gating out events with high side scatter (SSC)
870 and forward scatter (FSC). Live hand2:GFP-positive cells (pink cells in panel d) were sorted
871 based on high GFP and low DAPI signal. (a) Gating control experiment using unstained cells

872 from non-transgenic embryos. (b) Gating control experiment using unstained cells from
873 hand2:GFP transgenic embryos. (c) Gating control experiment using DAPI-stained cells from
874 non-transgenic embryos. (d) Gating of sorting sample for analysis using DAPI-stained cells from
875 hand2:GFP transgenic embryos. 0.6% of the total cells were sorted as live GFP-positive cells.
876 Plots are representative of 3 independent experiments.
877
878 **Supplementary Figure 2.** UMAP plots of 11-12 somite stage control and *rtf1* morphant embryo
879 integrated multimodal (ATAC + Gene Expression) single cell sequencing datasets split based
880 on sample (control on left vs. *rtf1* morphant on right). Cells are colored based on Seurat cluster.
881
882 **Supplementary Figure 3.** Clustered dot plot displaying expression of marker genes in cells
883 belonging to each Seurat cluster for the dataset displayed in Supplementary Figure 2. Each
884 cluster is further split by sample (control vs. *rtf1* morphant).
885
886 **Supplementary Figure 4.** Comparison of proportions of cells belonging to each Seurat cluster
887 in control and *rtf1* morphant samples for the dataset displayed in Supplementary Figure 2.
888 Predicted identities of each cell type based on marker gene expression are displayed to the left
889 of the Seurat cluster IDs. Cell type names are colored based on whether they are significantly
890 expanded (pink) or significantly reduced (light blue) in *rtf1* morphants based on Fisher's exact
891 test with an adjusted *p*-value cutoff of 0.05.
892
893 **Supplementary Data File 1.** DESeq2 results file containing differential gene expression
894 analysis data comparing FACS-isolated hand2:GFP-positive LPM from control and *rtf1*
895 morphant 10-12 somite stage embryos.
896

897 **Supplementary Data File 2.** Significantly restricted marker genes for Seurat cell clusters from
898 integrated multimodal (ATAC + Gene Expression) single cell sequencing dataset from control
899 and *rtf1* morphant embryos at the 11-12 somite stage.
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922

923 **References**:

924

925 1 Bruneau, B. G. Signaling and transcriptional networks in heart development and regeneration.
926 *Cold Spring Harb Perspect Biol* **5**, a008292, doi:10.1101/cshperspect.a008292 (2013).
927 2 Paige, S. L., Plonowska, K., Xu, A. & Wu, S. M. Molecular regulation of cardiomyocyte
928 differentiation. *Circ Res* **116**, 341-353, doi:10.1161/CIRCRESAHA.116.302752 (2015).
929 3 Rana, M. S., Christoffels, V. M. & Moorman, A. F. A molecular and genetic outline of cardiac
930 morphogenesis. *Acta Physiol (Oxf)* **207**, 588-615, doi:10.1111/apha.12061 (2013).
931 4 Lu, F., Langenbacher, A. D. & Chen, J. N. Transcriptional Regulation of Heart Development in
932 Zebrafish. *J Cardiovasc Dev Dis* **3**, doi:10.3390/jcdd3020014 (2016).
933 5 Schoenebeck, J. J., Keegan, B. R. & Yelon, D. Vessel and blood specification override cardiac
934 potential in anterior mesoderm. *Dev Cell* **13**, 254-267, doi:10.1016/j.devcel.2007.05.012 (2007).
935 6 Mao, L. M. F., Boyle Anderson, E. A. T. & Ho, R. K. Anterior lateral plate mesoderm gives rise to
936 multiple tissues and requires tbx5a function in left-right asymmetry, migration dynamics, and
937 cell specification of late-addition cardiac cells. *Dev Biol* **472**, 52-66,
938 doi:10.1016/j.ydbio.2021.01.007 (2021).
939 7 Duong, T. B., Holowiecki, A. & Waxman, J. S. Retinoic acid signaling restricts the size of the first
940 heart field within the anterior lateral plate mesoderm. *Dev Biol* **473**, 119-129,
941 doi:10.1016/j.ydbio.2021.02.005 (2021).
942 8 Sorrell, M. R. & Waxman, J. S. Restraint of Fgf8 signaling by retinoic acid signaling is required for
943 proper heart and forelimb formation. *Dev Biol* **358**, 44-55, doi:10.1016/j.ydbio.2011.07.022
944 (2011).

