# Ventricular, atrial and outflow tract heart progenitors arise from spatially and molecularly distinct regions of the primitive streak.

Kenzo Ivanovitch[1,*], Pablo Soro-Barrio[1], Probir Chakravarty[1], Rebecca A Jones[1], S. Neda Mousavy Gharavy[1], Despina Stamataki[1], Julien Delile[1], James C Smith[1] and James Briscoe[1]

[1]The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

[*] Corresponding author

E-mail: kenzo.ivanovitch@crick.ac.uk (K.I.)

## Abstract

The heart develops from two sources of mesoderm progenitors, the first and second heart field (FHF and SHF). Using a single cell transcriptomic assay in combination with genetic lineage tracing, we find the FHF and SHF are subdivided into distinct pools of progenitors in gastrulating mouse embryos at earlier stages than previously thought. Each subpopulation has a distinct origin in the primitive streak. The first progenitors to leave the primitive streak contribute to the left ventricle, shortly after right ventricle progenitor emigrate, followed by the outflow tract and atrial progenitors. Although cells allocated to the outflow tract and atrium leave the primitive streak at a similar stage, they arise from different regions. Outflow tract originate from distal locations in the primitive streak while atrial progenitors are positioned more proximally. Moreover, single cell RNA sequencing demonstrates that the primitive streak cells contributing to the ventricles have a distinct molecular signature from those forming the outflow tract and atrium. We conclude that cardiac progenitors are pre-patterned within the primitive streak and this prefigures their allocation to distinct anatomical structures of the heart. Together, our data provide a new molecular and spatial map of mammalian cardiac progenitors that will support future studies of heart development, function and disease.

## Introduction

The cells that comprise different components of an organ can arise from distinct origins and acquire their fate at different time during ontogeny. In the case of the heart, two mesodermal

30 derived sources of cardiac progenitors, named the first heart field (FHF) and the second heart
31 field (SHF), have been identified (reviewed in Meilhac and Buckingham, 2018). The FHF
32 forms mainly the left ventricle and constitutes the initial cardiac crescent. By contrast, the
33 cardiac progenitors of the SHF contribute to the right ventricle, outflow tract and atria in
34 addition to branchiomeric muscles.

35 Retrospective clonal analysis has suggested that the FHF and SHF progenitors segregate
36 either before or at the onset of gastrulation (Meilhac and Buckingham, 2018; Meilhac et al.,
37 2004). The analysis showed large clones spanning multiple compartments. Clones belonging
38 to the first lineage labelled the left ventricle and other compartments with the exception of the
39 outflow tract while the second lineage contributed to the outflow tract, atria and right
40 ventricle but never to the left ventricle. Subclones within the first and second lineages were
41 also observed. These were smaller in size and suggested that the individualisation of the
42 different regions of the heart happens at later stages (Meilhac et al., 2004).

43 Consistent with the notion that the cardiac lineages are established during gastrulation, clonal
44 analysis based on tracing the progeny of *Mesp1*-expressing cells, which is expressed in the
45 primitive streak and nascent mesoderm
46 (https://marionilab.cruk.cam.ac.uk/MouseGastrulation2018/; Arnold and Robertson, 2009;
47 Costello et al., 2011; Pijuan-Sala et al., 2019), indicated that independent sets of *Mesp1*
48 expressing cells contribute to FHF and SHF derivatives (Devine et al., 2014; Lescroart et al.,
49 2014). These can be distinguished by their time of appearance during gastrulation, with
50 embryonic day (E)6.5 *Mesp1*+ cells contributing to the FHF while SHF derivatives

51 preferentially derive from E7.5 *Mesp1*+ cells (Lescroart et al., 2014). The *Mesp1* clonal
52 analysis resulted in clones of small sizes restricted to compartments derived from the FHF or
53 SHF (Devine et al., 2014; Lescroart et al., 2014). Thus, these observations are consistent with
54 the existence of two groups of progenitors that subsequently became further restricted to
55 specific cardiac fates in the mesoderm.

56 In addition to a segregation between the FHF and SHF, genetic tracing experiments
57 uncovered a distinction between the atria and ventricular progenitors. Genetic lineage tracing
58 with the transcription factor *Foxa2* identifies a population of cardiac progenitors in the
59 anterior primitive streak at E6.5 that contribute to both left and right ventricles but not to the
60 atria (Bardot et al., 2017). This is in line with fate mapping studies in the chick showing that
61 the atrial and ventricular cells arise at different anterior-posterior positions in the primitive
62 streak (Garcia-Martinez and Schoenwolf, 1993; Yutzey and Bader, 1995). Although these
63 analyses did not resolve the clonal relationship between the cells, these findings are
64 compatible with a model in which sublineages within the FHF and SHF lineages already exist
65 in the primitive streak, such that the atrial lineage is distinct from the right ventricle lineage
66 within SHF progenitors. This would imply that two distinct cardiac progenitors exist solely at
67 the epiblast stage prior to gastrulation and that more subpopulations of cardiac progenitors
68 than first and second progenitors can be molecularly defined in the primitive streak. The most
69 rigorous way to decide if this is the case is to define the location of all the cardiac progenitors
70 in the streak and the precise embryonic stages at which they ingress. This would allow the
71 comparison of distinct subpopulations of cardiac precursors and reveal putative molecular
72 differences of cardiac progenitors in the primitive streak. Single cell transcriptomic assays

show *Mesp1*+ cardiac progenitors diversify into molecularly distinct FHF and SHF populations at late embryonic day (E7.25) stage once cells have migrated and reached their final location in the embryo (Lescroart et al., 2018). The signalling environment cells encounter during and after their migration might play a role in the patterning of the cardiac progenitors into FHF and SHF domains (Andersen et al., 2018; Evan et al 2020; Guzzetta et al., 2020; Hochgreb et al., 2003; Xavier-Neto et al., 1999). Whether initial molecular differences between different sets of cardiac progenitors already exist in the primitive streak remains unclear.

In this study, we use genetic tracing of *T* and *Foxa2* expressing cells and find there is an orderly allocation of primitive streak cells first into the left ventricle at mid-streak stage, then the right ventricle progenitors at late streak stage, and finally at the no bud (OB)-early bud (EB) stage into outflow tract and atrial progenitors. Consistent with this, we identified independent sets of *Foxa2*+ cells that are allocated to the left ventricle and right ventricle. Allocation of cells to the outflow tract and atria happens at similar gastrulation stages but from distinct locations within the primitive streak. The outflow tract originates from distal regions of the primitive streak, while proximal regions contribute to the atria. Moreover, the outflow tract forms from primitive streak cells that initially expressed *Foxa2*, but subsequently turned off *Foxa2* expression as they switched their contribution from right ventricle to the outflow tract. Crucially, by combining single cell transcriptomic assays with a lineage tracer to label cells contributing only to the poles of the heart, we uncover molecularly distinct subpopulations of cells that correspond to progenitors for right ventricle, outflow tract and atria in the SHF. Further single cell transcriptomic experiments at earlier

95 stages established that the primitive streak cells contributing to the ventricles and outflow
96 tract/atria are also molecularly distinct. Thus, rather than a simple subdivision of cardiac
97 progenitors into FHF and SHF, our analysis reveals a more elaborate map for the source of
98 cells that form the heart with distinct spatial and temporal origins for outflow tract and atrial
99 progenitors as well as left and right ventricular progenitors. We conclude that the cardiac
100 progenitors are pre-patterned within the primitive streak and this prefigures their contribution
101 to distinct anatomical structures of the heart both in time and space. This has implications for
102 the classification of congenital heart diseases based on the origin of malformation in a
103 specific mesodermal lineage as well as for the design of *in vitro* methods to generate specific
104 cardiac cells from pluripotent stem cells.

