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:

Transcriptional pause-release critically regulates cellular RNA biogenesis, yet how dysregulation of this process impacts embryonic development is not fully understood. Rtf1 is a multifunctional transcription regulatory protein involved in modulating promoter-proximal pausing of RNA polymerase II. Using zebrafish and mouse as model systems, we show that Rtf1 activity is essential for the differentiation of the myocardial lineage from mesoderm. Ablation of *rtf1* impairs the formation of *nkx2.5*+/*tbx5a*+ cardiac progenitor cells, resulting in the development of embryos without cardiomyocytes. Structure-function analysis demonstrates that Rtf1's cardiogenic activity requires its Plus3 domain, which confers interaction with the

pausing/elongation factor Spt5. In Rtf1-deficient embryos, the occupancy of RNA Pol II at transcription start sites was reduced relative to downstream occupancy, suggesting a reduction in transcriptional pausing. Intriguingly, attenuating pause release by pharmacological inhibition or morpholino targeting of CDK9 improved RNA polymerase II occupancy at the transcription start sites of key cardiac genes and restored cardiomyocytes in Rtf1-deficient embryos. Thus, our findings demonstrate the crucial role that Rtf1-mediated transcriptional pausing plays in controlling the precise spatio-temporal transcription programs that govern early heart development.

Keywords:

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

Introduction:

Cardiogenesis relies on tightly controlled transcriptional programs regulated at all stages of mRNA synthesis. Decades of studies have identified an evolutionarily conserved network of transcription factors that direct embryonic heart development. 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 (ALPM) into multiple fates including the cardiac mesoderm, which is bounded anteriorly by blood and endothelial progenitors and posteriorly by pectoral fin progenitors[5-8].

53

54 Along with the paradigmatic role played by cardiac transcription factors in transcriptional
55 initiation, accumulating evidence suggests that transcriptional control of cardiogenesis is also
56 accomplished by epigenetic and post-initiation mechanisms. For example, the BAF (Brg1/Brm-
57 associated factor) complex, an ATP-dependent chromatin remodeling complex that actively
58 positions nucleosomes, is critical for normal vertebrate cardiogenesis[9-11]. The BAF complex
59 directly interacts with cardiac transcription factors including Tbx5, Nkx2-5, and Gata4[12] and
60 controls temporal changes in chromatin accessibility during cardiac differentiation[9]. Intriguingly,
61 forced expression of the BAF component Baf60c along with Tbx5/Nkx2-5/Gata4 can induce
62 cardiac differentiation from non-cardiac mesoderm[13], suggesting that nucleosomal positioning
63 and chromatin accessibility are key mechanisms regulating the activity of cardiac transcription
64 factors during cardiogenesis.

65

66 Cardiac gene expression is also sensitive to transcription elongation regulation. Spt6
67 (Suppressor of Ty 6) is a histone chaperone that promotes RNA Polymerase II elongation and
68 co-transcriptionally modulates chromatin structure[14,15]. In zebrafish, Spt6 interacts genetically
69 with the transcriptional pausing/elongation factor Spt5 (Suppressor of Ty 5) to support Nkx2.5+
70 cardiac progenitor specification[16]. While both Spt6 and Spt5 promote transcription elongation,
71 they likely do so via distinct mechanisms. Spt5 directly interacts with Spt4 to form the DSIF
72 (DRB Sensitivity Inducing Factor) complex, which has both inhibitory and stimulatory roles in
73 transcription[17]. Unphosphorylated Spt5 fosters promoter-proximal pausing of RNA Polymerase
74 II, but phosphorylation of Spt5's C-terminal repeat region by Cdk9 (cyclin dependent kinase 9)
75 promotes transcriptional pause release, and converts DSIF into an elongation-promoting
76 factor[18-20]. How the transcription regulatory activities of these elongation and pausing factors
77 facilitate cardiac development is not yet understood.

78

79 The Polymerase Associated Factor 1 Complex (PAF1C) consists of five core proteins (Paf1,
80 Ctr9, Cdc73, Rtf1, and Leo1) and plays wide-ranging transcription regulatory roles including
81 promoting co-transcriptional epigenetic modification of histones, supporting both pausing and
82 elongation of RNA Polymerase II, influencing mRNA 3' end formation, and serving as a bridge
83 between specific transcription factors and RNA Polymerase II[21-23]. In zebrafish, PAF1C subunits
84 differentially regulate cardiac development[24]. Leo1 activity is dispensable for the formation of
85 cardiac progenitors, but is required for differentiation of the cardiac chambers[25]. Embryos
86 lacking Ctr9, Cdc73, or Paf1, on the other hand, exhibit a deficiency of cardiomyocytes and only
87 form small, dysmorphic hearts. Intriguingly, loss of the Rtf1 subunit of the PAF1C results in the
88 most dramatic defects in cardiogenesis. Rtf1-deficient embryos have severe reduction in
89 *nkx2.5*+ cardiac progenitors and fail to generate any cardiomyocytes[24].
90
91 Rtf1 is a multifunctional transcription regulatory protein that modulates pausing and elongation
92 of RNA polymerase II, as well as histone epigenetic modifications. Given Rtf1's numerous roles
93 in transcription regulation, it remains an open question how Rtf1 mechanistically regulates
94 cardiogenesis. In this study, we showed that zygotic *rtf1* mutants, like *rtf1* morphants, lack
95 expression of most cardiac progenitor marker genes and produce no cardiomyocytes in
96 zebrafish. We also found that knockout of Rtf1 in the cardiogenic mesoderm of mouse embryos
97 and knockdown of Rtf1 in mouse ESCs reduced the formation of cardiac tissue, indicating that
98 the role of Rtf1 in specification of the cardiac lineage is conserved among vertebrates. Using
99 FACS and transcriptomic profiling approaches, we show that differentiation of the lateral plate
100 mesoderm (LPM) into cardiac progenitors requires Rtf1 activity. We further show that Rtf1's
101 Plus3 domain is essential for cardiac progenitor formation, while its HMD (Histone modification
102 domain) is dispensable. Consistent with the Plus3 domain's function as an interaction point
103 between Rtf1 and the DSIF component Spt5, Rtf1-deficient embryos display a genome-wide
104 decrease in promoter-proximal pausing of RNA Polymerase II. Excitingly, attenuating pause

release by chemical and morpholino inhibition of CDK9 improved pausing at cardiac genes and restored cardiac progenitor formation to Rtf1-deficient embryos. Together, our findings suggest that during development, Rtf1 facilitates expression of the cardiac transcription program by supporting the promoter-proximal pausing of RNA Polymerase II.

Results:

Rtf1 is essential for heart development in zebrafish

We previously showed that knockdown of *rtf1* abolishes cardiac progenitor cell formation in zebrafish. To 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, which nearly eliminate the production of Rtf1 protein (Fig. 1a-c). Like *rtf1* morphants and the previously reported *rtf1*^{KT641}allele[24,26], both *rtf1*^{LA2678} and *rtf1*^{LA2679} mutants have poorly defined posterior somite boundaries but are otherwise visually indistinguishable from their wild type siblings during early somitogenesis, the developmental timeframe when cardiac progenitors emerge from the LPM. After one day of development, cardiac tissue is not detectable in *rtf1* mutant embryos by visual inspection or by *myl7* expression (Fig. 2a,b) and mutant embryos display a downwardly curved body shape and an absence of pigmentation at two days post-fertilization (Fig. 2 Supplement 1). In situ hybridization further showed that *rtf1*^{LA2678} and *rtf1*^{LA2679} mutant embryos, like *rtf1* morphants, fail to express normal levels of critical cardiac transcription factors including *mef2ca*, *nkx2.5, tbx5a and txb20* at the time when cardiac progenitors emerge from the ALPM (Fig. 2c-j)[24,27]. Together, these data show that morpholino knockdown faithfully recapitulates *rtf1* mutant phenotype [24] and demonstrate that Rtf1 activity is required for the

130 expression of multiple cardiac transcription factors, indicating a critical role of Rtf1 in regulating
131 the cardiac gene program.
132
133 *Conserved role of Rtf1 in mammalian cardiac development*
134
135 The sequence of the Rtf1 protein is highly conserved among vertebrates (83% identity at the
136 amino acid level between fish and humans). To examine if the crucial role of Rtf1 in
137 cardiogenesis we observed in zebrafish was conserved in mammals, we expressed Cre-
138 recombinase under the control of the Mesp1 promoter in Rtf1 flox mouse embryos (Fig.
139 3a)[28]. The Mesp1 promoter drives expression of Cre in the early cardiogenic mesoderm
140 (E6.25)[29], allowing for knockout of Rtf1 activity prior to the onset of heart development. Rtf1
141 Mesp1-Cre knockout mouse embryos die prior to birth, suggesting that Rtf1 function in the
142 cardiogenic mesoderm is essential for survival. We examined the developing hearts of Rtf1
143 Mesp1-Cre knockouts and their wild type siblings by in situ hybridization and found that while
144 Cre+;Rtf1 flox/+ and Cre- embryos display an *Nkx2.5*+/*Tbx20*+ heart tube undergoing looping at
145 E8.5, Rtf1 Mesp1-Cre knockouts have little or no expression of these markers (Fig. 3b,c).
146 These data suggest a conservation of the critical role of Rtf1 in differentiation of the myocardial
147 lineage from the mesoderm.
148
149 To further investigate the conservation of Rtf1's role in cardiac differentiation, we knocked down
150 Rtf1 activity in mouse embryonic stem cells (mESCs) using short hairpin RNA (Fig. 4a,b).
151 Unmanipulated embryoid bodies (EBs) derived from mESCs are capable of differentiating into
152 cardiac tissues and form beating patches (Fig. 4c,d). In sharp contrast, Rtf1-deficient mESCs
153 display a dramatic reduction in *Myh6*, *Nkx2.5*, and *ANF*/*Nppa* expression and produce
154 significantly fewer beating patches (~67% reduction) (Fig. 4a-d). Interestingly, the temporally

155 dynamic expression levels of *Brachyury* (mesoderm), *Afp* (endoderm) and *Ncam1* (ectoderm)

156 are not significantly affected by Rtf1 knockdown (Fig. 4e-g). Together, these data show that Rtf1
157 is dispensable for the formation of the three germ layers but is essential for differentiation of the
158 cardiac lineage from mesoderm in mammals. This phenotype is similar to what we observed in
159 zebrafish embryos, indicating a conservation of Rtf1's function in cardiogenesis between fish
160 and mammals.

161

162 *Rtf1 is required for expression of genes associated with differentiation of cardiac precursors*
163 *from the lateral plate mesoderm*

164

165 To gain insight into how Rtf1 deficiency impacts differentiation of the cardiac lineage, we profiled
166 LPM transcriptomes of control and *rtf1* morphant embryos. We previously showed that loss of
167 Rtf1 activity does not disrupt expression of the LPM marker *hand2*[24]. We therefore took
168 advantage of a hand2:GFP zebrafish transgenic reporter line to isolate LPM cells from wild type
169 and Rtf1-deficient embryos (Fig. 5a-c and Fig. 5 Supplement 1). Compared to wild type control
170 samples, the *rtf1* morphant LPM displayed 1420 significantly downregulated and 1185
171 significantly upregulated genes (adjusted p-value < 0.05) at the 10-12 somite stage when
172 cardiac progenitor formation is underway in the LPM (Supplementary Data File 1). We
173 examined the most significantly differentially expressed genes and observed that many of the
174 significantly downregulated genes, including *wnt2bb, mef2cb*, *mef2ca*, *isl2b*, *tnnt2a*, *atp2a2a*,
175 *rbfox1l*, and *nkx2.5*, are cardiac-lineage markers or have critical known roles in cardiac
176 development (Fig. 5d). We further found that the most significantly downregulated genes were
177 highly enriched for gene ontologies related to cardiac development, embryonic heart
178 morphogenesis, and muscle cell development (Fig. 5e). These data show that Rtf1 activity is

179 required for expression of a broad cardiac gene program in the LPM that is likely under the
180 control of multiple cardiac transcription factors.
181
182 To further analyze the deployment of the cardiac gene program during cardiac progenitor
183 formation, we performed Single Cell Multiome sequencing of control and *rtf1* morphant
184 embryos. Clustering of control and *rtf1* morphant cells based on gene expression and
185 chromatin accessibility resolved 39 cell types corresponding to expected major cell lineages
186 (Fig. 6 Supplement 1, Fig. 6 Supplement 2, Supplementary Table 2). Interestingly, while *rtf1*
187 morphant cells of each cell type were present, the frequencies of cell types differed significantly
188 between controls and *rtf1* morphants (Fig. 6 Supplement 3). Notably, LPM cells are significantly
189 more abundant in *rtf1* morphants and ALPM derivative cell types (putative cardiac mesoderm,
190 endothelial precursors and rostral blood precursors) are less abundant, indicating that Rtf1
191 deficiency negatively impacts ALPM differentiation (Fig. 6 Supplement 3).
192
193 We next examined developmental trajectories within the ALPM by reclustering ALPM cells
194 (*sema3e*+/*gata5*+ LPM) and ALPM derivatives (Fig. 6a,b). By calculating pseudotime
195 trajectories, we observed a gradual increase in the expression levels of cardiac markers
196 including *ttn.2*, *mef2cb*, *tnnt2a*, *ryr2b*, and *myh7bb* along the ALPM-cardiac developmental
197 trajectory in control embryos (Fig. 6c-e). In *rtf1* morphants, the proportion of cells representing
198 undifferentiated ALPM was increased (37% in *rtf1* morphants vs 15.5% in controls, Fig. 6b and
199 Fig. 6 Supplement 4). Some cells (26.6%) did appear to differentiate along the cardiac
200 trajectory in *rtf1* morphants; like control cardiac progenitors, they maintained the expression of
201 *gata5,* downregulated expression of the ALPM marker *sema3e*, and initiated expression of the
202 early cardiac lineage genes *bmp6*, *ttn.2*, and *mef2cb* (Fig. 6d-g, Fig. 6 Supplement 4, Fig. 6
203 Supplement 5). However, loss of Rtf1 activity prevented the proper expression of cardiac
204 lineage-restricted markers that are normally activated later in the ALPM-cardiac trajectory