945 9 Hota, S. K. *et al.* Dynamic BAF chromatin remodeling complex subunit inclusion promotes
946 temporally distinct gene expression programs in cardiogenesis. *Development* **146**,
947 doi:10.1242/dev.174086 (2019).
948 10 Takeuchi, J. K. *et al.* Chromatin remodelling complex dosage modulates transcription factor
949 function in heart development. *Nat Commun* **2**, 187, doi:10.1038/ncomms1187 (2011).
950 11 Lickert, H. *et al.* Baf60c is essential for function of BAF chromatin remodelling complexes in
951 heart development. *Nature* **432**, 107-112, doi:10.1038/nature03071 (2004).
952 12 Lou, X., Deshwar, A. R., Crump, J. G. & Scott, I. C. Smarcd3b and Gata5 promote a cardiac
953 progenitor fate in the zebrafish embryo. *Development* **138**, 3113-3123, doi:10.1242/dev.064279
954 (2011).
955 13 Takeuchi, J. K. & Bruneau, B. G. Directed transdifferentiation of mouse mesoderm to heart tissue
956 by defined factors. *Nature* **459**, 708-711, doi:10.1038/nature08039 (2009).
957 14 Endoh, M. *et al.* Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro.
958 *Mol Cell Biol* **24**, 3324-3336, doi:10.1128/MCB.24.8.3324-3336.2004 (2004).
959 15 Hartzog, G. A., Wada, T., Handa, H. & Winston, F. Evidence that Spt4, Spt5, and Spt6 control
960 transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. *Genes Dev* **12**, 357-
961 369, doi:10.1101/gad.12.3.357 (1998).
962 16 Keegan, B. R. *et al.* The elongation factors Pandora/Spt6 and Foggy/Spt5 promote transcription
963 in the zebrafish embryo. *Development* **129**, 1623-1632, doi:10.1242/dev.129.7.1623 (2002).
964 17 Wada, T. *et al.* DSIF, a novel transcription elongation factor that regulates RNA polymerase II
965 processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* **12**, 343-356,
966 doi:10.1101/gad.12.3.343 (1998).

967 18 Wier, A. D., Mayekar, M. K., Heroux, A., Arndt, K. M. & VanDemark, A. P. Structural basis for
968 Spt5-mediated recruitment of the Paf1 complex to chromatin. *Proc Natl Acad Sci U S A* **110**,
969 17290-17295, doi:10.1073/pnas.1314754110 (2013).
970 19 Chen, J. J. *et al.* Spt5 Phosphorylation and the Rtf1 Plus3 Domain Promote Rtf1 Function through
971 Distinct Mechanisms. *Mol Cell Biol* **40**, doi:10.1128/MCB.00150-20 (2020).
972 20 Hu, S. *et al.* SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains
973 the enhancer landscape. *Mol Cell* **81**, 4425-4439 e4426, doi:10.1016/j.molcel.2021.08.029
974 (2021).
975 21 Langenbacher, A. D. *et al.* The PAF1 complex differentially regulates cardiomyocyte
976 specification. *Dev Biol* **353**, 19-28, doi:10.1016/j.ydbio.2011.02.011 (2011).
977 22 Park, J., Park, S. & Lee, J. S. Role of the Paf1 complex in the maintenance of stem cell
978 pluripotency and development. *FEBS J* **290**, 951-961, doi:10.1111/febs.16582 (2023).
979 23 Van Oss, S. B., Cucinotta, C. E. & Arndt, K. M. Emerging Insights into the Roles of the Paf1
980 Complex in Gene Regulation. *Trends Biochem Sci* **42**, 788-798, doi:10.1016/j.tibs.2017.08.003
981 (2017).
982 24 Jaehning, J. A. The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim*
983 *Biophys Acta* **1799**, 379-388, doi:10.1016/j.bbagrm.2010.01.001 (2010).
984 25 Nguyen, C. T., Langenbacher, A., Hsieh, M. & Chen, J. N. The PAF1 complex component Leo1 is
985 essential for cardiac and neural crest development in zebrafish. *Dev Biol* **341**, 167-175,
986 doi:10.1016/j.ydbio.2010.02.020 (2010).
987 26 Lu, F., Langenbacher, A. & Chen, J. N. Tbx20 drives cardiac progenitor formation and
988 cardiomyocyte proliferation in zebrafish. *Dev Biol* **421**, 139-148,
989 doi:10.1016/j.ydbio.2016.12.009 (2017).