105 **Results**

106 **Genetic tracing of primitive streak cells using a tamoxifen inducible T reporter.**

107 During gastrulation, epiblast cells ingress through the primitive streak to form the mesoderm.
108 This process occurs over an extended period of time, from embryonic day (E)6 to E8 in the
109 mouse. The T-box transcription factor *T* is expressed in the primitive streak and is
110 downregulated shortly after ingression during migration within the nascent mesoderm. To
111 assess the developmental time points at which the *T*-expressing primitive streak cells are
112 destined to contribute to the heart, we performed genetic tracing using an inducible $T^{nGFP-}$
113 $^{CreERT2/+}$ mouse, expressing *CreERT2* downstream of the endogenous *T* (Imuta et al., 2013), in

114 combination with the *R26R*$^{tdtomato/+}$ reporter mice (*T*$^{nEGP\text{-}CreERT2/+}$; *R26R*$^{tdtomato/+}$, Fig 1A)
115 (Madisen et al., 2010).

116 We first assessed how long the activity of the tamoxifen persists after it is administered by
117 oral gavage in a pregnant mouse. We administered tamoxifen (0.08mg/bw) at E5, ~24 hours
118 prior to the initial onset of *T*-expression in the primitive streak (Rivera-Perez and Magnuson,
119 2005), and we detected recombined tdtomato-expressing cells within mesodermal derivatives
120 including the heart tube, head mesenchyme and endothelium, albeit at a low density, in E8.5
121 embryos, (S1A Fig.). Thus, the activity of the tamoxifen persists for at least around 24 hours
122 at this dose.

123 **Primitive streak cells contribute first to the ventricles and subsequently to the outflow**
124 **tract and atria myocardium.**

125 We next administered single doses of tamoxifen *T*$^{nEGP\text{-}CreERT2/+}$; *R26R*$^{tdtomato/+}$ between E6.5 and
126 E7.5. We reasoned this would result in pulse-labelling of most populations of *T*-expressing
127 cells from the time of administration for at least 24 hours (Fig 1B). The fate of their progeny
128 could then be followed and the developmental stages during which primitive streak cells
129 contribute to the left and right ventricle, atria and outflow tract could be deduced. To gain
130 better control over the embryonic stages, mice were synchronized in estrus and mated over
131 short periods of 2 hours for all the experiments. To quantify the contribution of the *T*-
132 expressing cells to the heart, we measured surface area occupied by the tdtomato-positive
133 cells in the heart myocardium at E12.5 and within each cardiac chamber.

134 An early administration of tamoxifen at E6+21h or E7 resulted in the population of left and
135 right ventricles, outflow tract and Atria with tdtomato-expressing cardiomyocytes (in 17 out of
136 17 hearts analysed, Fig 1Ci-iv, D and E). The contribution of the tdtomato-positive cells to
137 the left ventricle was lower compared to the right ventricle (fig 1e, for E6+21h, left ventricle:
138 22.2% ±11.1; right ventricle: 40.2% ± 12.2 and for E7, left ventricle: 24.8% ±16, and right
139 ventricle: 33.4% ± 9.9, mean ± sd). Contribution to the outflow tract and atria was the
140 highest, covering almost their entirety in all cases (Fig 1Ci-iv and Fig 1E, 89.1%± 8.4 and %
141 94.1± 6.9 on average, for outflow tract and atria respectively). Variability in embryonic
142 stages within litters at the time of tamoxifen administration can explain the variability in the
143 results (see below). For example, in 4/17 hearts, tdtomato-expressing cardiomyocytes
144 populated less than 10% of the total left ventricle surface area. In those hearts, the proportion
145 of tdtomato-positive cardiomyocytes found in the right ventricle was high (Fig 1Cii, iv and
146 D). We never observed the reverse, however, where a lower proportion of tdtomato positive
147 cardiomyocytes was found in the right ventricle compared to the left ventricle (Fig 1D).
148 These results suggest that subsets of the embryos at E6+21h and E7 were at more advanced
149 embryonic stages with most left ventricle precursors having already left the primitive streak
150 with right ventricular precursors remaining in the primitive streak.

151 We next asked if the administration of tamoxifen at a later time point would result in the
152 absence of tdtomato-expressing cardiomyocytes in the left ventricle and right ventricle. We
153 administered tamoxifen at E7+7h. Tdtomato-expressing cardiomyocytes were detected in the
154 left ventricle in only 2/13 hearts (Fig 1D). Tdtomato-positive cardiomyocytes were detected
155 more frequently in the right ventricle, in 7/13 hearts (Fig 1D). However, these cells covered

156 less than 5% of the total right ventricle surface area (yellow arrow in Fig 1Cv and 1E), except
157 for 2 hearts in which tdtomato-expressing cardiomyocytes were also detected in the left
158 ventricle. These results indicate that left ventricle and most right ventricle precursors have
159 left the primitive streak by E7+7h. In these embryos, tdtomato-expressing cardiomyocytes
160 were evident in the outflow tract and atria in 13/13 hearts, covering in some cases between
161 72.8% and 94% of the total surface area of the outflow tract and atria, including hearts
162 without tdtomato-expressing cardiomyocytes in the left ventricle and right ventricle. These
163 results suggest that the majority of the outflow tract and atria precursors are still located in
164 the primitive streak at stages after the ventricular precursors have already left the primitive
165 streak.

166 Finally, when tamoxifen was administered at a later stage (E7+11h), the contribution of the
167 tdtomato-expressing cardiomyocytes to the outflow tract and atria was lower (Fig 1Cvii-viii
168 and E, 12.6% ± 17.4 and 19.7% ± 20.8). In all cases, these embryos lacked tdtomato-
169 expressing cells in the left and right ventricles. In 5/14 hearts, tdtomato-expressing cells were
170 absent in the outflow tract but present in the atria albeit in low numbers. The reverse, *i.e.*
171 tdtomato expressing cells found only in the outflow tract but not in the atria, was never
172 observed, suggesting that the atrial precursors are the last cardiomyocyte precursors to leave
173 the primitive streak.

174 To analyse variability in embryonic stages at the time of tamoxifen administration, we
175 dissected litters of $T^{nEGP\text{-}CreERT2/+}$; $R26R^{tdtomato/+}$ embryos at E6+21h and E7+7h and staged
176 embryos according to morphological landmarks using the dissecting brightfield microscope

177 (Downs et al., 1993). As expected, we found a range of stages for each time point. Litters
178 dissected at E6+21h included early (17%), mid (56%) and late-streak stages (27%). Litters
179 dissected at E7+7h included late streak (10%), no bud (OB, 37%), early bud (EB, 32%) and
180 late bud stages (LB, 21%) (Fig 1F and H). PCR analysis showed that the *R26R*$^{tdtomato/+}$ locus
181 was recombined as early as two hours following oral gavage of tamoxifen (0.08mg/bw) (Fig
182 1G).

183 Together, these results indicate that cardiac progenitors ingress in an orderly sequence
184 through the primitive streak. We propose left ventricle progenitors leave the primitive streak
185 first, at the mid-streak stage, followed by right ventricle progenitors at the late-streak stage.
186 Outflow tract and atrial precursors leave the primitive streak at subsequent stages starting
187 around the OB stage.

188 **Independent *Foxa2*-expressing cells contribute to either the left or right ventricular**
189 **myocardium at similar stages to the *T*-expressing primitive streak cells.**

190 Genetic tracing experiments at around E6-E6.5 (Frank et al., 2007; Bardot et al., 2017)
191 showed that *Foxa2* positive cells contribute to cardiomyocytes in the ventricles but not in the
192 atria. These results prompted us to first test whether *Foxa2* positive cells contribute to
193 ventricular myocardium at similar stages to the *T* expressing cells *i.e.* at the mid-late streak
194 stage. We performed genetic tracing of the *Foxa2* positive cells using a *Foxa2*$^{nEGP\text{-}CreERT2/+}$ line
195 expressing *CreERT2* downstream of the endogenous *Foxa2* in combination with the
196 *R26R*$^{tdtomato/+}$ reporter line (Imuta et al., 2013) (Fig 2A-B). Mice were synchronized in estrus
197 and timed matings performed for periods of 2 hours to better control embryonic stages.