205 including *tnnt2a*, *ryr2b*, and *myh7bb* (Fig. 6d-g, Fig. 6 Supplement 4, Fig. 6 Supplement 5).
206 Taken together, these data suggest that Rtf1 activity is required for gene expression changes
207 associated with proper cardiac progenitor differentiation.
208
209 *Rtf1's Plus3 functional domain is necessary for cardiac progenitor formation*
210
211 The Rtf1 protein is a multifunctional platform for transcription regulation: it consists of several
212 defined functional domains that mediate its interaction with other transcription regulatory
213 proteins and serve as additional points of contact for these proteins with the transcription
214 apparatus (Fig. 1a and Fig. 7a)[30,31]. The N-terminal half of Rtf1 notably contains its Histone
215 Modification Domain (HMD), which mediates its interaction with E2 ubiquitin-conjugating
216 enzymes and supports the ubiquitylation of histone H2B K120[32-34]. Rtf1's C-terminus contains
217 the Plus3 domain, which interacts with the C-terminal repeat region of the pausing and
218 elongation factor Spt5[18,19,35], a Pol II-interaction domain that allosterically stimulates
219 transcription elongation[31], and a PAF1C interaction domain.
220
221 We investigated if Rtf1 could function in a modular fashion, and if a specific domain or domains
222 were required for the formation of cardiac progenitors during development. To this end, we
223 generated several constructs to express Flag-tagged wild type and mutant versions of Rtf1
224 lacking specific functional domains and tested whether they could support cardiogenesis in an
225 *rtf1* morphant background. These constructs also carry 8 silent mutations that alter the *rtf1*
226 translation blocking morpholino binding site to ensure successful translation of synthesized
227 mRNA. Wild type (Rtf1 wt*) and mutant FLAG-tagged Rtf1 proteins lacking the HMD (ΔHMD) or
228 Plus3 domain (ΔPlus3) were expressed at comparable levels and properly localized to the
229 nucleus demonstrating that the activity of these proteins was not compromised by instability or

230 mislocalization (Fig. 7b). Injection of Rtf1 wt* mRNA into *rtf1* morphant embryos was able to
231 robustly rescue the knockdown phenotype and support normal formation of the embryonic heart
232 tube (Fig. 7c). Interestingly, we found that Rtf1's HMD was dispensable for its role in cardiac
233 progenitor formation (100% of Rtf1 wt* injected *rtf1* morphants with *nkx2.5* signal, n=25 vs.
234 100% of Rtf1 ΔHMD *rtf1* morphants with *nkx2.5* signal, n=25) (Fig. 7c). On the other hand,
235 Rtf1's Plus3 domain was essential for supporting the formation of cardiac progenitors (0% of
236 Rtf1 ΔPlus3 *rtf1* morphants with *nkx2.5* signal, n=25) (Fig. 7c). These data suggest that Rtf1's
237 domains function in a modular fashion for regulating cardiac gene expression and point to
238 potential specific effects on transcriptional pausing and elongation.
239
240 *Rtf1 deficiency disrupts promoter-proximal transcriptional pausing*
241
242 To examine the occupancy of RNA Pol II, we performed ChIP-seq on embryos at the 10-12
243 somite stage, when cardiac progenitors are arising from the lateral plate mesoderm. We
244 compared the occupancy of Pol II between uninjected control and *rtf1* morphant embryos and
245 found that Pol II was substantially diminished at the transcription start site (TSS) of many genes
246 compared to the remainder of the gene body (Fig. 8 Supplement 1 and data not shown).
247 Pausing of RNA Pol II at the TSS is a normal aspect of transcription in vertebrates and has
248 been hypothesized to regulate both the timing and expression level of genes[36,37]. To quantify
249 transcriptional pausing at each gene, we calculated the Pause Release Ratio (PRR), a ratio of
250 Pol II occupancy in the promoter proximal region of a gene (-300 to +300) to the occupancy in
251 the downstream region (+301 to +1301). A higher PRR indicates increased pausing of Pol II
252 near the transcription start site (TSS), while a lower PRR reflects a gene in a less paused
253 state. We found that on a genome-wide scale, pausing is diminished upon loss of Rtf1 activity
254 during embryonic development (Fig. 8a,b). Out of 6,078 genes with substantial Pol II ChIP
255 signal that were analyzed, 5,799 displayed a significantly decreased PRR (Benjamini–Hochberg

256 false discovery rate of 10%) in Rtf1-deficient embryos compared to 37 genes that displayed a
257 significantly increased PRR (Fig. 8b). The median PRR in control embryos was 3.34, while 95%
258 (5,793/6,078) of genes had a PRR of less than 3.34 in Rtf1-deficient embryos (Figure 8a),
259 reflecting a widespread decrease in promoter proximally paused Pol II. These changes in PRR
260 suggest that Rtf1 is critical for normal levels of pausing at transcriptionally active genes during
261 the developmental window when cardiac progenitor cells arise from the lateral plate mesoderm.
262
263 *Attenuation of pause release permits cardiogenesis in an Rtf1-deficient background*
264
265 The transition between the pausing and elongation phases of transcription is regulated by
266 PTEF-b, a complex consisting of a cyclin (T1, T2a, or T2b) and the kinase Cdk9[38-
267 40]. Phosphorylation of the negative elongation factor (NELF) complex, the pausing/elongation
268 factor Spt5, and the C-terminal tail of RNA Pol II by Cdk9 promotes changes in the factors
269 associated with RNA Pol II and triggers a switch to processive elongation[41]. Pharmacological
270 inhibition of Cdk9 with the small molecule flavopiridol can prevent or attenuate pause release,
271 leading to genome-wide increases in pausing at most genes[42]. To explore the effect of Cdk9
272 inhibition on Rtf1-dependent transcription, we performed ChIP-seq for RNA Pol II in flavopiridol-
273 treated *rtf1* morphant embryos at the 10-12 somite stage. Flavopiridol treatment caused a small
274 but highly significant increase in promoter-proximal pausing in an Rtf1-deficient background
275 (Fig. 8a), with 87.1% of genes with significantly altered PRR values displaying an increase in
276 PRR in flavopiridol-treated *rtf1* morphants compared to untreated *rtf1* morphants (Fig. 8c).
277 Furthermore, we found that critical LPM and cardiac genes including *hand2*, *gata5*, *aplnrb*,
278 *bmp4*, *nkx2.7*, and *rbfox1l* exhibited reduced promoter-proximal pausing in Rtf1-deficient
279 embryos that was partially rescued by treatment with flavopiridol (Fig. 8 Supplement 1).
280

281 If diminished promoter-proximal pausing at genes crucial for cardiogenesis is responsible for the
282 absence of cardiac progenitor formation in Rtf1-deficient embryos, then we would predict that
283 reducing pause release may restore cardiogenesis in an Rtf1-deficient background. Indeed,
284 flavopiridol treatment (5 µg/mL dosage) potently restored cardiomyocyte formation in *rtf1* mutant
285 embryos as indicated by *myl7* expression (Fig. 8d,e). Similarly, knockdown of Cdk9 by injection
286 of antisense morpholino oligonucleotides[43,44] also restored the formation of *myl7*+
287 cardiomyocytes in an Rtf1-deficient background (Fig. 8f,g), confirming that decreased activity of
288 the Cdk9 kinase is responsible for rescuing the *rtf1* cardiac phenotype. These dramatic results
289 suggest that the primary role of Rtf1 in cardiac progenitor formation is regulating promoter-
290 proximal pausing of critical members of the cardiac gene program.
291
292 We next performed RNA-seq to examine what transcriptional changes might account for the
293 rescue of *myl7*+ cardiomyocyte formation in flavopiridol-treated Rtf1-deficient embryos.
294 Strikingly, we found that of the 4,102 genes that were significantly downregulated in *rtf1*
295 morphants at the 8-9 somite stage compared to stage-matched uninjected controls, the
296 expression of 1,020 genes was significantly rescued by treatment with flavopiridol (24.9%; Fig.
297 8h). Interestingly, we also observed a modest rescue effect on those genes that were
298 significantly downregulated in *rtf1* morphants (427 of 4,122 genes significantly rescued: 10.4%;
299 Fig. 8h). Those genes rescued by flavopiridol treatment were relevant to embryonic
300 development based on significant enrichment for gene ontologies including pattern specification,
301 embryonic organ development, embryonic organ morphogenesis, and cell fate commitment (Fig.
302 8i). We examined the expression levels of individual genes critical to cardiac development and
303 found that *nkx2.5*, *tbx5a*, *tbx20*, *mef2ca*, *mef2cb*, *ttn.2*, and *ryr2b* expression levels were all
304 significantly rescued by treatment of Rtf1-deficient embryos with flavopiridol (Fig. 8j).
305 Altogether, our data support a model in which loss of Rtf1 causes excessive pause release, a

306 dysregulation of the cardiac transcription program, and a failure in the formation of cardiac
307 progenitors.
308
309 **Discussion**:
310
311 In this study we showed that the PAF1C member Rtf1 plays a critical role in cardiogenesis in
312 zebrafish and mammals. Ablation of Rtf1 activity impedes mesoderm differentiation along the
313 myocardial trajectory resulting in a failure to produce cardiac progenitors with proper expression
314 of cardiac lineage-restricted genes such as *nkx2.5* and *tbx5a*, and the subsequent loss of the
315 cardiomyocyte population. Structure-function analysis revealed that Rtf1's cardiogenic activity
316 requires its Plus3 domain, a site of interaction with the pausing/elongation factor Spt5, and that
317 loss of Rtf1 diminishes promoter-proximal pausing of RNA Polymerase II. These observations,
318 together with the finding that partial inhibition of transcriptional pause release can support
319 cardiogenesis and restore cardiomyocyte formation in an Rtf1-deficient background,
320 demonstrate a crucial role for Rtf1-mediated transcriptional pausing in early heart development.
321 These findings also raise an intriguing possibility of potentiating cardiac differentiation through
322 manipulating transcriptional pausing.
323
324 How exactly pausing regulates gene expression is not fully understood and likely involves both
325 general and gene/locus-specific mechanisms[36,45]. By limiting the release of RNA Polymerase II
326 into the elongation phase of transcription, pausing can restrain gene expression. This inhibitory
327 role of pausing on transcription allows for another layer of transcription regulation and has been
328 shown to modulate the expression of heat shock and immune stimulus-induced genes[46-48].
329 Pausing-dependent restraint of gene expression is also important for the timing of gene
330 expression in a developmental context. Supporting this notion, Spt5 and NELF-deficient
331 zebrafish embryos exhibit aberrant upregulation of TGFβ signaling and a resulting inhibition of

332 hematopoiesis[44]. Intriguingly, the PAF1C has also been shown in mammalian cells to restrain
333 transcription from a subset of promoters by inhibiting nearby enhancer activation[49]. On the other
334 hand, pausing also positively regulates gene expression in some contexts. At genes where
335 nucleosome formation is favored in promoter regions, pausing of RNA Polymerase II can
336 promote a chromatin architecture that is permissive to transcription by competing with
337 nucleosomes for promoter occupancy[50,51]. This mechanism has been suggested to support
338 expression of the IFN-gamma pathway during zebrafish development[44]. Pausing may also
339 positively regulate gene expression during neural crest development. Loss of Paf1 activity
340 disrupts neural crest differentiation, but this defect can be partially rescued by inactivation of
341 Cdk9. Interestingly, zygotic loss of Cdk9 alone is sufficient to expand the neural crest population
342 in zebrafish embryos, suggesting that promoter-proximal pausing promotes the neural crest
343 gene program[52].
344
345 Experiments performed using a rapid protein degradation system in human cell culture has shed
346 light on the mechanisms by which the PAF1C regulates transcriptional pausing and gene
347 expression outputs[53]. In DLD-1 cells, acute depletion of PAF1 destabilizes the PAF1C and
348 stimulates pause release. Loss of PAF1 also decreases chromatin recruitment of INTS11, a
349 component of the Integrator-PP2A complex (INTAC) which interacts with PAF1C and
350 counteracts the activity of P-TEFb. This disruption in the balance of INTAC and P-TEFb
351 activities at transcribed genes results in hyperphosphorylation of SPT5 and likely explains the
352 increase in pause release caused by PAF1 depletion. Similar to our finding that Rtf1-deficiency
353 positively and negatively regulates gene expression, depletion of PAF1 resulted in both
354 upregulation and downregulation of many genes. Intriguingly, the gene expression changes
355 caused by PAF1 depletion correlated with the release frequency and processivity of RNA
356 Polymerase II at a given gene, with genes exhibiting low levels of release frequency tending to
357 display decreases in their expression level. These observations suggest that the combined

358 regulatory functions of pausing and elongation determine the transcriptional output of PAF1C-
359 regulated genes. Future studies focused on the chromatin architecture of cardiac gene
360 promoters and other regulatory regions, and the dynamics of pause release and elongation at
361 cardiac genes are needed to provide a complete picture of the manner by which promoter-
362 proximal pausing fosters their expression.
363
364 **Materials and Methods**:
365
366 *Zebrafish*
367
368 Zebrafish colonies were cared for and bred under standard conditions[54]. Developmental stages
369 of zebrafish embryos were determined based on somite number or hours of development post
370 fertilization at 28.5 °C[54]. Fish husbandry and experiments were performed according to the
371 Institutional Approval for Appropriate Care and Use of Laboratory Animals by the UCLA
372 Institutional Animal Care and Use Committee (Protocols ARC-2000-051-TR-001 and BUA-2018-
373 195-002-CR).
374
375 *CRISPR mutagenesis*
376
377 Template DNA for *rtf1* exon 3 guideRNA (rtf1-e3 gRNA) synthesis was amplified with KOD
378 polymerase (MilliporeSigma) using the pXTW-gRNA plasmid and the primers rtf1-e3-F (5'-
379 TAATACGACTCACTATAGGAAGAAGGGGAAACCGAGCAAGTTTTAGAGCTAGAAATAGC-3')
380 and gRNA-R (5'-AAAAAAAGCACCGACTCGGTGCCAC-3'). rtf1-e3 gRNA (target sequence: 5'-
381 GGAAGAAGGGGAAACCGAGCAA-3') was synthesized using the MEGAshortscript T7
382 Transcription Kit (ThermoFisher Scientific) and purified using the RNeasy Mini Kit (Qiagen).
383 Zebrafish embryos were injected at the 1-cell stage with 300 pg of Cas9 mRNA and 200 pg of

384 rtf1-e3 gRNA. Injected embryos were raised to adulthood and then outcrossed to wild type AB
385 fish to establish lines carrying specific *rtf1* mutations.
386
387 *Mice*
388
389 All mice were maintained in the C57BL/6 background according to the Guide for the Care and
390 Use of Laboratory Animals published by the US National Institute of Health. Housing and
391 experiments were performed according to the Institutional Approval for Appropriate Care and
392 Use of Laboratory Animals by the UCLA Institutional Animal Care and Use Committee
393 (Protocols ARC-2020-045-TR-001 and ARC-2020-043-TR-001). ES cells were obtained from
394 KOMP and used to rederive the Rtf1$^{tm1a(KOMP)Wtsi}$ mouse line. This line was then crossed to a
395 ROSA26::FLPe strain[55] (The Jackson Laboratory, JAX stock #009086) to create an Rtf1
396 conditional knockout allele where exon 3 of Rtf1 is flanked by two loxP sites (Rtf1 flox). These
397 mice were crossed into the Mesp1$^{tm2(cre)Ysa}$ background[29] to produce a strain capable of
398 generating Mesp-Cre^{+};Rtf1 flox/flox offspring.
399
400 *Whole mount in situ hybridization*
401
402 Whole-mount in situ hybridization was performed as described previously[56]. Embryos for in situ
403 hybridization were fixed in 4% paraformaldehyde in 1xPBS. After a brief proteinase K digestion,
404 embryos were incubated with digoxigenin-labeled antisense RNA probes overnight at 65-70°C.
405 Probes were detected using an anti-digoxigenin antibody conjugated with alkaline phosphatase
406 (Roche 11093274910). The zebrafish antisense RNA probes used in this study include *myl7*,
407 *nkx2.5*, *mef2ca*, *tbx20*, and *tbx5a*. The mouse probes used include *Tbx20* and *Nkx2.5*.
408
409 Genotypes of zebrafish embryos were confirmed by PCR with primers:

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

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

412

413 Genotypes of mouse embryos were confirmed by PCR with primers:

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

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

416 Cre-F: 5'-GTAGCTGATGATCCGAATAACTAC-3'

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

418

419 *Constructs*

420

421 To prevent complementarity between *rtf1* morpholino and mRNA sequences, PCR was used to

422 introduce 8 silent mutations to the morpholino target sequence using pCS2/3XFLAG-rtf1 as a

423 template[24]. The morpholino-resistant wild type *rtf1* sequence was cloned into pCS2/3XFLAG to

424 produce an N-terminally FLAG-tagged morpholino-resistant rtf1 construct (pCS2/3XFLAG-rtf1-

425 mr). Rtf1 deletion mutant sequences were amplified from pCS2/3XFLAG-rtf1-mr by SOE-PCR

426 to produce Rtf1ΔHMD (missing a.a. 126-211 of Rtf1) and Rtf1ΔPlus3 (missing a.a. 319-450 of

427 Rtf1), which were subsequently cloned into pCS2/3XFLAG to produce N-terminally tagged

428 constructs.