990 27 Langenbacher, A. D. *et al.* Rtf1 Transcriptionally Regulates Neonatal and Adult Cardiomyocyte
991 Biology. *J Cardiovasc Dev Dis* **10**, doi:10.3390/jcdd10050221 (2023).
992 28 Saga, Y. *et al.* MesP1 is expressed in the heart precursor cells and required for the formation of a
993 single heart tube. *Development* **126**, 3437-3447, doi:10.1242/dev.126.15.3437 (1999).
994 29 Warner, M. H., Roinick, K. L. & Arndt, K. M. Rtf1 is a multifunctional component of the Paf1
995 complex that regulates gene expression by directing cotranscriptional histone modification. *Mol*
996 *Cell Biol* **27**, 6103-6115, doi:10.1128/MCB.00772-07 (2007).
997 30 Vos, S. M., Farnung, L., Linden, A., Urlaub, H. & Cramer, P. Structure of complete Pol II-DSIF-PAF-
998 SPT6 transcription complex reveals RTF1 allosteric activation. *Nat Struct Mol Biol* **27**, 668-677,
999 doi:10.1038/s41594-020-0437-1 (2020).
1000 31 Cucinotta, C. E., Hildreth, A. E., McShane, B. M., Shirra, M. K. & Arndt, K. M. The nucleosome
1001 acidic patch directly interacts with subunits of the Paf1 and FACT complexes and controls
1002 chromatin architecture in vivo. *Nucleic Acids Res* **47**, 8410-8423, doi:10.1093/nar/gkz549 (2019).
1003 32 Van Oss, S. B. *et al.* The Histone Modification Domain of Paf1 Complex Subunit Rtf1 Directly
1004 Stimulates H2B Ubiquitylation through an Interaction with Rad6. *Mol Cell* **64**, 815-825,
1005 doi:10.1016/j.molcel.2016.10.008 (2016).
1006 33 Piro, A. S., Mayekar, M. K., Warner, M. H., Davis, C. P. & Arndt, K. M. Small region of Rtf1 protein
1007 can substitute for complete Paf1 complex in facilitating global histone H2B ubiquitylation in
1008 yeast. *Proc Natl Acad Sci U S A* **109**, 10837-10842, doi:10.1073/pnas.1116994109 (2012).
1009 34 de Jong, R. N. *et al.* Structure and DNA binding of the human Rtf1 Plus3 domain. *Structure* **16**,
1010 149-159, doi:10.1016/j.str.2007.10.018 (2008).
1011 35 Core, L. & Adelman, K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene
1012 regulation. *Genes Dev* **33**, 960-982, doi:10.1101/gad.325142.119 (2019).

1013 36 Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in
1014 metazoans. *Nat Rev Genet* **13**, 720-731, doi:10.1038/nrg3293 (2012).
1015 37 Peng, J., Zhu, Y., Milton, J. T. & Price, D. H. Identification of multiple cyclin subunits of human P-
1016 TEFb. *Genes Dev* **12**, 755-762, doi:10.1101/gad.12.5.755 (1998).
1017 38 Price, D. H. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol*
1018 *Cell Biol* **20**, 2629-2634, doi:10.1128/MCB.20.8.2629-2634.2000 (2000).
1019 39 Paparidis, N. F., Durvale, M. C. & Canduri, F. The emerging picture of CDK9/P-TEFb: more than
1020 20 years of advances since PITALRE. *Mol Biosyst* **13**, 246-276, doi:10.1039/c6mb00387g (2017).
1021 40 Fujinaga, K., Huang, F. & Peterlin, B. M. P-TEFb: The master regulator of transcription elongation.
1022 *Mol Cell* **83**, 393-403, doi:10.1016/j.molcel.2022.12.006 (2023).
1023 41 Jonkers, I., Kwak, H. & Lis, J. T. Genome-wide dynamics of Pol II elongation and its interplay with
1024 promoter proximal pausing, chromatin, and exons. *Elife* **3**, e02407, doi:10.7554/eLife.02407
1025 (2014).
1026 42 Huang, M. *et al.* Negative Elongation Factor (NELF) Inhibits Premature Granulocytic
1027 Development in Zebrafish. *Int J Mol Sci* **23**, doi:10.3390/ijms23073833 (2022).
1028 43 Yang, Q. *et al.* RNA polymerase II pausing modulates hematopoietic stem cell emergence in
1029 zebrafish. *Blood* **128**, 1701-1710, doi:10.1182/blood-2016-02-697847 (2016).
1030 44 Gaertner, B. & Zeitlinger, J. RNA polymerase II pausing during development. *Development* **141**,
1031 1179-1183, doi:10.1242/dev.088492 (2014).
1032 45 Gilchrist, D. A. *et al.* Regulating the regulators: the pervasive effects of Pol II pausing on
1033 stimulus-responsive gene networks. *Genes Dev* **26**, 933-944, doi:10.1101/gad.187781.112
1034 (2012).