198 Administration of tamoxifen at E6+21h resulted in large patches of tdtomato-expressing
199 cardiomyocytes populating the left and right ventricle and outflow tract in 8/15 hearts (Fig
200 2Ciii). In 3/14 hearts, smaller tdtomato-expressing domains formed either in the left or right
201 ventricular myocardium. This indicates that independent groups of *Foxa2* expressing
202 progenitors exist in the primitive streak that contribute to the left and right myocardium (Fig
203 2Ci-ii and D). We conclude that *Foxa2* positive primitive streak cells contribute to the
204 ventricular myocardium at similar stages to the *T*-expressing cells. Analysis of the embryos at
205 later stages, when the cardiac crescent is differentiating into cardiomyocytes, confirmed that
206 cells arising from *Foxa2* positive cells contribute to the myocardium in addition to other
207 lineages including the epicardium and pericardium but not the endocardium located below
208 the presumptive myocardial epithelium (S2A-B Fig).

209 This result prompted us to analyse the location of *Foxa2* and *T* expressing cells at the mid
210 and late streak stages. Whole mount immunostaining for T and Foxa2 revealed coexpression
211 in individual primitive streak cells at the mid-streak position of mid and late streak stages
212 (Fig 3Bii and Dii). This corresponds to where cardiomyocytes precursors have previously
213 been identified by fate mapping (Kinder et al., 1999). Double positive cells are also detected
214 at the distal tip of the embryo where the node forms (Fig 3Biii and Diii) and in the definitive
215 endoderm (Fig 3Bi and Di). These results indicate that *T* and *Foxa2* are not always expressed
216 in mutually exclusive populations of cells as previously reported (Burtscher and Lickert,
217 2009). Instead, it suggests *Foxa2* and *T* are co-expressed in a population of anterior primitive
218 streak cells that contribute to the ventricular myocardium and never to the atria. Finally,
219 Foxa2 is undetectable in the primitive streak from the OB stage onwards, the stages during

220 which primitive streak cells switched their contribution from the ventricles to the outflow
221 tract and atria.

222 **Distal primitive streak cells downregulate Foxa2 expression as they switch their**
223 **contribution from the ventricles towards the outflow tract myocardium.**

224 Following tamoxifen administration in *Foxa2*$^{nEGP\text{-}CreERT2/+}$ *R26R*$^{tdtomato/+}$ embryos at E6+21h,
225 we found tdtomato-expressing cells located in the outflow tract myocardium (Fig 2Ciii and D
226 and S2A). When we administered tamoxifen at E7+7h, no tdtomato-positive cardiomyocytes
227 could be detected in either the left ventricle, right ventricle or outflow tract (in 11 out of 13
228 hearts, Fig 2Civ). This could indicate that a subset of the *Foxa2* positive primitive streak
229 cells, at E6+21h, are outflow tract progenitors and not solely ventricular progenitors.
230 Alternatively, primitive streak cells contributing to the outflow tract may arise from cells that
231 had previously expressed *Foxa2* but no longer do so as the primitive streak switches its
232 contribution from right ventricular to outflow tract myocardium by E7+7h (Fig 4B).
233 Consistent with this latter hypothesis, the presence of tdtomato-expressing cardiomyocytes in
234 the outflow tract was always associated with the presence of tdtomato-expressing
235 cardiomyocytes in the right ventricle (in 8/8 hearts, Fig 2D).

236 We next administered tamoxifen in *Foxa2*$^{nEGP\text{-}CreERT2/+}$ *R26R*$^{tdtomato/+}$ embryos at E6+21h and
237 fixed the embryos 10 hours later. We immunostained the embryos for Foxa2 and T protein.
238 Computationally masking cells with Foxa2 expression (*i.e.* endodermal and axial mesoderm
239 cells, Fig 4Ai-ii) revealed tdtomato-expressing mesodermal cells (*i.e.* cells that had expressed

240 Foxa2) forming a stream of migratory cells originating from distal regions of the primitive
241 streak and finishing in the anterior region of the embryo. This corresponds to the prospective
242 cardiac crescent (dotted white arrow in Fig 4Aii and red arrows in Fig 4Aiii).

243 We then analysed all tdtomato-expressing cells within the distal primitive streak region and
244 found these cells expressed T protein, (yellow arrows in Fig 2Aiv), although none
245 coexpressed Foxa2 and T. Cells expressing Foxa2 but not T corresponded to endodermal
246 cells, as indicated by their location in the outer layer of the embryo (see green arrow in Fig
247 4Aiv). All these experiments show that the outflow tract myocardium arises from the distal
248 primitive streak while the atria progenitors are instead located in the proximal primitive
249 streak. The outflow tract progenitors are *Foxa2* negative, but they arise from cells that had
250 expressed *Foxa2* previously. We propose that distal primitive streak cells downregulate
251 *Foxa2* as they switch their contribution from the right ventricle to the outflow tract
252 myocardium (Fig 4B). Atrial cells instead arise from the proximal primitive streak, which had
253 not previously expressed *Foxa2*.

254 **Right ventricle, outflow tract and Atrial progenitors have distinct locations and**
255 **transcriptional profile within the heart fields.**

256 Having identified distinct population of T-expressing cells descendants, we reasoned their
257 final locations could correlate with the previously described heart fields. We analysed the
258 locations of the descendants of early E6+21h and late E7+7h *T*-expressing cells. These cells
259 formed the ventricles and the majority of the outflow tract/atria, respectively, at heart field
260 stage (see above). We used a BMP reporter (S5A-A" Fig) and assayed for Phospho(P)-

261 Smad1/5/8 to mark the FHF – the cardiac crescent - since the FHF is characterised by high
262 levels of BMP signalling activity at these stages (Klaus et al., 2007; Lescroart et al., 2018).
263 Progenitors derived from the early E6+21h *T*-expressing cells contributed to anterior
264 structures, distinct from the region occupied by late E7+7h *T*-expressing cells (Fig 5A). The
265 descendants from the late E7+7h were located in the posterior mesoderm anterior to the node
266 and were excluded from the FHF (left ventricle progenitors) domain in which BMP activity is
267 high, and also from the cranial paraxial mesoderm with no detectable BMP activity (Fig
268 5Aiii-v and S3 Fig).

269 To ask whether molecular differences exist between the right ventricle, outtflow and atrial
270 cardiac progenitors, we performed single-cell transcriptome analysis of four $T^{nEGP\text{-}CreERT2/+}$
271 $R26R^{tdtomato/tdtomato}$ embryos dissected 7 hours after tamoxifen administration at E7+7h (early
272 head fold -EHF- stage, Fig 5B). The SHF is partitioned into distinct anterior and posterior
273 regions (AHF and pSHF). We reasoned we could use the presence or absence of tdtomato
274 transcripts to discriminate the transcriptome of the tdtomato-positive outflow tract and atrial
275 progenitors from the tdtomato-negative right ventricle progenitors within the AHF and pSHF.
276 We obtained 3494 high-quality single cell transcriptomes. We clustered cells into
277 subpopulations, guided by available single cell transcriptomic data (S4A-C Fig) (Lescroart et
278 al., 2018; Pijuan-Sala et al., 2019). Within these subpopulations, cardiac progenitors
279 belonging to the previously defined FHF, AHF and pSHF. Among the cardiac progenitors,
280 the pSHF had the highest proportion of tdtomato-expressing cells (39.5%) followed by the
281 AHF (7.3%). The FHF had very few tdtomato-positive cells (1%) (S4C Fig).

282 By increasing the resolution of the clustering, we observed the AHF and pSHF clusters could
283 be subdivided into four discernible clusters (Fig 5C and D). Two of these clusters were
284 enriched for tdtomato-positive cells (26% and 46.3%). The two other clusters were depleted
285 of tdtomato-positive cells (5.8% and 4.8%). We hypothesised that the tdtomato positive
286 clusters corresponded to progenitors contributing to the outflow and atria. One of these two
287 clusters was derived from both the AHF and the pSHF while the other cluster was derived
288 from the pSHF with no contribution from the AHF (Fig 5C). Conversely, we reasoned that
289 the tdtomato negative clusters comprised progenitors derived from earlier *T*-expressing cells
290 that contribute to the right ventricle and head mesenchyme. These two clusters derived
291 mostly from the AHF with no contribution from the pSHF. We have summarised these results
292 as a fate map for the different cardiac regions (Fig 7A-B) and we propose that right ventricle,
293 outflow and atrial progenitors form molecularly distinct populations within the AHF and
294 pSHF mirroring their distinct origin in the primitive streak. We describe below the rationale
295 behind the classification of these new subclusters into right ventricle, outflow tract, atria and
296 head mesenchyme progenitor populations based on gene expression patterns (Fig 5E and F).