429

430 *Microinjections*

431

432 0.5 ng of *rtf1* morpholino (rtf1MO; 5'-CTTTCCGTTTCTTTACATTCACCAT-3') was injected at

433 the 1-cell stage[24,26] along with 1 ng of *p53* morpholino (5'-GCGCCATTGCTTTGCAAGAATTG-

434 3') to prevent p53-dependent apoptosis[57]. For *cdk9* knockdown rescue experiments, 5 ng of

435 *cdk9* morpholino (cdk9MO; 5'-ACACACAAACATCAAATACTCACCC-3')[43,44] was injected into an

436 incross of *rtf1*^{LA2679} heterozygotes at the 1-cell stage. 150-200 pg of *rtf1* or *rtf1* deletion mutant
437 mRNAs were co-injected with *rtf1* and *p53* morpholinos at the 1-cell stage for mRNA rescue
438 experiments.
439
440 *Mouse ESC culture and differentiation*
441
442 E14TG2a mESCs (ATCC) were cultured on 0.1% gelatin covered plates without feeder cells.
443 The feeder-free mESCs were maintained in Dulbecco's modified Eagle's medium (Invitrogen)
444 supplemented with 15% fetal bovine serum (FBS; Invitrogen, USA), 1% non-essential amino
445 acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 1% (v/v) penicillin-streptomycin-
446 Glutamine (ThermoFisher Scientific), and 10 μg/μL recombinant mouse leukemia inhibitory
447 factor (LIF; Sigma-Aldrich) in a 5% CO_2 incubator at 37°C. The ESC culture medium was
448 changed every two days. mESCs were subcultured every 3 days using 0.25% trypsin-EDTA
449 (Invitrogen). mESCs were differentiated in hanging drops (600 cells/20 μL) without LIF under
450 non-adherent conditions. Day 5 embryoid bodies (EBs) were plated onto 0.1% gelatin covered
451 plates. Media was changed every two days during adherent culture.
452
453 *Lentivirus transduction*
454
455 To knockdown Rtf1 activity in mESCs, Rtf1 or NT (non-target) shRNA lentivirus harboring
456 puromycin resistance were purchased from Sigma-Aldrich. mESCs were transduced with the
457 indicated virus at a multiplicity of infection of 1 and selected in 1.2 μg/mL puromycin. Rtf1
458 shRNA knockdown efficacy was tested by Western blotting analysis that detected Rtf1 protein
459 level. Antibodies used were anti-Rtf1 (Bethyl Labs A301-329A) and anti-β-actin (Sigma A1978).
460 Relative protein levels were determined by densitometry measurements using Adobe
461 Photoshop.

462

463 *Quantitative RT-PCR*

464

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

469

470 *Antibody staining*

471

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

478

479 *Western Blotting*

480

481 Embryos were lysed in Rubinfeld's Lysis Buffer at 1 day post fertilization as previously
482 described[24]. Two embryo equivalents were loaded per lane of an 8% polyacrylamide gel.
483 Following electrophoresis, proteins were transferred to a nitrocellulose membrane and detected
484 with antibodies against Rtf1 (Bethyl Labs A301-329A, 1:4000) and beta-actin (Sigma A1978,
485 1:5000). ImageJ was used to perform densitometry comparing the levels of Rtf1 protein
486 between samples using beta-actin as the control.

487

488 *Chemical treatment*

489

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

494

495 *FACS*

496

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

506

507 *mRNA-seq*

508

509 RNA was purified from FACS-sorted cells with the NucleoSpin RNA kit (Machery Nagel). Total
510 RNA was extracted from whole embryos using TRIzol RNA isolation reagents (ThermoFisher
511 Scientific) and then purified with the NucleoSpin RNA kit (Machery Nagel). Approximately 5-10
512 ng of input RNA was used to produce libraries with the NEBNext Single Cell/Low Input RNA
513 Library Prep Kit for Illumina (NEB) which were sequenced on an Illumina HiSeq 3000 to produce

514 50 bp single-end reads. Reads from triplicate samples were mapped to the danRer7 genome
515 using Tophat v2.1.1[59] with default parameters. Mapped reads were counted with
516 FeatureCounts v2.0.0[60] and differential expression analysis was carried out with DESeq2
517 v1.32.0[61]. Gene ontology enrichment was examined using the R package clusterProfiler[62] in R
518 v4.1.0. For gene ontology analysis in LPM FACS experiments, the most significantly
519 downregulated genes in *rtf1* morphant LPM were selected as those genes that were
520 downregulated among the 300 most significantly differentially expressed genes based on
521 adjusted *p*-value. For gene expression analysis of flavopiridol rescue experiments, genes were
522 considered significantly rescued if flavopiridol treatment resulted in a significant expression
523 change in the direction opposite of that caused by *rtf1* knockdown (adjusted *p*-value < 0.1).
524
525 *ChIP-seq*
526
527 Embryos for ChIP were dissociated with 4 mg/mL Collagenase Type IV and 0.25% Trypsin at
528 room temperature with intermittent pipetting. Dissociation was stopped with 10% FBS in DMEM
529 and cells were washed once with PBS. Fixation and preparation of chromatin was carried out
530 using the Covaris truChIP Chromatin Shearing Kit and 7 minutes of sonication in a Covaris
531 E220 ultrasonicator using the recommended settings. Immunoprecipitations were carried out as
532 previously described[63] on approximately 5 µg of chromatin (minimum 2.5 µg, maximum 8.8 µg)
533 using 5 µL of total RNA Pol II antibody (Cell Signaling D8L4Y). Libraries were prepared using
534 the NEBNext Ultra II DNA Library Kit for Illumina (NEB) and sequenced with an Illumina
535 Novaseq 6000 system to produce 50 bp paired-end reads. Reads from triplicate samples were
536 mapped to the danRer11 genome using bowtie2 v2.4.2[64] (options: --local -X 1000) and read
537 densities of UCSC RefSeq gene regions were quantified using the package csaw v1.24.3[65,66] in
538 R v4.0.3. For our pausing analyses, we defined the transcription start site (TSS) region as -300
539 bp to +300 bp surrounding the TSS. RNA Pol II ChIP-seq read densities for the TSS regions of

540 genes displayed a bimodal distribution which was assumed to be two partially overlapping
541 distributions of genes with high and low RNA Pol II TSS signal. We examined those genes with
542 a substantial level of RNA Pol II signal in the TSS region of control samples, defined as those
543 genes present in the upper mode of the distribution of log-transformed TSS read densities
544 based on a cutoff for type-I error of 10% using the R cutoff package v0.1.0
545 (https://github.com/choisy/cutoff). The read density in the TSS region was divided by the read
546 density in the region extending from +301 to +1301 bp with respect to the TSS to calculate the
547 pause release ratio (PRR). ChIP-seq data was visualized in Python 3.10.0 using the plotting
548 tool SparK v2.6.2[67].
549
550 *Nuclei isolation for single cell sequencing*
551
552 Nuclei were isolated from 200 11-12 somite stage zebrafish embryos for 10x Genomics
553 Multiome sequencing using recommended procedures with slight modifications. Embryos were
554 deyolked in Deyolking buffer (55 mM NaCl, 1.8 mM KCl, 1.25 mM $NaHCO_3$) and then embryonic
555 cells were partially dissociated by pipetting in DMEM/F12 media (Gibco 21041025). Cells were
556 pelleted and then lysed on ice for 5 minutes in 100 µL of chilled 0.1X Lysis buffer (10 mM Tris
557 pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.01% Tween-20, 0.01% IGEPAL CA-630, 0.001% digitonin,
558 1% BSA, 1 mM DTT). 1 mL of Wash buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$,
559 0.1% Tween-20, 1% BSA, 1 mM DTT, 0.1 U/µL Sigma Protector RNase Inhibitor) was then
560 added to the cell lysis and the nuclei were pelleted by centrifuging at 500xg for 5 minutes at
561 4°C. Two more washes with 1 mL of Wash buffer were performed. Nuclei were then
562 resuspended in 50 µL of Nuclei Dilution Buffer (1X 10x Genomics Single Cell Multiome Nuclei
563 Buffer, 1 mM DTT, 1 U/µL Sigma Protector RNase Inhibitor) and filtered with a 40 µm Flowmi
564 cell strainer. Approximately 10,000 nuclei per sample were targeted for capture with the 10x
565 Chromium Single Cell Multiome ATAC + Gene Expression system by the UCLA Technology

566 Center for Genomics & Bioinformatics.

567

568 *Single cell sequencing analysis*

569

570 10x Chromium Single Cell Multiome ATAC + Gene Expression libraries were prepared by the
571 UCLA Technology Center for Genomics & Bioinformatics following the manufacturer's
572 recommended protocols and sequenced with an Illumina Novaseq 6000 system. Reads from
573 RNA-seq and ATAC-seq assays were aligned to the zebrafish GRCz11 genome with the
574 Lawson laboratory's improved transcriptome annotation[68] using 10x Genomics' Cell Ranger
575 ARC software v2.0.0 with default parameters. Mapped reads from single cell RNA-seq and
576 ATAC-seq assays were analyzed in R v4.2.2 using Seurat v4.3.0[69]. Good quality cells were
577 retained by filtering using Seurat with the following thresholds: 15000 > RNA UMI > 2000,
578 Genes > 1500, 20000 > ATAC UMI > 5000, TSS enrichment score > 3.5. Filtered control and
579 *rtf1* morphant cell datasets were integrated using Harmony v0.1.1[70], and preliminary
580 dimensional reduction, cell clustering, and marker gene identification were carried out in Seurat.
581 Clusters of cells that exhibited enriched expression of mitochondrial and ribosomal transcript
582 expression levels but lacked enrichment for other genes were considered abnormal cells and
583 were manually removed from the dataset. Following manual filtering, a final dataset of 14340
584 control and 13839 *rtf1* morphant cells was again integrated with Harmony and multimodal
585 dimensional reduction (weighted-nearest neighbor method with RNA and ATAC assays), cell
586 clustering, and marker gene identification were carried out in Seurat. Cluster cell number
587 proportions and gene expression levels were visualized with Seurat and with published code[71].
588 Based on marker gene expression, cells that were members of clusters corresponding to
589 anterior lateral plate mesoderm (*sema3e*+), endothelium (*flt4*+), cardiac precursors (*mef2cb*+),
590 and rostral blood (*spi1b*+) were extracted, independently integrated with Harmony, and
591 dimensional reduction (RNA assay only), cell clustering, and marker gene identification were

592 carried out in Seurat. Pseudotime analysis of Harmony/Seurat cell embeddings and plotting of
593 gene expression over pseudotime was performed with Monocle v3.0[72].
594
595 *Statistical analysis*
596
597 Statistical significance for qualitative differences in gene expression in in situ hybridization
598 experiments was determined using Fisher's Exact Test, with a *p*-value of less than 0.05
599 considered to be significant. Statistical significance of gene expression changes in qPCR
600 analyses was determined based on 2 (for *Nkx2-5* and *Afp* expression) or 3 (for *Myh6*, *Nppa*,
601 *Brachyury*, and *Ncam1* expression) independent experiments using a two-tailed Student's *t*-test
602 with a *p*-value of less than 0.05 considered to be significant. Statistical significance of the
603 difference between beating cardiomyocyte cluster formation in control and Rtf1 shRNA mESCs
604 was determined based on 4 independent experiments using a two-tailed Student's *t*-test with a
605 *p*-value of less than 0.05 considered to be significant. To compare PRR means between
606 uninjected, *rtf1* morphant, and flavopiridol-treated *rtf1* morphant embryos, mean PRR values
607 were first log-transformed (log10) and then a Welch's paired two-tailed *t*-test was carried out
608 with a *p*-value of less than 0.05 considered to be significant. Significantly changed PRRs for
609 individual genes were determined based on 3 independent replicates per condition using the
610 rowttests function in the R package genefilter v1.72.1 to perform two-tailed *t*-tests. The p.adjust
611 function in the R stats package v4.0.3 using a Benjamini and Hochberg (false discovery rate)
612 correction for multiple comparisons was then applied with adjusted *p*-values of less than 0.1
613 considered to be significant.
614
615 **Acknowledgements**
616

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).