1035 46 Adelman, K. *et al.* Immediate mediators of the inflammatory response are poised for gene
1036 activation through RNA polymerase II stalling. *Proc Natl Acad Sci U S A* **106**, 18207-18212,
1037 doi:10.1073/pnas.0910177106 (2009).
1038 47 Rougvie, A. E. & Lis, J. T. The RNA polymerase II molecule at the 5' end of the uninduced hsp70
1039 gene of D. melanogaster is transcriptionally engaged. *Cell* **54**, 795-804, doi:10.1016/s0092-
1040 8674(88)91087-2 (1988).
1041 48 Chen, F. X. *et al.* PAF1 regulation of promoter-proximal pause release via enhancer activation.
1042 *Science* **357**, 1294-1298, doi:10.1126/science.aan3269 (2017).
1043 49 Gilchrist, D. A. *et al.* Pausing of RNA polymerase II disrupts DNA-specified nucleosome
1044 organization to enable precise gene regulation. *Cell* **143**, 540-551,
1045 doi:10.1016/j.cell.2010.10.004 (2010).
1046 50 Gilchrist, D. A. *et al.* NELF-mediated stalling of Pol II can enhance gene expression by blocking
1047 promoter-proximal nucleosome assembly. *Genes Dev* **22**, 1921-1933, doi:10.1101/gad.1643208
1048 (2008).
1049 51 Jurynec, M. J. *et al.* The Paf1 complex and P-TEFb have reciprocal and antagonist roles in
1050 maintaining multipotent neural crest progenitors. *Development* **146**, doi:10.1242/dev.180133
1051 (2019).
1052 52 Wang, Z. *et al.* Coordinated regulation of RNA polymerase II pausing and elongation progression
1053 by PAF1. *Sci Adv* **8**, eabm5504, doi:10.1126/sciadv.abm5504 (2022).
1054 53 Westerfield, M. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio).* 4th
1055 ed. edn, (Univ. of Oregon Press, 2000).
1056 54 Farley, F. W., Soriano, P., Steffen, L. S. & Dymecki, S. M. Widespread recombinase expression
1057 using FLPeR (flipper) mice. *Genesis* **28**, 106-110 (2000).

1058 55 Chen, J. N. & Fishman, M. C. Zebrafish tinman homolog demarcates the heart field and initiates
1059 myocardial differentiation. *Development* **122**, 3809-3816 (1996).
1060 56 Akanuma, T., Koshida, S., Kawamura, A., Kishimoto, Y. & Takada, S. Paf1 complex homologues
1061 are required for Notch-regulated transcription during somite segmentation. *EMBO Rep* **8**, 858-
1062 863, doi:10.1038/sj.embor.7401045 (2007).
1063 57 Robu, M. E. *et al.* p53 activation by knockdown technologies. *PLoS Genet* **3**, e78,
1064 doi:10.1371/journal.pgen.0030078 (2007).
1065 58 Yin, C., Kikuchi, K., Hochgreb, T., Poss, K. D. & Stainier, D. Y. Hand2 regulates extracellular matrix
1066 remodeling essential for gut-looping morphogenesis in zebrafish. *Dev Cell* **18**, 973-984,
1067 doi:10.1016/j.devcel.2010.05.009 (2010).
1068 59 Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions,
1069 deletions and gene fusions. *Genome Biol* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).
1070 60 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for
1071 assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930,
1072 doi:10.1093/bioinformatics/btt656 (2014).
1073 61 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
1074 RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
1075 62 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological
1076 themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).
1077 63 Chicas, A. *et al.* Dissecting the unique role of the retinoblastoma tumor suppressor during
1078 cellular senescence. *Cancer Cell* **17**, 376-387, doi:10.1016/j.ccr.2010.01.023 (2010).
1079 64 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-
1080 359, doi:10.1038/nmeth.1923 (2012).

1081 65 Lun, A. T. & Smyth, G. K. csaw: a Bioconductor package for differential binding analysis of ChIP-
1082 seq data using sliding windows. *Nucleic Acids Res* **44**, e45, doi:10.1093/nar/gkv1191 (2016).
1083 66 Lun, A. T. & Smyth, G. K. De novo detection of differentially bound regions for ChIP-seq data
1084 using peaks and windows: controlling error rates correctly. *Nucleic Acids Res* **42**, e95,
1085 doi:10.1093/nar/gku351 (2014).
1086 67 Kurtenbach, S. & Harbour, J. W. SparK: A Publication-quality NGS Visualization Tool. *bioRxiv*,
1087 845529, doi:10.1101/845529 (2019).
1088 68 Lawson, N. D. *et al.* An improved zebrafish transcriptome annotation for sensitive and
1089 comprehensive detection of cell type-specific genes. *Elife* **9**, doi:10.7554/eLife.55792 (2020).
1090 69 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529,
1091 doi:10.1016/j.cell.2021.04.048 (2021).
1092 70 Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat*
1093 *Methods* **16**, 1289-1296, doi:10.1038/s41592-019-0619-0 (2019).
1094 71 Song, M. *et al.* GATA4/5/6 family transcription factors are conserved determinants of cardiac
1095 versus pharyngeal mesoderm fate. *Sci Adv* **8**, eabg0834, doi:10.1126/sciadv.abg0834 (2022).
1096 72 Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**,
1097 496-502, doi:10.1038/s41586-019-0969-x (2019).

1098