297 Cells from the outflow tract and atrial clusters expressed *Raldh2* and *Hoxb*1 (Fig 5F).
298 Previous lineage tracing experiments have demonstrated that retinoic acid activated cells
299 contribute to the outflow and atria (Dolle et al., 2010). Additionally, *Hoxb1* positive cells
300 have been shown to contribute to the atria and to the inferior wall of the outflow tract that
301 subsequently forms the myocardium at the base of the pulmonary trunk, in addition to
302 patches of myocardium in the ventricles (Bertrand et al., 2011; Roux et al., 2015). Our data
303 show *Tbx1* and *Raldh2* positive cells formed complementary cell populations within the

304 outflow tract cluster. This is consistent with lineage tracing experiments demonstrating that
305 *Tbx1* expressing cells within the AHF contribute to the outflow tract (Huynh et al., 2007). We
306 also found that *Wnt2* was absent in the outflow tract cluster but expressed in the atrial cluster.
307 This is consistent with *Wnt2* expression in posterior mesoderm in addition to the anterior
308 mesoderm where the prospective cardiac crescent forms (Costello et al., 2011; Monkley et
309 al., 1996).

310 *Fgf8,* a known marker of the AHF (Lescroart et al. 2018), was enriched in both the outflow
311 tract (tdtomato positive) and right ventricle clusters (tdtomato negative). Lineage tracing
312 experiments have shown that *Tbx1* marks part of the right ventricle progenitor domain within
313 the AHF (Baldini et al., 2017; Huynh et al., 2007) while *Fgf8* marks it entirely (Crossley and
314 Martin, 1995; Ilagan et al., 2006). Consistent with these results, we found *Tbx1* to be
315 enriched within a subset of the right ventricle cluster while *Fgf8* is expressed throughout.
316 Also consistent with the hypothesis that these cells correspond to right ventricle progenitors
317 is the enrichment of *Jag1* in this cluster, a regulator of ventricular cardiomyocytes (Evan et
318 al. 2020). *Jag1* is expressed anteriorly, complementary to the region where the cardiac
319 crescent resides (Przemeck et al., 2003). *Dkk1*, a Wnt modulator controlling cardiac
320 differentiation (David et al., 2008; Guo et al., 2019; Phillips et al., 2011), is also uniquely
321 enriched in the right ventricle cluster. *Id1* and *2,* which are essential for the specification of
322 the heart tube-forming progenitors (Cunningham et al., 2017) are enriched in both the left and
323 right ventricle clusters but absent in the outflow and atrial clusters.

324 As mentioned above, the head mesenchyme cluster was depleted of tdtomato-positive cells
325 and may correspond to progenitors mostly derived from early *T*-expressing cells. Notably,

326 *Fgf8* expression was reduced in the head mesenchyme cluster consistent with the absence of
327 *Fgf8* expression in this cell population (Ilagan et al., 2006). However, this cluster was
328 uniquely enriched in transcripts marking the cranial paraxial mesoderm anterior to the *Tbx1*
329 positive domain at E7-E7-5 or the head mesenchyme at E8-8.5 *(Pitx2*, *Alx1*, *Rspo1, Shisa2*)
330 (Campione et al., 1999; Filipe et al., 2006; Nam et al., 2007; Zhao et al., 1996).

331 **The ventricular and outflow tract myocardium are prepatterned within the primitive**
332 **streak.**

333 The signalling environment that cells encounter during migration influences the patterning of
334 the heart fields (Andersen et al., 2018; Evan et al. 2020; Guzzetta et al., 2020; Hochgreb et
335 al., 2003; Klaus et al., 2007; Sirbu et al., 2008). However, our findings that the ventricles and
336 the outflow/atria arise from distinct primitive streak cell populations raise the possibility that
337 these progenitors are already molecularly distinct in the primitive streak. To address this
338 question, we analysed the transcripts of cells from of tamoxifen treated $T^{nEGP\text{-}CreERT2/+}$
339 $R26R^{tdtomato/tdtomato}$ embryos at E6+21h (3635 cells from nine embryos selected at the mid-late
340 streak stages, following tamoxifen administration at E6+5h, Fig 6A) and E7+3h (3994 cells
341 from four embryos selected at the OB-EB stages, following tamoxifen administration at
342 E6+21h, Fig 6A). In the E6+21h mid-late streak stage embryos, tdtomato-positive cells were
343 identified in the extra-embryonic mesoderm, consistent with the idea that the earliest
344 progenitors to emerge from the primitive streak are contributing to the yolk sac blood cells
345 and vasculature endothelium (Kinder et al., 1999). Tdtomato-positive cells were also sparsely
346 found in the embryonic mesoderm. In the E7+3h OB-EB stage embryos, tdtomato-positive

347 cells were found in the embryonic mesoderm where ventricular progenitors reside, and
348 outflow and atrial progenitors are leaving the primitive streak at these stages.

349 We clustered cells guided by available atlases of mouse gastrulation (Pijuan-Sala et al., 2019)
350 (Fig 6B-C and S6A-C Fig*)*. We analysed cells corresponding to the primitive streak and
351 anterior primitive streak (aPS) clusters at the mid-late streak stages and cells in primitive
352 streak cluster at the OB-EB stages (Fig 6E). The transcriptome of the EB-OB primitive streak
353 cells mapped to a subset of the E7.5 primitive streak cells analysed from the mouse
354 gastrulation atlas because it contained a wider range of embryonic stages from the late streak
355 to the LB stages (Pijuan-Sala et al., 2019). The primitive streak cells at the mid-late streak
356 and OB-EB stages did not express the *Cdx1/4* TFs associated with a posterior mesodermal
357 identity. Nor did these cells express *Raldh2*, suggesting that the outflow and atrial progenitors
358 start expressing *Raldh2* once the cells have reached their final location in the embryo (S6E
359 Fig.).

360 Differential gene expression analysis revealed major molecular differences between primitive
361 streak cells contributing to either the ventricles or the outflow/atria (Fig 6E and F). The T-
362 box transcription factor *Eomes* was strongly expressed in the mid-late streak stage primitive
363 streak (ventricular progenitors) while its expression was lower at the OB-EB stages in the
364 outflow/atrial progenitors (Fig 6G). This is in line with the function of *Eomes* in establishing
365 the anterior mesoderm and specifying cardiac progenitors upstream of *Mesp1* (Costello et al.,
366 2011). *T* had the opposite pattern of expression and showed a stronger expression at the OB-
367 EB stages compared to the mid-late streak stages (Fig 6G). Fate mapping in the mouse also
368 identified the anterior primitive streak adjacent to the organiser as a source of cardiac

369 progenitors (Kinder et al., 2001) and *Hes1*, *hHex* (Thomas et al., 1998), *Gsc* (Blum et al.,
370 1992), *Foxa2* (Ang and Rossant, 1994), *Upp1*, *Tdgf1* (Ding et al., 1998; Dono et al., 1993),
371 *Bmp2, Epha2, Lhx1* (Costello et al., 2015; Shimono and Behringer, 1999), *Otx2* (Acampora
372 et al., 2003) and *Chrd* (Bachiller et al., 2000) were among the genes enriched in the anterior
373 primitive streak at the mid-late streak stages (Fig 6E and H). *Sp8*, *Amt*, and *Notch2* (Williams
374 et al., 1995) were preferentially enriched in the primitive streak at the OB-EB stages (Fig 6E
375 and H). *Ccnd2* is expressed in the proximal regions of the primitive streak where the atrial
376 progenitors are located (Wianny et al., 1998) and where BMP signalling activity is high
377 (Morgani et al., 2018). *Mixl1* (Hart *et al.* 2002), *Fst* (Albano et al., 1994), *Gas1* (Lee and Fan,
378 2001), *Frzb* (Hoang et al., 1998) and *Apln* were commonly expressed in the primitive streak
379 at both the mid-late streak and OB-EB stages (Fig 6E). We conclude that ventricular, outflow
380 and atrial progenitors derive from molecularly distinct groups of cells that occupy spatially
381 and temporally discrete regions of the primitive streak.