Figures:

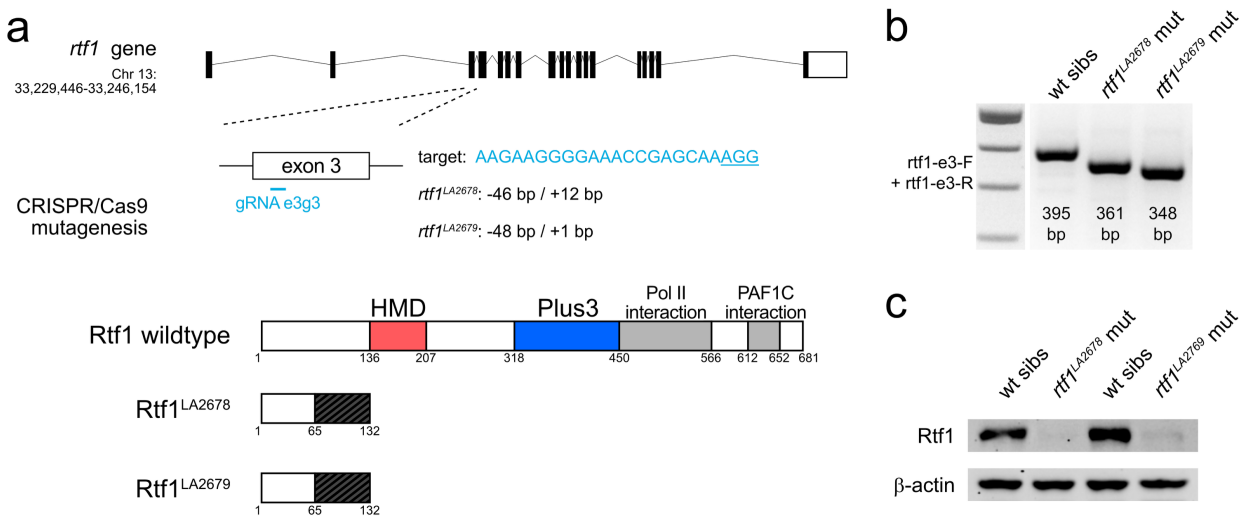


Figure 1. Generation of zebrafish *rtf1* null mutants with CRISPR/Cas9. (a) Schematic of CRISPR/Cas9 mutagenesis of the *rtf1* gene. Two mutant alleles, *rtf1*^{LA2678} and *rtf1*^{LA2679}, were recovered after targeting of *rtf1* exon 3. Both alleles are predicted to disrupt translation of the Rtf1 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.

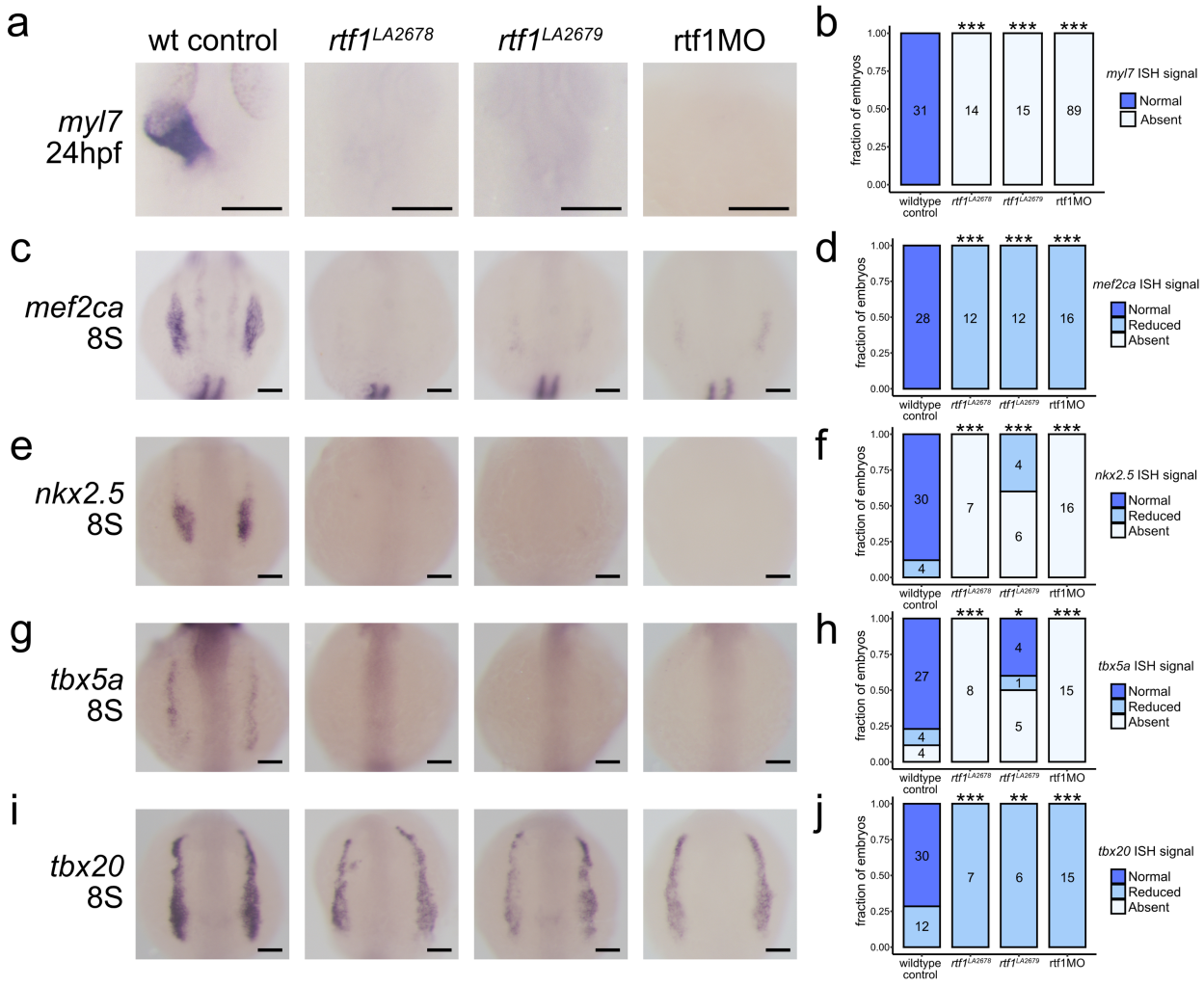


Figure 2. Characterization of early cardiac development in Rtf1-deficient embryos. (a) Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf zebrafish embryos. (b) Quantification of *myl7* signal intensity in control and Rtf1-deficient embryos at 24 hpf. Numbers on 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. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ based on Fisher's Exact Test.

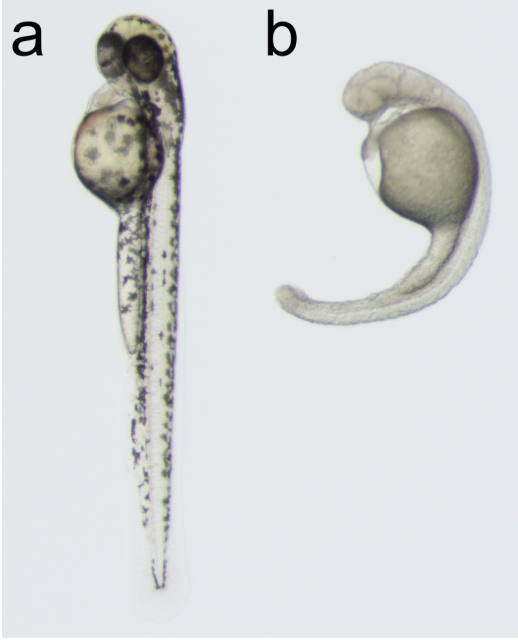


Figure 2 Supplement 1. Bright-field microscopic images of two days post-fertilization (d2) wild type (a) and *rtf1*^{LA2679} mutant (b) embryos.

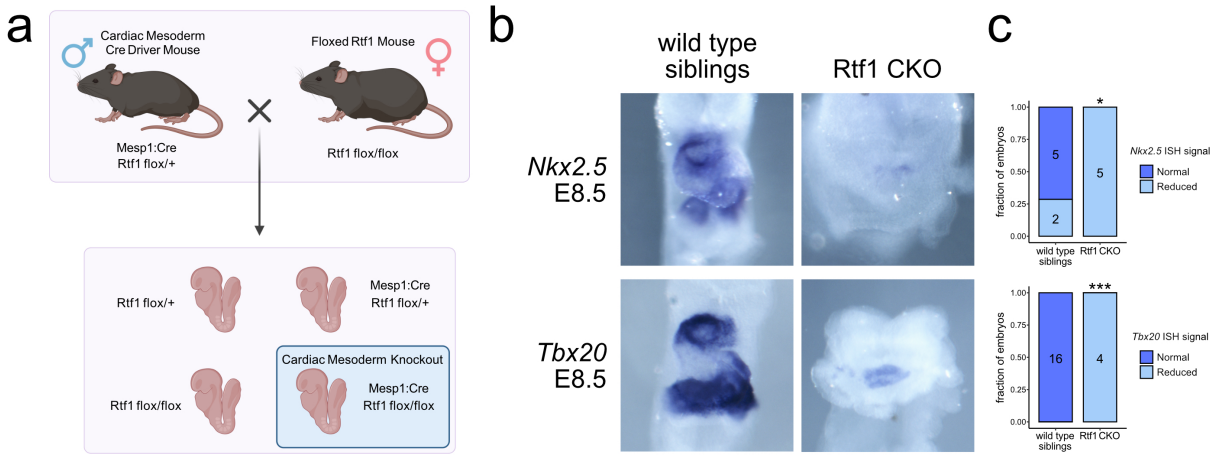


Figure 3. Failure of cardiac differentiation upon loss of Rtf1 in the mammalian cardiac mesoderm. (a) Diagram of generation of cardiac mesoderm-specific Rtf1 knockout mouse embryos. A Mesp1:Cre insertion allele was bred into an Rtf1 exon 3 floxed background. Heterozygous Rtf1 floxed male mice with Mesp1:Cre 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) and *Tbx20* (wild type n = 16, Rtf1 CKO n = 4) gene expression in Cre-negative (wild type siblings) and Cre-positive homozygous Rtf1 floxed (Rtf1 CKO) embryos at E8.5. (c) Quantification of *Nkx2.5* and *Tbx20* signal intensity in Cre-negative (wild type siblings) and Cre-positive homozygous Rtf1 floxed (Rtf1 CKO) embryos at E8.5. Numbers on bars indicate the number of embryos analyzed. *: $p < 0.05$, ***: $p < 0.001$ based on Fisher's Exact Test.

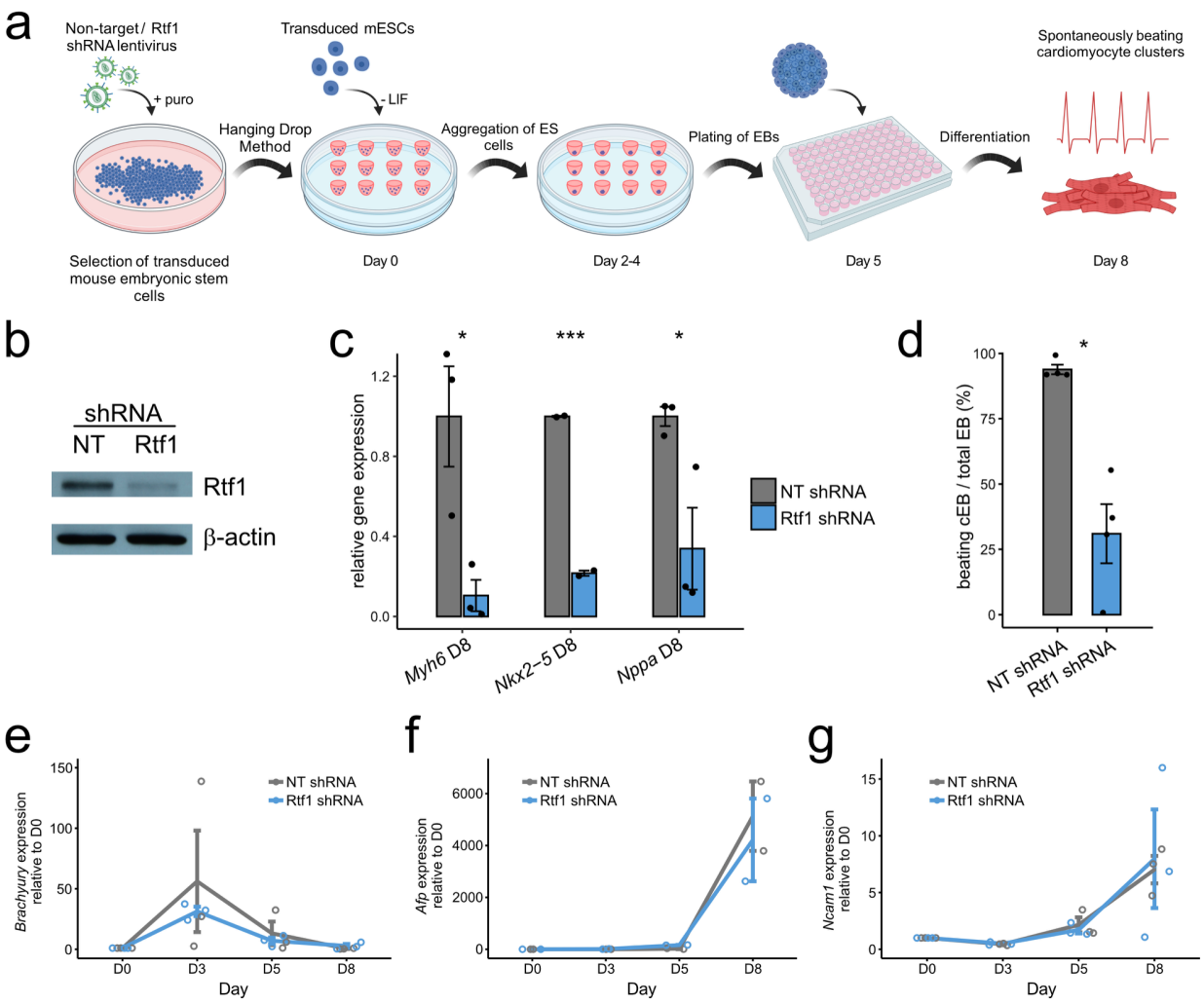


Figure 4. Failure of cardiac differentiation from Rtf1 knockdown mouse embryonic stem cells. (a) Diagram of shRNA knockdown of Rtf1 in mouse embryonic stem cells (mESCs) and differentiation of plated EBs. mESCs were transduced with non-target control or Rtf1 shRNA lentivirus and selected with puromycin. 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

743 shRNA samples and error bars indicate the standard error of the mean. (d) Mean percentage (±
744 standard error) of plated EBs exhibiting beating cardiomyocyte differentiation (cEBs) in NT and
745 Rtf1 shRNA plated EBs. (e) qPCR analysis of *Brachyury* expression in non-target control (NT)
746 and Rtf1 shRNA plated EBs. (f) qPCR analysis of *Afp* expression in non-target control (NT) and
747 Rtf1 shRNA plated EBs. (g) qPCR analysis of *Ncam1* expression in non-target control (NT) and
748 Rtf1 shRNA plated EBs. Data in e, f, and g are normalized to the mean expression level in NT
749 shRNA samples at day 0 (D0) and error bars indicate the standard error of the mean. *: $p <$
750 0.05, ***: $p < 0.001$.
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768