382 **Discussion**

383 Our findings, summarised in Fig 7A-C, reveal a temporal and spatial order at which different
384 cardiac lineages arise within the primitive streak. The left ventricular progenitors are the first
385 to leave the primitive streak at the mid-streak stage followed shortly after by the right
386 ventricular progenitors at the late-streak stage. Progenitors contributing to the poles of the
387 heart leave the primitive streak at the OB-EB stages. The ventricular and outflow progenitors
388 arise from distal regions of the primitive streak while the atrial progenitors are located in
389 proximal regions of the primitive streak. These different subpopulations constitute

390 molecularly distinct groups of progenitors within the heart field and the transcriptional
391 differences in the primitive streak cells suggest that cardiac progenitors are prepatterned in
392 the primitive streak prior to their migration.

393 ***T* is expressed at low level in the ventricular progenitors and at high level in the atrial**
394 **and outflow tract progenitors.**

395 Upon administration of tamoxifen to $T^{nEGP\text{-}CreERT2/+}$ $R26R^{tdtomato/tdtomato}$ mice, we found tdtomato-
396 expressing cells occupied a maximum of 50.8 percent of the total ventricular myocardial
397 surface area (Figure 1E). The contribution to the poles of the heart instead approached 100
398 percent in some cases. This result raises the possibility that not all ventricular progenitor
399 express *T*, whereas *T* is expressed in all outflow tract and atrial progenitors. Alternatively, all
400 ventricular progenitors might express *T* but at low levels or transiently, compared to the
401 outflow tract and atrial progenitors. This may result in lower levels of cre expression, below
402 the threshold required for reporter recombination in all ventricular progenitors. We found that
403 another T-box transcription factor *Eomes* is strongly expressed in ventricular progenitors in
404 line with other studies also demonstrating *Eomes* function upstream of *Mesp1* to generate
405 anterior mesoderm and specify a cardiac fate (Costello et al., 2011). A similar pulse labelling
406 approach using an *Eomes* inducible line (Pimeisl et al., 2013) would test whether all
407 ventricular progenitors can be traced back to the mid-late streak stages.

408 **Lineage relationships between cardiac progenitors.**

409 The *Foxa2* lineage tracing results demonstrate the existence of a common lineage between
410 the right ventricle and the outflow. This suggests primitive streak cells can divide
411 asymmetrically to generate successively right ventricle and outflow tract progenitors while
412 self-renewing (see model in Fig 7A-C). Live-imaging experiments will provide a more direct
413 indication of the existence of asymmetric cell divisions within the primitive streak. While the
414 ventricular and outflow tract progenitors are located in the anterior/distal primitive streak,
415 atrial progenitors are located in the proximal primitive streak. This spatial segregation
416 suggests that atrial progenitors are likely to constitute a pool of progenitors independent from
417 other cardiac progenitors because of their physical separation. This is consistent with the
418 original clonal analysis from Meilhac et al. (Meilhac et al., 2004). In that study, a
419 retrospective approach based on the random activation of an nLacZ reporter targeted to the
420 cardiac-actin locus (*Actc1*) was adopted. The timing of recombination was estimated based in
421 the size of the clones. By analysing clones of small sizes, Meilhac *et al*. predicted that atrial
422 cells became clonally distinct before the presumptive right ventricle segregate from the
423 outflow (Meilhac et al. 2004). Subsequent clonal experiments tracing *Mesp1* positive
424 progenitors (Devine et al., 2014; Lescroart et al., 2014) also resulted in small clones spanning
425 both the right ventricle and outflow tract compartment while independent clones were found
426 in the atrium (Lescroart et al., 2014). Thus, the atrial progenitors are likely to constitute an
427 independent pool of progenitors in the primitive streak distinct from a common progenitor for
428 right ventricular and outflow tract (see model in Fig 7C).

429 Meilhac et al. also observed two classes of larger clones extending across multiple
430 compartments. This demonstrated the existence of two cardiac lineages in the embryo
431 corresponding to the two heart fields (Meilhac and Buckingham, 2018; Meilhac et al., 2004).
432 Our analysis suggests the segregation of two cardiac lineages could only occur before the
433 onset of gastrulation, for example at the epiblast stage. A further subdivision into progenitors
434 with more restricted lineages then arises by the time epiblast cells ingress through the
435 primitive streak during gastrulation. This is reflected by the existence of progenitor
436 subpopulations within the FHF and SHF corresponding to the prospective left and right
437 ventricle, outflow tract and atria.

438 **A combination of intrinsic factors and inductive events specify and pattern the cardiac**
439 **progenitors.**

440 Depending on the initial timing and site of ingression through the primitive streak, cardiac
441 progenitors will be exposed to different signaling environments, which may influence their
442 cardiac fate. BMPs and Fgfs secreted by the anterior endoderm (Alsan and Schultheiss, 2002;
443 de Soysa et al., 2019; Lough and Sugi, 2000; Schultheiss et al., 1997), retinoic acid expressed
444 in the posterior lateral plate mesoderm (Hochgreb et al., 2003; Xavier-Neto et al., 1999) and
445 signals expressed in the anterior intestinal portal (AIP) at later stages (Anderson et al., 2016)
446 all contribute to the patterning of the cardiac progenitors. A Hedgehog-Fgf signaling axis has
447 also been recently proposed to pattern the anterior mesoderm to allocate the head and heart
448 during mesodermal migration (Guzzetta et al., 2020). Our analysis shows the primitive streak
449 cells that contribute to either the ventricles or to the poles of the heart are molecularly distinct

450 prior to migration. For example, changes in the transcriptional profile of primitive streak
451 cells, including the downregulation of *Foxa2* expression, accompany the transition from
452 contributing to the right ventricle to supplying the outflow tract. This mode of cellular
453 diversification is reminiscent of temporal programmes in other tissues, such as neural
454 progenitors (Doe, 2017), where sequential expression of distinct sets of transcription factors
455 in progenitors defines the later differentiated cell types. These findings raise the possibility
456 that cardiac progenitors are specified via a combination of both initial conditions set in the
457 primitive streak and inductive events happening during migration (Evan et al., 2020;
458 Guzzetta et al., 2020) or once cells have reached their final location in the embryo.

459 Our analysis could not resolve putative differences between left and right progenitors in the
460 primitive streak, and a more precise staging of embryos separating the mid from the late
461 streak stages within a single cell transcriptomic assay will be required for this.
462 Transcriptional differences between left and right ventricular progenitors may be established
463 during migration by exposure to BMP and FGF signaling, both of which are known to
464 mediate medio-lateral patterning of the mesoderm (Dorey and Amaya, 2010; Ilagan et al.,
465 2006; Kelly et al., 2001; Tuazon and Mullins, 2015).

466 The observation that initial molecular differences in the primitive streak lead to the formation
467 of groups of mesodermal cells that generate distinct populations of cardiac progenitors might
468 inform *in vitro* cardiac differentiation protocols. We show the atria are located in the
469 proximal primitive streak where BMP signaling marked by P-Smad1/5/8 is high (Morgani et
470 al., 2018). By contrast, ventricular progenitors originate in anterior primitive streak regions

471 where cells are exposed to high Nodal signaling. This is consistent with methods to generate
472 cardiac cells from human ESCs in vitro that rely on the exposure to different levels of Nodal
473 and BMP signaling (Lee et al., 2017). A higher ratio of BMP4 to activin A signaling is
474 required for the generation of Raldh2 positive mesoderm, which forms atrial cells.
475 Conversely, the generation of ventricular cardiomyocytes relies on a higher level of activin to
476 BMP4 signaling and the formation of Raldh2 negative CD235a positive mesoderm (Lee et
477 al., 2017).

478 We conclude that cardiac progenitors are pre-patterned within the primitive streak both in
479 time and space and this prefigures their allocation to distinct anatomical structures of the
480 heart. Further work will be required to test if generating the correct population of primitive
481 streak cells can help to obtain purer populations of cardiomyocytes from pluripotent stem
482 cells. Our analysis will also help to identify the initial mesodermal population that when
483 dysregulated lead to specific malformations in the heart. For example, left ventricle
484 hypoplasia results from a reduction in the number of specified cardiomyocytes within the
485 mesoderm and this is due to a reduction in the expression of key cardiac transcription factors
486 (Santos-Ledo et al., 2020). Whether such cardiac defects can arise from mutations affecting
487 initial primitive streak cell populations and mesoderm remains to be addressed.

488 # Methods

489 ## Experimental Model and Subject Details

490 All animal procedures were performed in accordance with the Animal (Scientific Procedures)
491 Act 1986 under the UK Home Office project licenses PF59163DB and PIL IA66C8062.

492 ## Mice

493 The *T*$^{nEGFP\text{-}CreERT2/+}$(MGI:5490031) and *Foxa2*$^{nEGP\text{-}CreERT2/+}$ (MGI:5490029) lines (Imuta *et al.*
494 2013) were obtained from Hiroshi Sasaki. *R26* $^{Tomato\ Ai14/\ Tomato\ Ai14}$ (*Gt(ROSA)26Sor*$^{tm14(CAG\text{-}}$
495 $^{tdTomato)Hze}$; MGI:3809524) was obtained from the Jackson Laboratory.

496 ## BRE:H2B-Turquoise BMP reporter line

497 The MLP:H2B-Turquoise sequence (originally from BRE-MLP-H2B-Turquoise pGL3 basic,
498 Briscoe lab, unpublished) was cloned using In-Fusion Cloning (Takara, 639650) between two
499 chicken insulators at the BamHI site of plasmid pJC5-4 (Chung et al., 1993) without the LCR
500 fragment. The 1.6kb MLP:H2B-Turquoise fragment was amplified using Phusion High-
501 Fidelity PCR Master Mix (ThermoFisher Scientific, F532L) according to the manufacturer's
502 instructions. Subsequently the 92bp BRE element was isolated with MluI/XhoI digest from
503 BRE-MLP-H2B-Turquoise pGL3 basic and cloned 5' to the MLP sequence of MluI/XhoI
504 digested MLP:H2B-Turquoise pJC4-4 plasmid. The final plasmid was linearized with NdeI
505 and used for pronuclear injection using fertilized embryos from the F1 hybrid strain

506 (C57BL6/CBA). Mice with BRE reporter activity were verified by genotyping by a
507 commercial vendor (Transnetyx). 5'-3' Forward primer:
508 CACAAGCTGGAGTACAACTACATCAGCGA Reporter 1: TCTATATCACCGCCGAC
509 Reverse Primer: GGCGGATCTTGAAGTTGGCCTTGA. The BRE reporter line was
510 maintained on the F1 background by crossing heterozygous *BRE:H2B-Turquoise* mice to F1
511 wild-type mice.

512 **Lineage tracing of the *T* and *Foxa2* expressing cells**

513 Lineage tracing of *T and Foxa2* mesodermal progenitors was performed by crossing $T^{nEGFP-}$
514 $^{CreERT2/+}$ and $Foxa2^{nEGP-CreERT2/+}$ *with R26* $^{Tomato\ Ai14/\ Tomato\ Ai14}$ mice. To gain better control over
515 embryonic staging, mice were synchronized in estrus by by introducing soiled bedding from
516 a male's cage into the females' three days in advance. On the fourth day, mice were crossed
517 over a two-hours period from 7am to 9am. Vaginal plugs were checked at 9 am and if
518 positive, the embryonic day was defined at E0. This was followed by tamoxifen oral gavage
519 with tamoxifen at 0.08mg/body weight (T5648 SIGMA) dissolved in corn oil at indicated
520 times.

521 **Wholemount immunofluorescence and image acquisition**

522 Embryos were fixed for 20min (for early E6.5-E7.5 embryos) or overnight (for late E12.5
523 hearts) in 2% PFA at 4°C, then permeabilized in PBST (PBS containing 0.5% Triton X-100)
524 for 15min (for early E6.5-E7.5 embryos) or 1 hour (for late E12.5 hearts) and blocked for 5

525 hours (5% donkey serum, Abcam ab138579). Embryos were incubated overnight at 4°C with
526 antibodies diluted in PBST (PBS 0.1% Triton X-100): goat anti-T (1:250, R&D Systems
527 AF2085), rabbit anti-FOXA2 (1:250, Abcam ab108422); mouse anti-cTNT2 (1:250, Thermo
528 Fischer Scientific Systems, MS295P0); rabbit anti-PhosphoSMAD1/5/8 (1:250, Cell
529 Signalling D5B10 Rabbit mAb #13820). After washing in freshly prepared PBST at 4°C,
530 embryos were incubated with secondary antibodies (Molecular Probes) coupled to
531 AlexaFluor 488 or 647 fluorophores as required at 1:250 overnight at 4°C. Before imaging,
532 embryos were washed in PBST at room temperature. Confocal images were obtained on an
533 inverted Zeiss 710 confocal microscope with a 20X water objective (for early E6.5-E7.5
534 embryos) or a 10X air objective (0.4 NA) (for late E12.5 hearts) at a 1024 × 1024 pixels
535 dimension with a z-step of 3-6- µm (2 × 2 tile scale, for the late E12.5 hearts). Embryos were
536 systematically imaged throughout from top to bottom. Images were processed using Fiji
537 software (Schindelin et al., 2012).

538 **Image Analysis**

539 To segment the Foxa2 positive cells a Gaussian filter whose radius is adjusted to the typical
540 size of a cell was first applied with Fiji (Schindelin et al., 2012). The resulting image was
541 next converted to a mask by thresholding. When objects touched each other, a watershed on
542 the binary mask. Finally, particle analyser generated a binary image with objects outside
543 specified size (50) and circularity (0.75-1.00) removed. This process was repeated in each
544 optical z-section of the z-stack. To quantify the surface area of the tdtomato labelled

545 myocardial cells, patches of tdtomato cells were manually segmented and area measured with
546 Fiji (Schindelin et al., 2012) for each individual heart within each litter.

547 **PCR analysis**

548 Primers were designed to span the floxed SV40 poly(A) signal sequence removed from the
549 genome following Cre recombinase-mediated recombination: 5'-3' Forward primer:
550 CGTGCTGGTTATTGTGCTGT; Reverse: CATGAACTCTTTGATGACCTCCTCGC.
551 Primers yield a 1145 bp product from unrecombined DNA and a 274 bp product following
552 recombination. gDNA was extracted using the HotSHOT method (Truett *et al*., 2000) from
553 either the ear clip of an adult *R26* $^{Tomato\ Ai14/\ Tomato\ Ai14}$ mouse (unrecombined) or E9.5 embryos
554 dissected from a pregnant $T^{nEGFP\text{-}CreERT2/+}$ *R26* $^{Tomato\ Ai14/\ Tomato\ Ai14}$ mouse, at 2, 4 and 12 hrs
555 following oral gavage with Tamoxifen at 0.08mg/body weight (T5648 SIGMA) dissolved in
556 corn oil. gDNA was then used in Q5 High-Fidelity 2X Master Mix (NEB) PCR reactions
557 according to the manufacturer's instructions. Amplicons were resolved on a 2% agarose gel
558 with a 100 bp ladder.

559 **Sample preparation for single cell RNA sequencing**

560 $T^{nEGFP\text{-}CreERT2/+}$ *R26* $^{Tomato\ Ai14/\ Tomato\ Ai14}$ mouse embryos were imaged prior to dissociation with
561 a Leica stereo fluorescent microscope with an exposure of ~ 1mn to reveal the tdtomato
562 signal. Sample preparation was done using previously established method (Delile et al.
563 2019). Mouse embryos were dissected in Hanks Balanced Solution without calcium and
564 magnesium (HBSS, Life Technologies, 14185045) supplemented with 5% heat- inactivated

565 foetal bovine serum (FBS). The samples were then incubated on FACSmax cell dissociation
566 solution (Amsbio, T200100) with 10× Papain (30 U/mg, Sigma-Aldrich, 10108014001) for
567 11 min at 37°C to dissociate the cells. To generate a single cell suspension, samples were
568 transferred to HBSS, with 5% FBS, rock inhibitor (10 μM, Stemcell Technologies, Y-27632)
569 and 1× non-essential amino acids (Thermo Fisher Scientific, 11140035), disaggregated
570 through pipetting, and filtered once through 0.35 μm filters and once through 0.20 μm
571 strainers (Miltenyi Biotech, 130-101-812). Quality control was assayed by measuring live
572 cells versus cell death, cell size and number of clumps and 10,000 cells per sample were
573 loaded for sequencing.