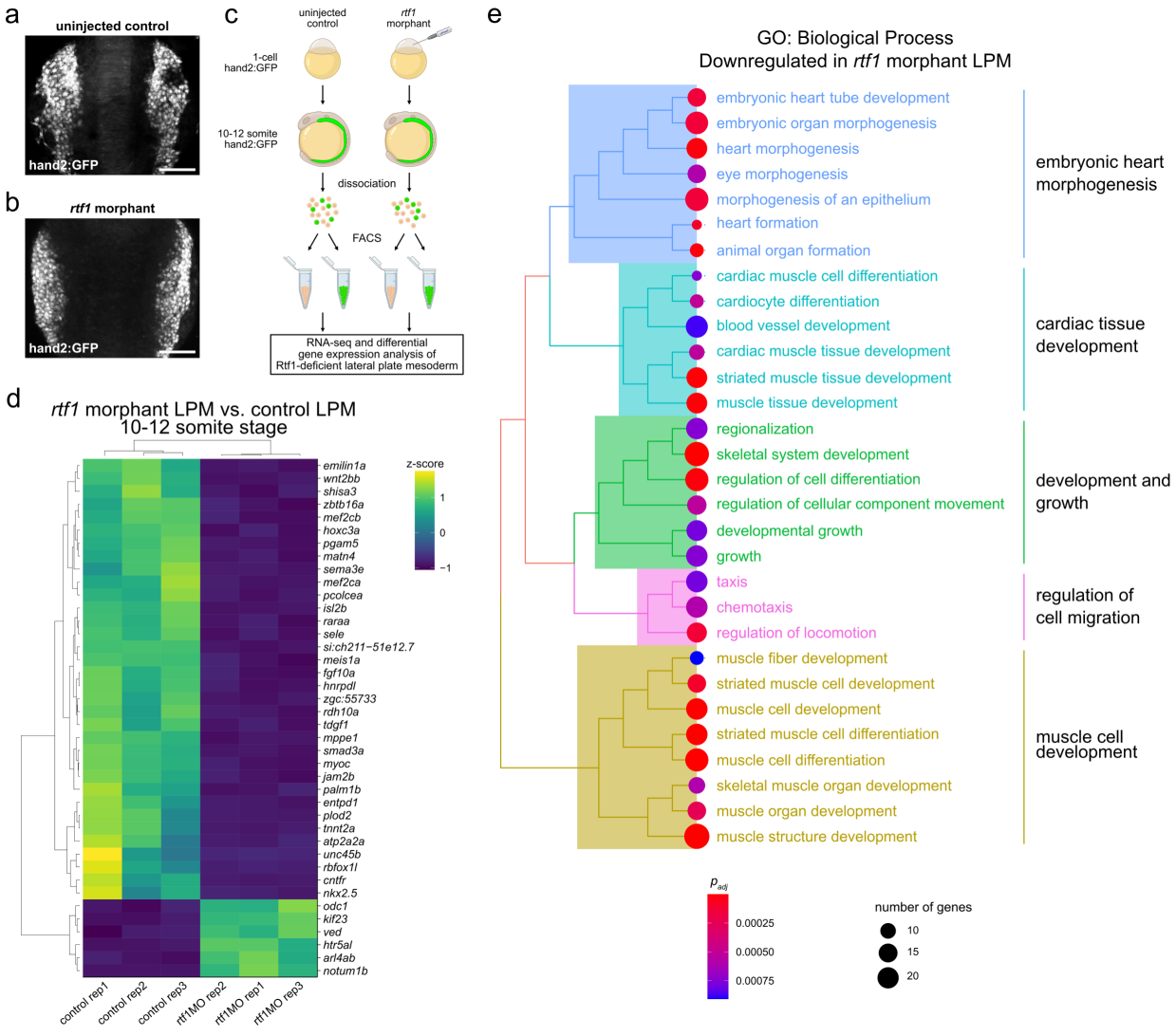


Figure 5. Dysregulation of cardiac gene expression in the Rtf1-deficient lateral plate mesoderm. (a,b) Projections of confocal z-stacks of hand2:GFP signal in 11 somite stage uninjected control (a) and *rtf1* morphant (b) embryos. Scale bars represent 100 μm. (c) Diagram of lateral plate mesoderm (LPM)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

779 hand2:GFP-positive LPM. High and low z-scores represent high and low gene expression,
780 respectively. (e) Hierarchical clustering of Biological Process gene ontology (GO) terms based
781 on semantic similarity for the genes most significantly downregulated in the *rtf1* morphant
782 hand2:GFP-positive LPM.
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804

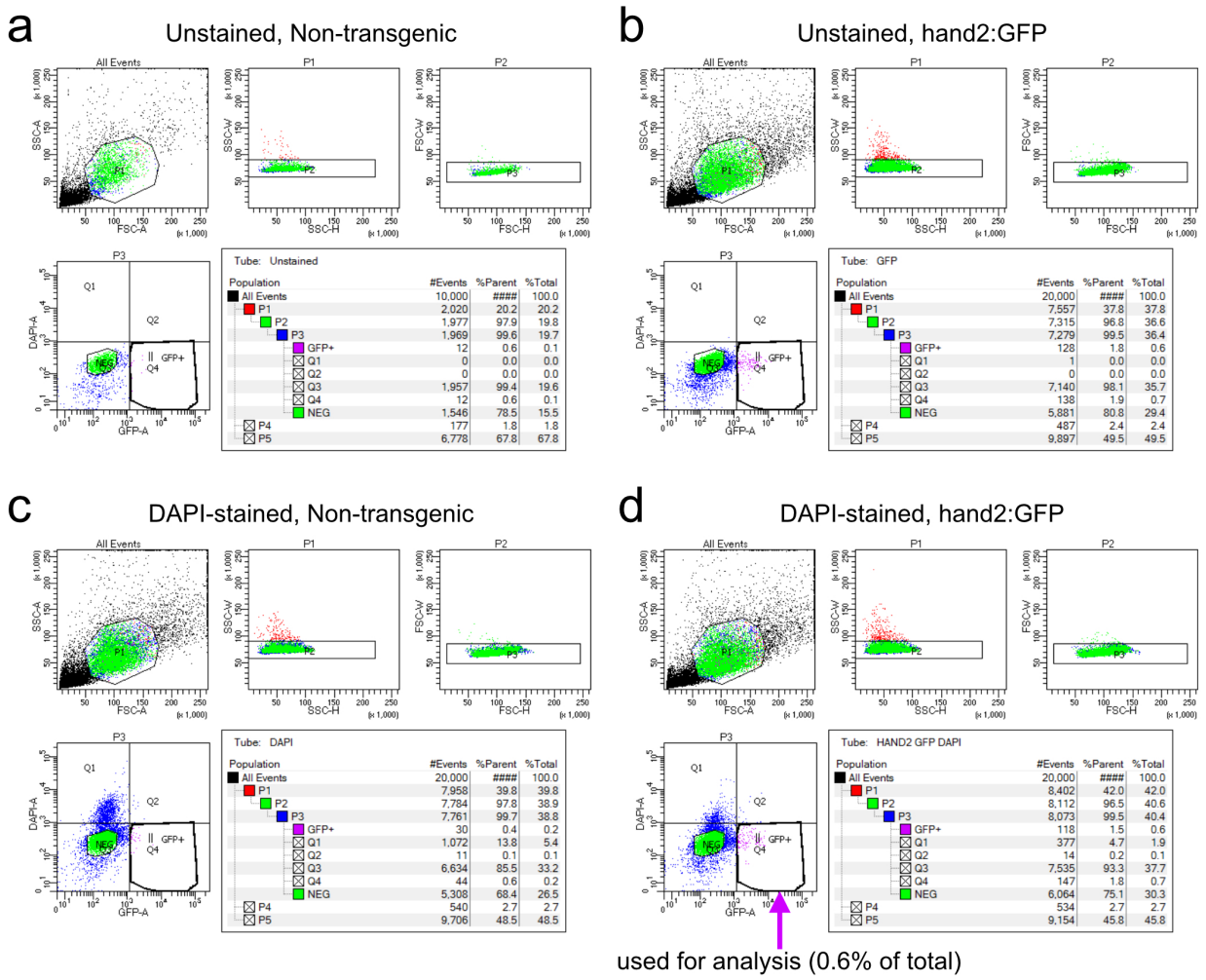


Figure 5 Supplement 1. Gating strategy for 10-12 somite stage zebrafish hand2:GFP FACS. Abnormal cells or multiplets were excluded by gating out events with high side scatter (SSC) and forward scatter (FSC). Live hand2:GFP-positive cells (pink cells in panel d) were sorted based on high GFP and low DAPI signal. (a) Gating control experiment using unstained cells from non-transgenic embryos. (b) Gating control experiment using unstained cells from hand2:GFP transgenic embryos. (c) Gating control experiment using DAPI-stained cells from non-transgenic embryos. (d) Gating of sorting sample for analysis using DAPI-stained cells from hand2:GFP transgenic embryos. 0.6% of the total cells were sorted as live GFP-positive cells. Plots are representative of 4 independent experiments.

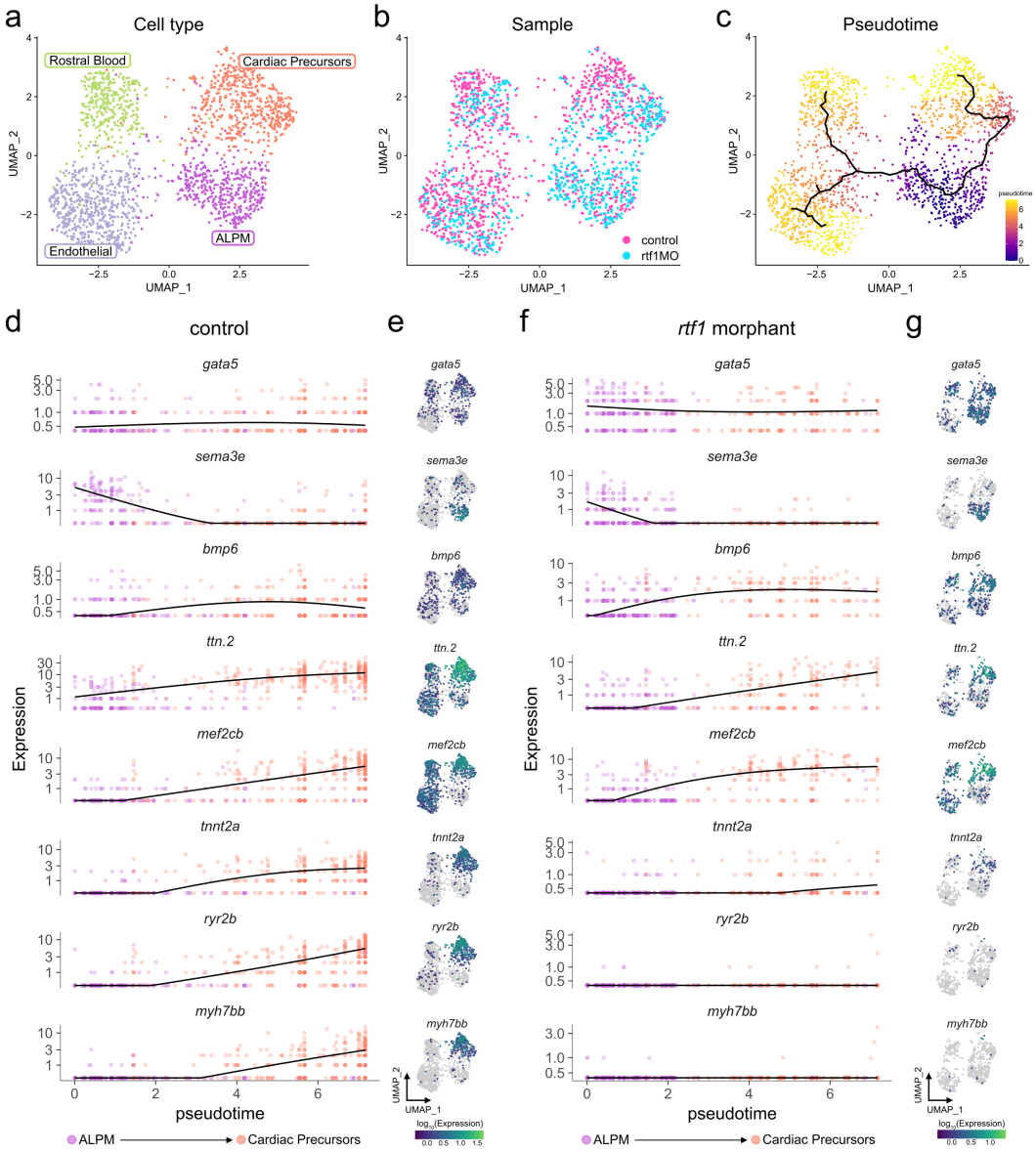


Figure 6. Rtf1 is required for transcriptional progression of cardiac precursors. (a-c) UMAP plots of anterior lateral plate mesoderm (ALPM) and derivatives from merged and integrated single cell RNA-seq datasets of 11-12 somite stage uninjected control and *rtf1* morphant embryos. Cells in plots are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

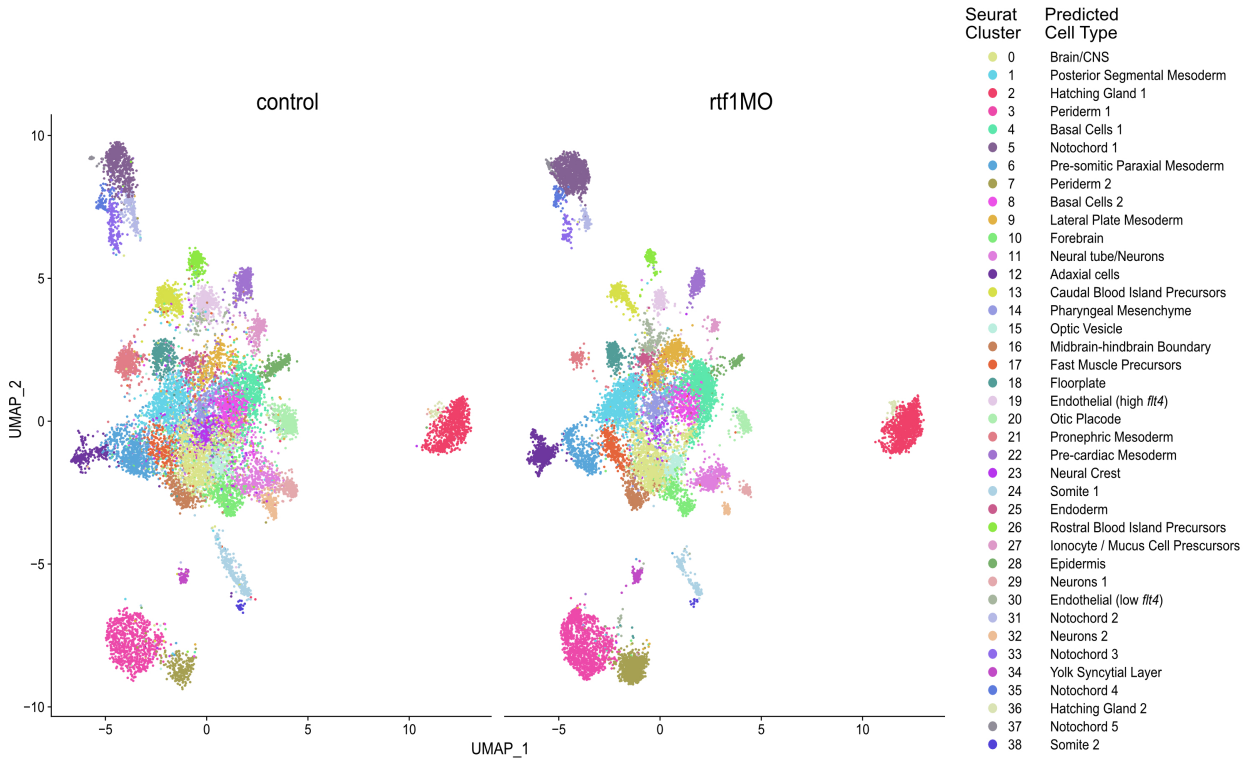


Figure 6 Supplement 1. UMAP plots of 11-12 somite stage control and *rtf1* morphant embryo integrated multimodal (ATAC + Gene Expression) single cell sequencing datasets split based on sample (control on left vs. *rtf1* morphant on right). Cells are colored based on Seurat cluster. Predicted identities of each cell type are based on manual inspection of marker gene expression.