574 **Analysis of scRNA-seq data**

575 A suspension of 10,000 single cells was loaded onto the 10x Genomics Single Cell 3′ Chip,
576 and cDNA synthesis and library construction were performed as per the manufacturer's
577 protocol for the Chromium Single Cell 3′ v2 protocol (10x Genomics; PN-120233) and
578 sequenced on an Hiseq4000 (Illumina). 10x CellRanger (version 3.0.2) was used to generate
579 single cell count data for each sample using a custom transcriptome built from the Ensembl
580 mouse GRCm38 release 86 with the addition of the sequence from tdtomato Ai9 plasmid
581 (Madisen et al., 2010). Due to 10x's poly A bias, an additional 225 bases between the
582 tdtomato gene stop codon and the bGH poly(A) signal was included to represent the WPRE
583 gene. All subsequent analyses were performed with R (v.3.6.1) (R Core Team (2013)) using
584 the Seurat (v3) package (Stuart et al., 2019). Primary filtering was performed on each dataset
585 by removing from consideration: cells expressing unique feature counts fewer than 500,

586 number of Unique Molecular Identifiers (UMIs) fewer than 1000 and cells for which
587 mitochondrial genes made up greater than 3 times the standard deviation value of all
588 expressed genes. Each dataset was normalised using the 'LogNormalize' function, with a
589 scale factor of 10,000. The top 2000 highly variable genes were found using the
590 'FindVariableGenes' function and the data centred and scaled using the 'ScaleData' function.
591 PCA decomposition was performed and after consideration of the eigenvalue 'elbow-plots',
592 the first 50 components were used to construct Uniform Manifold Approximation Projection
593 (UMAP) plots. The resolution and number of clusters per experiment was guided by
594 expression of known marker genes (Pijuan-Sala et al., 2019; Lescroart et al., 2018). Clusters
595 relating to surface ectoderm, epiblast, caudal epiblast, primitive streak, intermediate
596 mesoderm, caudal and presomitic mesoderm, pSHF, AHF, FHF, ex-mesoderm and allantois
597 were identified. Cluster specific gene markers were identified using a Wilcoxon rank sum test
598 and a selection of the top genes ranked were used to generate the heatmaps. The clustering
599 analysis was then repeated by increasing the resolution parameter and discernible clusters
600 with distinct marker genes corresponding to the left and right ventricles, head mesenchyme,
601 outflow tract and atria were identified. Samples E6+21h and E7+3h were integrated using
602 Seurat standard integration workflow. Common anchors features between both experiments
603 were selected using the 'FindIntegrationAnchors' functions, using the first 30 dimensions.
604 Clusters were labelled and grouped using pre-existing cell markers and differential
605 expression between clusters were determined using the FindMarkers Function, without logfc
606 or minimum percentage expressed thresholds. Genes showing the largest logFC and smallest
607 adjusted p value between clusters were used to generate heatmaps.

608 **Data availability**

609 Single cell RNA sequencing data have been deposited in NCBI under the accession number
610 GSE153789.

611 ## Acknowledgements

612 The authors would like to thank the Science Technology Platforms at the Francis Crick
613 Institute. In particular, we thank the Advanced Light Microscopy facility, the Advanced
614 Sequencing Facility, Bioinformatics and Biostatistics Facility and the Biological Research
615 Facility for their ongoing support and access to equipment. We are grateful to Robert
616 Goldstone and Amelia Edwards for excellent support with single cell sequencing, Teresa
617 Rayon for assistance with single cell preparation, Gavin Kelly for support with single cell
618 data processing and Joe Brock for research illustration. We thank Teresa Rayon, Florencia
619 Cavodeassi and Peter Scambler for comments on the manuscript and members of the Smith
620 and Briscoe lab for useful discussion. K.I. has received funding from HFSP LTF
621 (LT000609/2015-L). Work in the J.C.S. and J.B.'s lab was supported by the Francis Crick
622 Institute, which receives its core funding from Cancer Research UK (FC001-157, FC001-
623 051) the UK Medical Research Council (FC001-157, FC001-051), and the Welcome Trust
624 (FC001-157, FC001-051). Work in the J.B.'s lab was also supported by the the European
625 Research Council under European Union (EU) Horizon 2020 research and innovation
626 program grant 742138.

627 ## Author contributions

628 K.I.: Conceptualization, Investigation, Formal analysis, Visualisation, Writing – Original
629 Draft Preparation. P.S.B. and P.C.: Formal analysis, J.D.: Data curation, N.M.G., D.S.:
630 Investigation, R.A.J.: Investigation, Visualisation, J.C.S.: Funding acquisition, Supervision,
631 Writing-Review & Editing, J.B.: Funding acquisition, Conceptualization, Supervision,
632 Writing-Review & Editing.

633 ## References


634 Acampora, D., Annino, A., Puelles, E., Alfano, I., Tuorto, F., and Simeone, A. (2003). OTX1
635 compensates for OTX2 requirement in regionalisation of anterior neuroectoderm. Gene Expr
636 Patterns *3*, 497-501.

637 Albano, R.M., Arkell, R., Beddington, R.S., and Smith, J.C. (1994). Expression of inhibin
638 subunits and follistatin during postimplantation mouse development: decidual expression of
639 activin and expression of follistatin in primitive streak, somites and hindbrain. Development
640 *120*, 803-813.

641 Alsan, B.H., and Schultheiss, T.M. (2002). Regulation of avian cardiogenesis by Fgf8
642 signaling. Development *129*, 1935-1943.

643 Andersen, P., Tampakakis, E., Jimenez, D.V., Kannan, S., Miyamoto, M., Shin, H.K., Saberi,
644 A., Murphy, S., Sulistio, E., Chelko, S.P., *et al*. (2018). Precardiac organoids form two heart
645 fields via Bmp/Wnt signaling. Nat Commun *9*, 3140.

646 Anderson, C., Khan, M.A.F., Wong, F., Solovieva, T., Oliveira, N.M.M., Baldock, R.A.,
647 Tickle, C., Burt, D.W., and Stern, C.D. (2016). A strategy to discover new organizers
648 identifies a putative heart organizer. Nat Commun *7*, 12656.

649 Ang, S.L., and Rossant, J. (1994). HNF-3 beta is essential for node and notochord formation
650 in mouse development. Cell *78*, 561-574.

651 Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and
652 axis patterning in the early mouse embryo. Nat Rev Mol Cell Biol *10*, 91-103.

653 Bachiller, D., Klingensmith, J., Kemp, C., Belo, J.A., Anderson, R.M., May, S.R., McMahon,
654 J.A., McMahon, A.P., Harland, R.M., Rossant, J., *et al*. (2000). The organizer factors
655 Chordin and Noggin are required for mouse forebrain development. Nature *403*, 658-661.

656 Baldini, A., Fulcoli, F.G., and Illingworth, E. (2017). Tbx1: Transcriptional and
657 Developmental Functions. Curr Top Dev Biol *122*, 223-243.

658 Bardot, E., Calderon, D., Santoriello, F., Han, S., Cheung, K., Jadhav, B., Burtscher, I.,
659 Artap, S., Jain, R., Epstein, J., *et al*. (2017). Foxa2 identifies a cardiac progenitor population
660 with ventricular differentiation potential. Nat Commun *8*, 14428.

661 Bertrand, N., Roux, M., Ryckebusch, L., Niederreither, K., Dolle, P., Moon, A., Capecchi,
662 M., and Zaffran, S. (2011). Hox genes define distinct progenitor sub-domains within the
663 second heart field. Dev Biol *353*, 266-274.

664 Blum, M., Gaunt, S.J., Cho, K.W., Steinbeisser, H., Blumberg, B., Bittner, D., and De
665 Robertis, E.M. (1992). Gastrulation in the mouse: the role of the homeobox gene goosecoid.
666 Cell *69*, 1097-1106.