Figure 6 Supplement 2. Clustered dot plot displaying expression of marker genes in cells belonging to each Seurat cluster identified in the integrated multimodal single cell sequencing analysis of 11-12 somite stage control and *rtf1* morphant embryos. Each cluster is further split by sample (control vs. *rtf1* morphant).

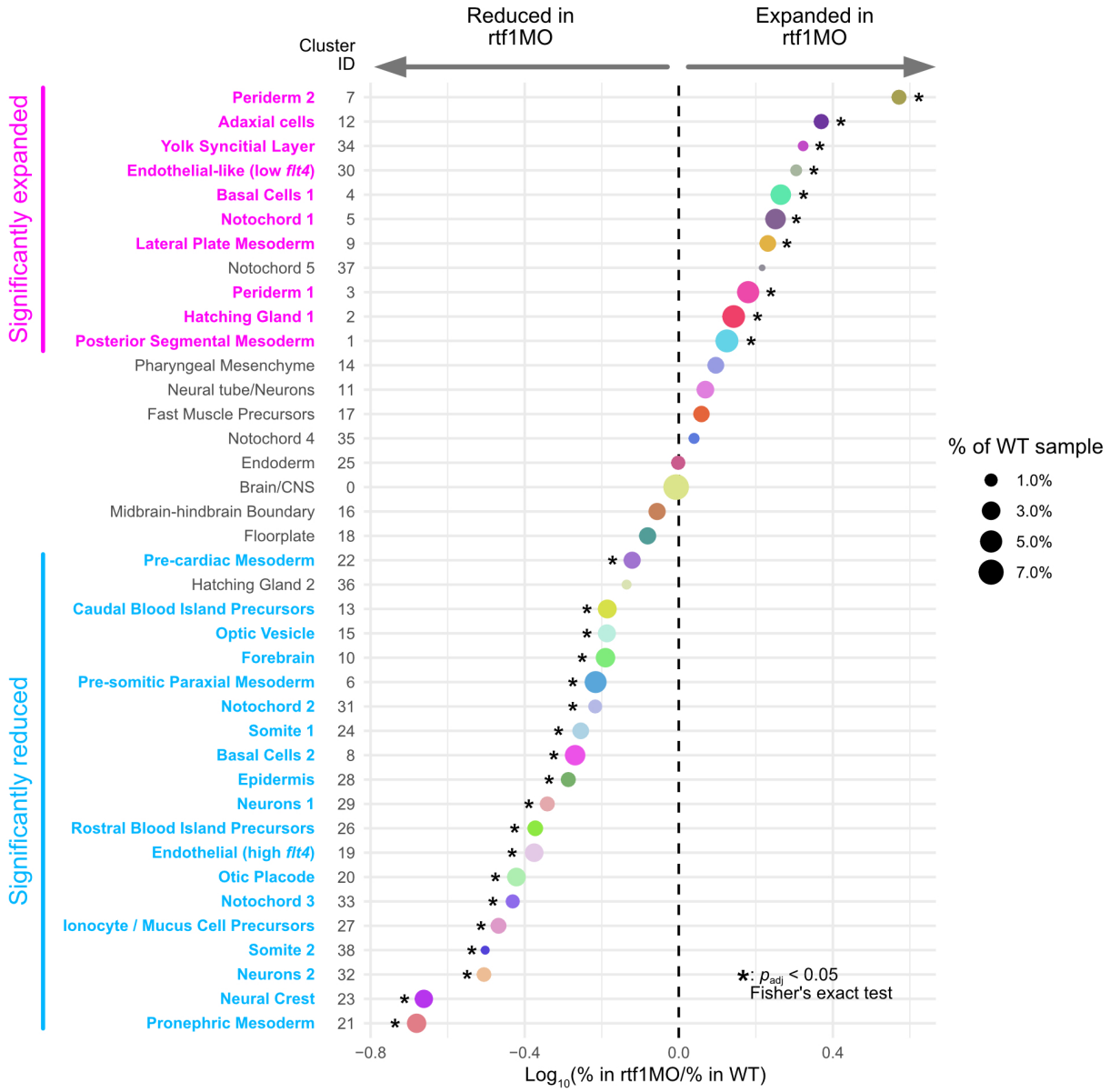


Figure 6 Supplement 3. Comparison of proportions of cells belonging to each Seurat cluster identified in the integrated multimodal single cell sequencing analysis of 11-12 somite stage control and *rtf1* morphant embryos. Predicted identities of each cell type based on marker gene expression are displayed to the left of the Seurat cluster IDs. Cell type names are colored based on whether they are significantly expanded (pink) or significantly reduced (light blue) in *rtf1* morphants based on Fisher's exact test with an adjusted *p*-value cutoff of 0.05.

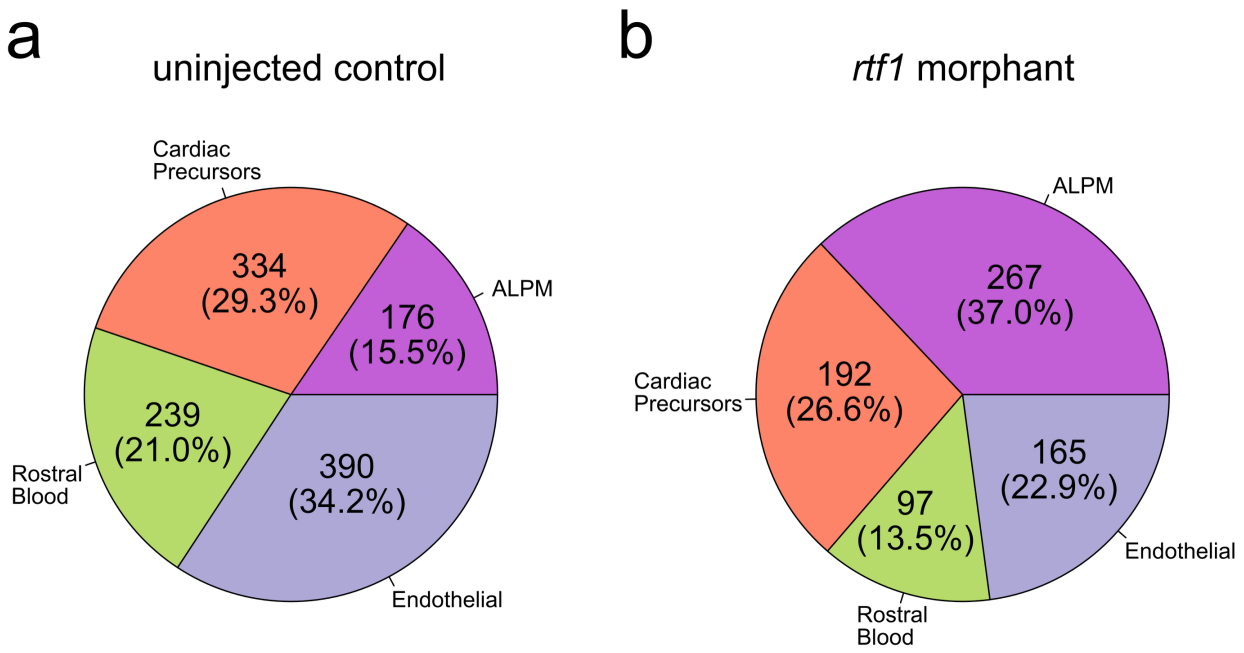


Figure 6 Supplement 4. Pie charts displaying the abundances of cell types identified in analysis of re-clustered anterior lateral plate mesoderm (ALPM) and derivatives from merged and integrated single cell RNA-seq datasets of 11-12 somite stage uninjected control (a) and *rtf1* morphant (b) embryos.

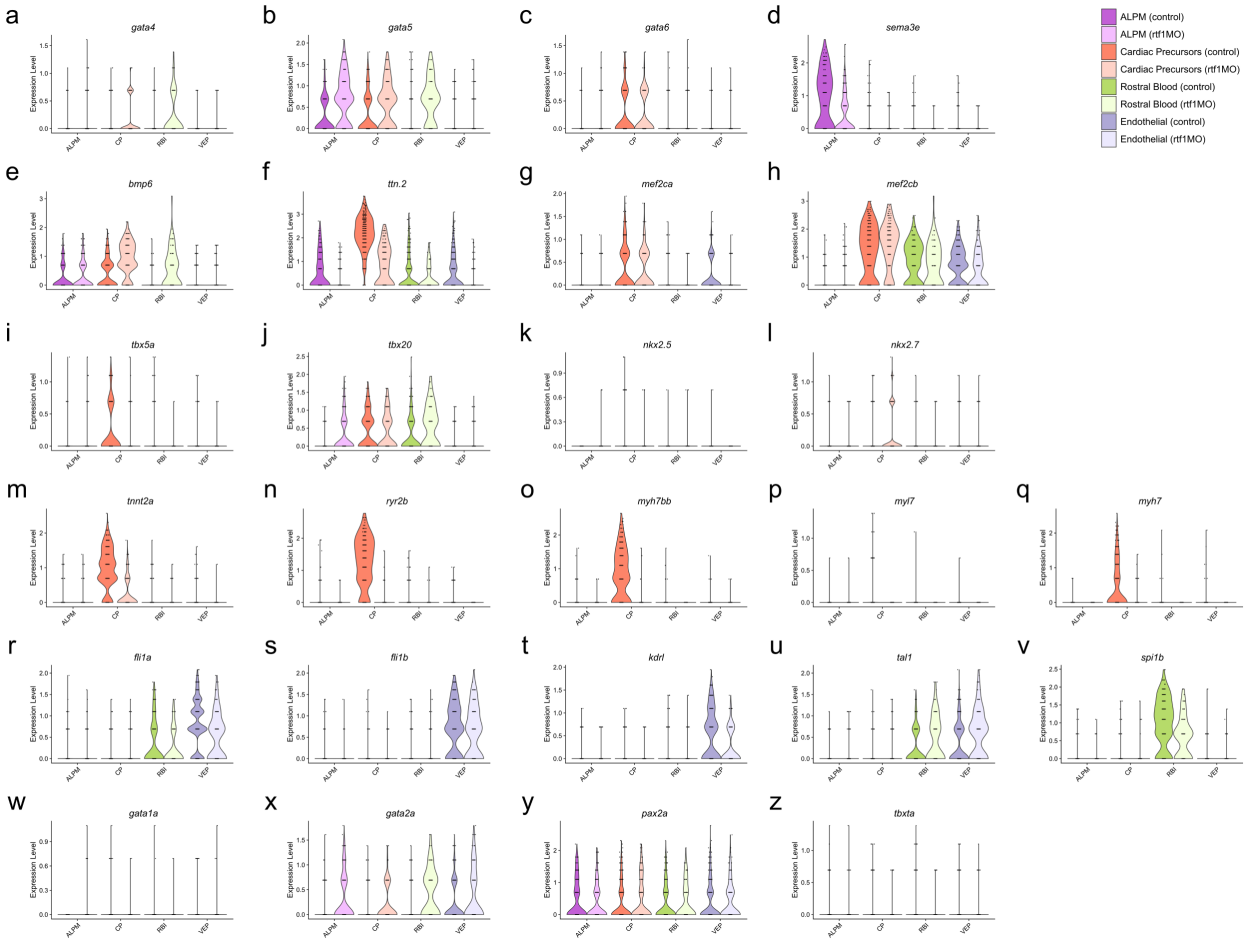
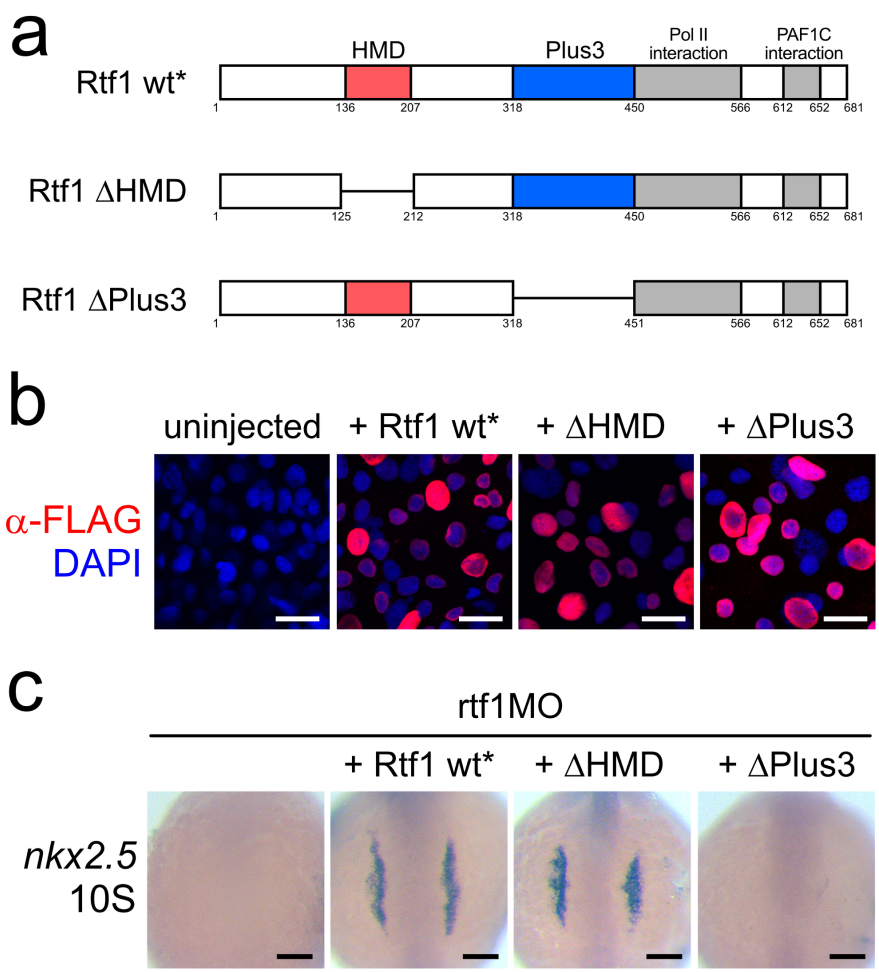


Figure 6 Supplement 5. (a-z) Violin plots showing the expression of genes relevant to the development of the lateral plate mesoderm and its derivatives. Each dot represents the expression level of a gene in a single cell. Violin-shaped areas are color-coded based on cell types identified by re-clustering ALPM and derivative cells from merged and integrated single cell RNA-seq datasets of 11-12 somite stage uninjected control and rtf1 morphant embryos. Shaded violins are not present when statistically significant expression of a particular gene was not detected.



887

888 **Figure 7. Rtf1's Plus3 domain is necessary for cardiac progenitor formation.** (a)
889 Schematics of Rtf1 wildtype (wt*) and Rtf1 mutant constructs (ΔHMD and ΔPlus3). Domains are
890 indicated by colored boxes (HMD: red, Plus3: blue, PolII and PAF1C interaction: grey). Deleted
891 regions are represented by lines. Numbers refer to the amino acid positions in wildtype Rtf1.
892 Rtf1 wt*, ΔHMD, and ΔPlus3 also harbor 8 silent mutations that alter the *rtf1* translation blocking
893 morpholino binding site. (b) Projections of confocal z-stacks of whole mount immunostaining
894 detecting N-terminal FLAG-tagged Rtf1 constructs (red) expressed in 75% epiboly zebrafish
895 embryos. Nuclei are labeled by DAPI staining (blue). Scale bars represent 20 μm. (c)
896 Representative images of RNA in situ hybridization detecting *nkx2.5* expression in 10 somite
897 stage (10S) zebrafish embryos coinjected with *rtf1* morpholino and mRNA encoding Rtf1
898 wildtype (wt*) or mutant (ΔHMD and ΔPlus3) proteins. Scale bars represent 0.1 mm.

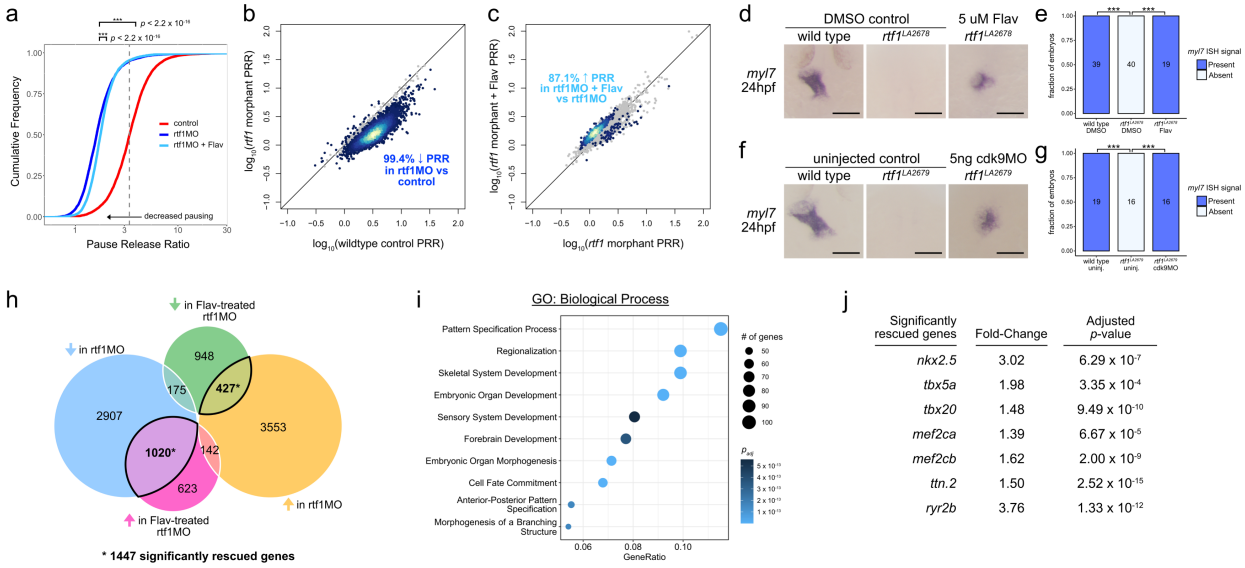


Figure 8. Rtf1-dependent transcriptional pausing regulates cardiogenesis. (a) Cumulative frequency plot of pause release ratios (PRRs) of 6,078 genes with substantial RNA Pol II signal. Both control (red) and flavopiridol treated *rtf1* morphant (light blue) samples displayed significantly different PRR frequencies compared to untreated *rtf1* morphants (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) Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf control and Rtf1-deficient (± flavopiridol) zebrafish embryos. (e) Quantification of *myl7* signal intensity in control and Rtf1-deficient (± flavopiridol) embryos at 24 hpf. Numbers on bars indicate the number of embryos analyzed. ***: $p < 0.001$. (f) Representative images of RNA in situ hybridization detecting *myl7* expression in 24 hpf control and Rtf1-deficient (± *cdk9* morpholino)

917 zebrafish embryos. (g) Quantification of *myl7* signal intensity in control and Rtf1-deficient (±
918 *cdk9* morpholino) embryos at 24 hpf. Numbers on bars indicate the number of embryos
919 analyzed. ***: p <0.001. (h) Venn diagram of significantly altered gene expression in 8-9 somite
920 stage *rtf1* morphant embryos (± flavopiridol treatment). Shaded circles represent genes that are
921 significantly: downregulated in *rtf1* morphants vs. uninjected controls (blue), upregulated in *rtf1*
922 morphants vs. uninjected controls (yellow), downregulated in flavopiridol-treated *rtf1* morphants
923 vs. vehicle-treated (DMSO) *rtf1* morphants (green), and upregulated in flavopiridol-treated *rtf1*
924 morphants vs. vehicle-treated *rtf1* morphants (pink). Overlapping regions with thick black
925 outlines represent the 1,447 genes (*) with expression levels that were significantly rescued in
926 8-9 somite stage *rtf1* morphants by flavopiridol treatment. (i) Plot of enriched Biological Process
927 gene ontology terms for the set of genes that were significantly rescued in *rtf1* morphants by
928 flavopiridol treatment. Dot sizes indicate the number of genes associated with a given ontology
929 term. Dot color indicates the significance level (adjusted p-value). (j) Selected list of genes that
930 are critical to cardiac development and were significantly rescued in *rtf1* morphants by
931 flavopiridol treatment. Fold-changes and adjusted p-values were calculated with DESeq2. Scale
932 bars in panels d and f represent 0.1 mm.
933
934
935
936
937
938
939
940
941
942

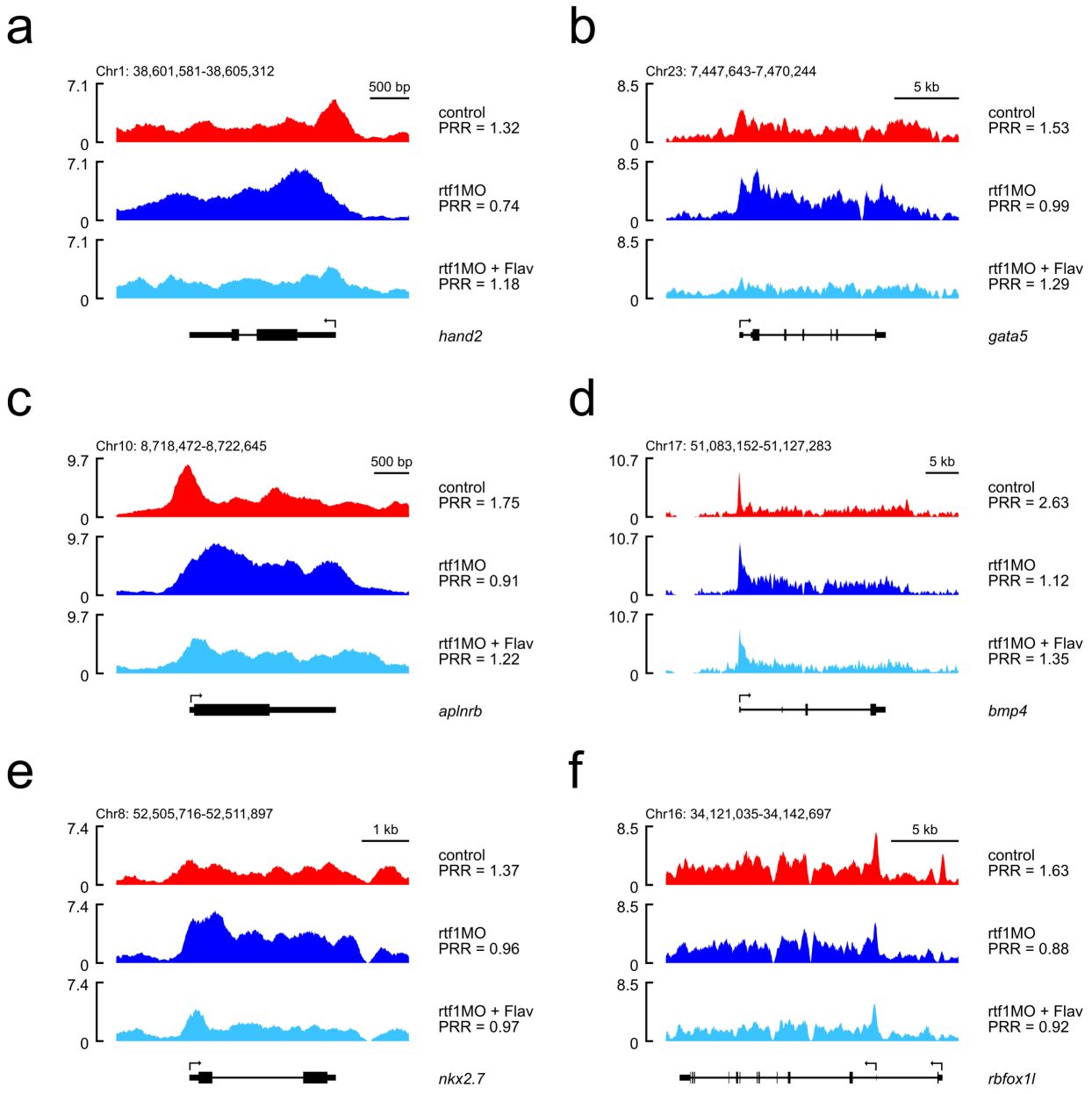


Figure 8 Supplement 1. Representative ChIP-seq tracks displaying RNA Pol II read densities (y-axis) for control (red), *rtf1* morphant (blue), and flavopiridol-treated *rtf1* morphant (light blue) embryos at cardiac mesoderm-related genes including *hand2* (a), *gata5* (b), *aplnrb* (c), *bmp4* (d), *nkx2.7* (e), and *rbfox1l* (f).

950 **Supplementary Tables:**

951

952 **Supplementary Table 1.** Sequences of primers used for quantitative real-time PCR analysis of

953 gene expression in mouse embryonic stem cells.

954

955 **Supplementary Table 2.** Predicted identities of cell types belonging to each Seurat cluster

956 identified in analysis of 11-12 somite stage control and *rtf1* morphant embryo integrated

957 multimodal (ATAC + Gene Expression) single cell sequencing datasets. Predicted identities

958 were based on manual inspection of marker gene expression.

959

960 **Supplementary Data Files:**

961

962 **Supplementary Data File 1.** DESeq2 results file containing differential gene expression

963 analysis data comparing FACS-isolated hand2:GFP-positive LPM from control and *rtf1*

964 morphant 10-12 somite stage embryos.

965

966 **Supplementary Data File 2.** Significantly restricted marker genes for Seurat cell clusters from

967 integrated multimodal (ATAC + Gene Expression) single cell sequencing dataset from control

968 and *rtf1* morphant embryos at the 11-12 somite stage.

969

970

971

972

973

974

975

976 **References**:

977

978 1 Bruneau, B. G. Signaling and transcriptional networks in heart development and regeneration.

979 *Cold Spring Harb Perspect Biol* **5**, a008292, doi:10.1101/cshperspect.a008292 (2013).

980 2 Paige, S. L., Plonowska, K., Xu, A. & Wu, S. M. Molecular regulation of cardiomyocyte

981 differentiation. *Circ Res* **116**, 341-353, doi:10.1161/CIRCRESAHA.116.302752 (2015).

982 3 Rana, M. S., Christoffels, V. M. & Moorman, A. F. A molecular and genetic outline of cardiac

983 morphogenesis. *Acta Physiol (Oxf)* **207**, 588-615, doi:10.1111/apha.12061 (2013).

984 4 Lu, F., Langenbacher, A. D. & Chen, J. N. Transcriptional Regulation of Heart Development in

985 Zebrafish. *J Cardiovasc Dev Dis* **3**, doi:10.3390/jcdd3020014 (2016).

986 5 Schoenebeck, J. J., Keegan, B. R. & Yelon, D. Vessel and blood specification override cardiac

987 potential in anterior mesoderm. *Dev Cell* **13**, 254-267, doi:10.1016/j.devcel.2007.05.012 (2007).

988 6 Mao, L. M. F., Boyle Anderson, E. A. T. & Ho, R. K. Anterior lateral plate mesoderm gives rise to

989 multiple tissues and requires tbx5a function in left-right asymmetry, migration dynamics, and

990 cell specification of late-addition cardiac cells. *Dev Biol* **472**, 52-66,

991 doi:10.1016/j.ydbio.2021.01.007 (2021).

992 7 Duong, T. B., Holowiecki, A. & Waxman, J. S. Retinoic acid signaling restricts the size of the first

993 heart field within the anterior lateral plate mesoderm. *Dev Biol* **473**, 119-129,

994 doi:10.1016/j.ydbio.2021.02.005 (2021).

995 8 Sorrell, M. R. & Waxman, J. S. Restraint of Fgf8 signaling by retinoic acid signaling is required for

996 proper heart and forelimb formation. *Dev Biol* **358**, 44-55, doi:10.1016/j.ydbio.2011.07.022

997 (2011).

998 9 Hota, S. K. *et al.* Dynamic BAF chromatin remodeling complex subunit inclusion promotes
999 temporally distinct gene expression programs in cardiogenesis. *Development* **146**,
1000 doi:10.1242/dev.174086 (2019).
1001 10 Takeuchi, J. K. *et al.* Chromatin remodelling complex dosage modulates transcription factor
1002 function in heart development. *Nat Commun* **2**, 187, doi:10.1038/ncomms1187 (2011).
1003 11 Lou, X., Deshwar, A. R., Crump, J. G. & Scott, I. C. Smarcd3b and Gata5 promote a cardiac
1004 progenitor fate in the zebrafish embryo. *Development* **138**, 3113-3123, doi:10.1242/dev.064279
1005 (2011).
1006 12 Lickert, H. *et al.* Baf60c is essential for function of BAF chromatin remodelling complexes in
1007 heart development. *Nature* **432**, 107-112, doi:10.1038/nature03071 (2004).
1008 13 Takeuchi, J. K. & Bruneau, B. G. Directed transdifferentiation of mouse mesoderm to heart tissue
1009 by defined factors. *Nature* **459**, 708-711, doi:10.1038/nature08039 (2009).
1010 14 Endoh, M. *et al.* Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro.
1011 *Mol Cell Biol* **24**, 3324-3336, doi:10.1128/MCB.24.8.3324-3336.2004 (2004).
1012 15 Hartzog, G. A., Wada, T., Handa, H. & Winston, F. Evidence that Spt4, Spt5, and Spt6 control
1013 transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. *Genes Dev* **12**, 357-
1014 369, doi:10.1101/gad.12.3.357 (1998).
1015 16 Keegan, B. R. *et al.* The elongation factors Pandora/Spt6 and Foggy/Spt5 promote transcription
1016 in the zebrafish embryo. *Development* **129**, 1623-1632, doi:10.1242/dev.129.7.1623 (2002).
1017 17 Wada, T. *et al.* DSIF, a novel transcription elongation factor that regulates RNA polymerase II
1018 processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* **12**, 343-356,
1019 doi:10.1101/gad.12.3.343 (1998).

1020 18 Wier, A. D., Mayekar, M. K., Heroux, A., Arndt, K. M. & VanDemark, A. P. Structural basis for
1021 Spt5-mediated recruitment of the Paf1 complex to chromatin. *Proc Natl Acad Sci U S A* **110**,
1022 17290-17295, doi:10.1073/pnas.1314754110 (2013).
1023 19 Chen, J. J. *et al.* Spt5 Phosphorylation and the Rtf1 Plus3 Domain Promote Rtf1 Function through
1024 Distinct Mechanisms. *Mol Cell Biol* **40**, doi:10.1128/MCB.00150-20 (2020).
1025 20 Hu, S. *et al.* SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains
1026 the enhancer landscape. *Mol Cell* **81**, 4425-4439 e4426, doi:10.1016/j.molcel.2021.08.029
1027 (2021).
1028 21 Park, J., Park, S. & Lee, J. S. Role of the Paf1 complex in the maintenance of stem cell
1029 pluripotency and development. *FEBS J* **290**, 951-961, doi:10.1111/febs.16582 (2023).
1030 22 Van Oss, S. B., Cucinotta, C. E. & Arndt, K. M. Emerging Insights into the Roles of the Paf1
1031 Complex in Gene Regulation. *Trends Biochem Sci* **42**, 788-798, doi:10.1016/j.tibs.2017.08.003
1032 (2017).
1033 23 Jaehning, J. A. The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim*
1034 *Biophys Acta* **1799**, 379-388, doi:10.1016/j.bbagrm.2010.01.001 (2010).
1035 24 Langenbacher, A. D. *et al.* The PAF1 complex differentially regulates cardiomyocyte
1036 specification. *Dev Biol* **353**, 19-28, doi:10.1016/j.ydbio.2011.02.011 (2011).
1037 25 Nguyen, C. T., Langenbacher, A., Hsieh, M. & Chen, J. N. The PAF1 complex component Leo1 is
1038 essential for cardiac and neural crest development in zebrafish. *Dev Biol* **341**, 167-175,
1039 doi:10.1016/j.ydbio.2010.02.020 (2010).
1040 26 Akanuma, T., Koshida, S., Kawamura, A., Kishimoto, Y. & Takada, S. Paf1 complex homologues
1041 are required for Notch-regulated transcription during somite segmentation. *EMBO reports* **8**,
1042 858-863, doi:10.1038/sj.embor.7401045 (2007).

1043 27 Lu, F., Langenbacher, A. & Chen, J. N. Tbx20 drives cardiac progenitor formation and
1044 cardiomyocyte proliferation in zebrafish. *Dev Biol* **421**, 139-148,
1045 doi:10.1016/j.ydbio.2016.12.009 (2017).
1046 28 Langenbacher, A. D. *et al.* Rtf1-dependent transcriptional pausing regulates cardiogenesis.
1047 *bioRxiv*, doi:10.1101/2023.10.13.562296 (2023).
1048 29 Saga, Y. *et al.* MesP1 is expressed in the heart precursor cells and required for the formation of a
1049 single heart tube. *Development* **126**, 3437-3447, doi:10.1242/dev.126.15.3437 (1999).
1050 30 Warner, M. H., Roinick, K. L. & Arndt, K. M. Rtf1 is a multifunctional component of the Paf1
1051 complex that regulates gene expression by directing cotranscriptional histone modification. *Mol*
1052 *Cell Biol* **27**, 6103-6115, doi:10.1128/MCB.00772-07 (2007).
1053 31 Vos, S. M., Farnung, L., Linden, A., Urlaub, H. & Cramer, P. Structure of complete Pol II-DSIF-PAF-
1054 SPT6 transcription complex reveals RTF1 allosteric activation. *Nat Struct Mol Biol* **27**, 668-677,
1055 doi:10.1038/s41594-020-0437-1 (2020).
1056 32 Cucinotta, C. E., Hildreth, A. E., McShane, B. M., Shirra, M. K. & Arndt, K. M. The nucleosome
1057 acidic patch directly interacts with subunits of the Paf1 and FACT complexes and controls
1058 chromatin architecture in vivo. *Nucleic Acids Res* **47**, 8410-8423, doi:10.1093/nar/gkz549 (2019).
1059 33 Van Oss, S. B. *et al.* The Histone Modification Domain of Paf1 Complex Subunit Rtf1 Directly
1060 Stimulates H2B Ubiquitylation through an Interaction with Rad6. *Mol Cell* **64**, 815-825,
1061 doi:10.1016/j.molcel.2016.10.008 (2016).
1062 34 Piro, A. S., Mayekar, M. K., Warner, M. H., Davis, C. P. & Arndt, K. M. Small region of Rtf1 protein
1063 can substitute for complete Paf1 complex in facilitating global histone H2B ubiquitylation in
1064 yeast. *Proc Natl Acad Sci U S A* **109**, 10837-10842, doi:10.1073/pnas.1116994109 (2012).
1065 35 de Jong, R. N. *et al.* Structure and DNA binding of the human Rtf1 Plus3 domain. *Structure* **16**,
1066 149-159, doi:10.1016/j.str.2007.10.018 (2008).

1067 36 Core, L. & Adelman, K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene
1068 regulation. *Genes Dev* **33**, 960-982, doi:10.1101/gad.325142.119 (2019).
1069 37 Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in
1070 metazoans. *Nat Rev Genet* **13**, 720-731, doi:10.1038/nrg3293 (2012).
1071 38 Peng, J., Zhu, Y., Milton, J. T. & Price, D. H. Identification of multiple cyclin subunits of human P-
1072 TEFb. *Genes Dev* **12**, 755-762, doi:10.1101/gad.12.5.755 (1998).
1073 39 Price, D. H. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol*
1074 *Cell Biol* **20**, 2629-2634, doi:10.1128/MCB.20.8.2629-2634.2000 (2000).
1075 40 Paparidis, N. F., Durvale, M. C. & Canduri, F. The emerging picture of CDK9/P-TEFb: more than
1076 20 years of advances since PITALRE. *Mol Biosyst* **13**, 246-276, doi:10.1039/c6mb00387g (2017).
1077 41 Fujinaga, K., Huang, F. & Peterlin, B. M. P-TEFb: The master regulator of transcription elongation.
1078 *Mol Cell* **83**, 393-403, doi:10.1016/j.molcel.2022.12.006 (2023).
1079 42 Jonkers, I., Kwak, H. & Lis, J. T. Genome-wide dynamics of Pol II elongation and its interplay with
1080 promoter proximal pausing, chromatin, and exons. *Elife* **3**, e02407, doi:10.7554/eLife.02407
1081 (2014).
1082 43 Huang, M. *et al.* Negative Elongation Factor (NELF) Inhibits Premature Granulocytic
1083 Development in Zebrafish. *Int J Mol Sci* **23**, doi:10.3390/ijms23073833 (2022).
1084 44 Yang, Q. *et al.* RNA polymerase II pausing modulates hematopoietic stem cell emergence in
1085 zebrafish. *Blood* **128**, 1701-1710, doi:10.1182/blood-2016-02-697847 (2016).
1086 45 Gaertner, B. & Zeitlinger, J. RNA polymerase II pausing during development. *Development* **141**,
1087 1179-1183, doi:10.1242/dev.088492 (2014).
1088 46 Gilchrist, D. A. *et al.* Regulating the regulators: the pervasive effects of Pol II pausing on
1089 stimulus-responsive gene networks. *Genes Dev* **26**, 933-944, doi:10.1101/gad.187781.112
1090 (2012).

1091 47 Adelman, K. *et al.* Immediate mediators of the inflammatory response are poised for gene
1092 activation through RNA polymerase II stalling. *Proc Natl Acad Sci U S A* **106**, 18207-18212,
1093 doi:10.1073/pnas.0910177106 (2009).
1094 48 Rougvie, A. E. & Lis, J. T. The RNA polymerase II molecule at the 5' end of the uninduced hsp70
1095 gene of D. melanogaster is transcriptionally engaged. *Cell* **54**, 795-804, doi:10.1016/s0092-
1096 8674(88)91087-2 (1988).
1097 49 Chen, F. X. *et al.* PAF1 regulation of promoter-proximal pause release via enhancer activation.
1098 *Science* **357**, 1294-1298, doi:10.1126/science.aan3269 (2017).
1099 50 Gilchrist, D. A. *et al.* Pausing of RNA polymerase II disrupts DNA-specified nucleosome
1100 organization to enable precise gene regulation. *Cell* **143**, 540-551,
1101 doi:10.1016/j.cell.2010.10.004 (2010).
1102 51 Gilchrist, D. A. *et al.* NELF-mediated stalling of Pol II can enhance gene expression by blocking
1103 promoter-proximal nucleosome assembly. *Genes Dev* **22**, 1921-1933, doi:10.1101/gad.1643208
1104 (2008).
1105 52 Jurynec, M. J. *et al.* The Paf1 complex and P-TEFb have reciprocal and antagonist roles in
1106 maintaining multipotent neural crest progenitors. *Development* **146**, doi:10.1242/dev.180133
1107 (2019).
1108 53 Wang, Z. *et al.* Coordinated regulation of RNA polymerase II pausing and elongation progression
1109 by PAF1. *Sci Adv* **8**, eabm5504, doi:10.1126/sciadv.abm5504 (2022).
1110 54 Westerfield, M. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio).* 4th
1111 ed. edn, (Univ. of Oregon Press, 2000).
1112 55 Farley, F. W., Soriano, P., Steffen, L. S. & Dymecki, S. M. Widespread recombinase expression
1113 using FLPeR (flipper) mice. *Genesis* **28**, 106-110 (2000).

1114 56 Chen, J. N. & Fishman, M. C. Zebrafish tinman homolog demarcates the heart field and initiates
1115 myocardial differentiation. *Development* **122**, 3809-3816 (1996).
1116 57 Robu, M. E. *et al.* p53 activation by knockdown technologies. *PLoS Genet* **3**, e78,
1117 doi:10.1371/journal.pgen.0030078 (2007).
1118 58 Yin, C., Kikuchi, K., Hochgreb, T., Poss, K. D. & Stainier, D. Y. Hand2 regulates extracellular matrix
1119 remodeling essential for gut-looping morphogenesis in zebrafish. *Dev Cell* **18**, 973-984,
1120 doi:10.1016/j.devcel.2010.05.009 (2010).
1121 59 Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions,
1122 deletions and gene fusions. *Genome Biol* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).
1123 60 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for
1124 assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930,
1125 doi:10.1093/bioinformatics/btt656 (2014).
1126 61 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
1127 RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
1128 62 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological
1129 themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).
1130 63 Chicas, A. *et al.* Dissecting the unique role of the retinoblastoma tumor suppressor during
1131 cellular senescence. *Cancer Cell* **17**, 376-387, doi:10.1016/j.ccr.2010.01.023 (2010).
1132 64 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-
1133 359, doi:10.1038/nmeth.1923 (2012).
1134 65 Lun, A. T. & Smyth, G. K. csaw: a Bioconductor package for differential binding analysis of ChIP-
1135 seq data using sliding windows. *Nucleic Acids Res* **44**, e45, doi:10.1093/nar/gkv1191 (2016).

1136 66 Lun, A. T. & Smyth, G. K. De novo detection of differentially bound regions for ChIP-seq data
1137 using peaks and windows: controlling error rates correctly. *Nucleic Acids Res* **42**, e95,
1138 doi:10.1093/nar/gku351 (2014).
1139 67 Kurtenbach, S. & Harbour, J. W. SparK: A Publication-quality NGS Visualization Tool. *bioRxiv*,
1140 845529, doi:10.1101/845529 (2019).
1141 68 Lawson, N. D. *et al.* An improved zebrafish transcriptome annotation for sensitive and
1142 comprehensive detection of cell type-specific genes. *Elife* **9**, doi:10.7554/eLife.55792 (2020).
1143 69 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529,
1144 doi:10.1016/j.cell.2021.04.048 (2021).
1145 70 Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat*
1146 *Methods* **16**, 1289-1296, doi:10.1038/s41592-019-0619-0 (2019).
1147 71 Song, M. *et al.* GATA4/5/6 family transcription factors are conserved determinants of cardiac
1148 versus pharyngeal mesoderm fate. *Sci Adv* **8**, eabg0834, doi:10.1126/sciadv.abg0834 (2022).
1149 72 Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**,
1150 496-502, doi:10.1038/s41586-019-0969-x (2019).

1151