667 Burtscher, I., and Lickert, H. (2009). Foxa2 regulates polarity and epithelialization in the
668 endoderm germ layer of the mouse embryo. Development *136*, 1029-1038.

669 Campione, M., Steinbeisser, H., Schweickert, A., Deissler, K., van Bebber, F., Lowe, L.A.,
670 Nowotschin, S., Viebahn, C., Haffter, P., Kuehn, M.R., *et al*. (1999). The homeobox gene
671 Pitx2: mediator of asymmetric left-right signaling in vertebrate heart and gut looping.
672 Development *126*, 1225-1234.

673 Chung, J.H., Whiteley, M., and Felsenfeld, G. (1993). A 5' element of the chicken beta-
674 globin domain serves as an insulator in human erythroid cells and protects against position
675 effect in Drosophila. Cell *74*, 505-514.

676 Costello, I., Nowotschin, S., Sun, X., Mould, A.W., Hadjantonakis, A.K., Bikoff, E.K., and
677 Robertson, E.J. (2015). Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern
678 anterior mesendoderm, node, and midline development. Genes Dev *29*, 2108-2122.

679 Costello, I., Pimeisl, I.M., Drager, S., Bikoff, E.K., Robertson, E.J., and Arnold, S.J. (2011).
680 The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac
681 mesoderm during mouse gastrulation. Nat Cell Biol *13*, 1084-1091.

682 Crossley, P.H., and Martin, G.R. (1995). The mouse Fgf8 gene encodes a family of
683 polypeptides and is expressed in regions that direct outgrowth and patterning in the
684 developing embryo. Development *121*, 439-451.

685 Cunningham, T.J., Yu, M.S., McKeithan, W.L., Spiering, S., Carrette, F., Huang, C.T.,
686 Bushway, P.J., Tierney, M., Albini, S., Giacca, M.*, et al*. (2017). Id genes are essential for
687 early heart formation. Genes Dev *31*, 1325-1338.

688 David, R., Brenner, C., Stieber, J., Schwarz, F., Brunner, S., Vollmer, M., Mentele, E.,
689 Muller-Hocker, J., Kitajima, S., Lickert, H.*, et al*. (2008). MesP1 drives vertebrate
690 cardiovascular differentiation through Dkk-1-mediated blockade of Wnt-signalling. Nat Cell
691 Biol *10*, 338-345.

692 Delile, J., Rayon, T., Melchionda, M., Edwards, A., Briscoe, J., and Sagner, A. (2019). Single
693 cell transcriptomics reveals spatial and temporal dynamics of gene expression in the
694 developing mouse spinal cord. Development *146*.

695 de Soysa, T.Y., Ranade, S.S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H.T.,
696 Schricker, A., Del Sol, A., Gifford, C.A., and Srivastava, D. (2019). Single-cell analysis of
697 cardiogenesis reveals basis for organ-level developmental defects. Nature *572*, 120-124.

698 Devine, W.P., Wythe, J.D., George, M., Koshiba-Takeuchi, K., and Bruneau, B.G. (2014).
699 Early patterning and specification of cardiac progenitors in gastrulating mesoderm. Elife
700 2014;3:e03848

701 Ding, J., Yang, L., Yan, Y.T., Chen, A., Desai, N., Wynshaw-Boris, A., and Shen, M.M.
702 (1998). Cripto is required for correct orientation of the anterior-posterior axis in the mouse
703 embryo. Nature *395*, 702-707.

704 Doe, C.Q. (2017). Temporal Patterning in the Drosophila CNS. Annu Rev Cell Dev Biol *33*,
705 219-240.

706 Dolle, P., Fraulob, V., Gallego-Llamas, J., Vermot, J., and Niederreither, K. (2010). Fate of
707 retinoic acid-activated embryonic cell lineages. Dev Dyn *239*, 3260-3274.

708 Dono, R., Scalera, L., Pacifico, F., Acampora, D., Persico, M.G., and Simeone, A. (1993).
709 The murine cripto gene: expression during mesoderm induction and early heart
710 morphogenesis. Development *118*, 1157-1168.

711 Dorey, K., and Amaya, E. (2010). FGF signalling: diverse roles during early vertebrate
712 embryogenesis. Development *137*, 3731-3742.

713 Downs, K.M., and Davies, T. (1993). Staging of gastrulating mouse embryos by
714 morphological landmarks in the dissecting microscope. Development *118*, 1255-1266.

715 Evan S. Bardot, Bharati Jadhav, Nadeera Wickramasinghe, Amélie Rezza, Michael Rendl,
716 Andrew J. Sharp, Nicole C. Dubois. Notch Signalling Commits Mesoderm to the Cardiac
717 Lineage. bioRxiv 2020.02.20.958348; doi: https://doi.org/10.1101/2020.02.20.958348

718 Filipe, M., Goncalves, L., Bento, M., Silva, A.C., and Belo, J.A. (2006). Comparative
719 expression of mouse and chicken Shisa homologues during early development. Dev Dyn *235*,
720 2567-2573.

721 Frank, D.U., Elliott, S.A., Park, E.J., Hammond, J., Saijoh, Y., and Moon, A.M. (2007).
722 System for inducible expression of cre-recombinase from the Foxa2 locus in endoderm,
723 notochord, and floor plate. Dev Dyn *236*, 1085-1092.

724 Garcia-Martinez, V., and Schoenwolf, G.C. (1993). Primitive-streak origin of the
725 cardiovascular system in avian embryos. Dev Biol *159*, 706-719.

726 Guo, Y., Dorn, T., Kuhl, S.J., Linnemann, A., Rothe, M., Pfister, A.S., Vainio, S., Laugwitz,
727 K.L., Moretti, A., and Kuhl, M. (2019). The Wnt inhibitor Dkk1 is required for maintaining
728 the normal cardiac differentiation program in Xenopus laevis. Dev Biol *449*, 1-13.

729 Guzzetta, A., Koska, M., Rowton, M., Sullivan, K.R., Jacobs-Li, J., Kweon, J., Hidalgo, H.,
730 Eckart, H., Hoffmann, A.D., Back, R., *et al*. (2020). Hedgehog-FGF signaling axis patterns
731 anterior mesoderm during gastrulation. Proc Natl Acad Sci U S A.

732 Hoang, B.H., Thomas, J.T., Abdul-Karim, F.W., Correia, K.M., Conlon, R.A., Luyten, F.P.,
733 and Ballock, R.T. (1998). Expression pattern of two Frizzled-related genes, Frzb-1 and Sfrp-
734 1, during mouse embryogenesis suggests a role for modulating action of Wnt family
735 members. Dev Dyn *212*, 364-372.

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

739 Huynh, T., Chen, L., Terrell, P., and Baldini, A. (2007). A fate map of Tbx1 expressing cells
740 reveals heterogeneity in the second cardiac field. Genesis *45*, 470-475.

741 Ilagan, R., Abu-Issa, R., Brown, D., Yang, Y.P., Jiao, K., Schwartz, R.J., Klingensmith, J.,
742 and Meyers, E.N. (2006). Fgf8 is required for anterior heart field development. Development
743 *133*, 2435-2445.

744 Imuta, Y., Kiyonari, H., Jang, C.W., Behringer, R.R., and Sasaki, H. (2013). Generation of
745 knock-in mice that express nuclear enhanced green fluorescent protein and tamoxifen-
746 inducible Cre recombinase in the notochord from Foxa2 and T loci. Genesis *51*, 210-218.

747 Kelly, R.G., Brown, N.A., and Buckingham, M.E. (2001). The arterial pole of the mouse
748 heart forms from Fgf10-expressing cells in pharyngeal mesoderm. Dev Cell *1*, 435-440.

749 Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., and Tam, P.P.
750 (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the
751 anteroposterior axis during gastrulation of the mouse embryo. Development *126*, 4691-4701.

752 Kinder, S.J., Tsang, T.E., Wakamiya, M., Sasaki, H., Behringer, R.R., Nagy, A., and Tam,
753 P.P. (2001). The organizer of the mouse gastrula is composed of a dynamic population of
754 progenitor cells for the axial mesoderm. Development *128*, 3623-3634.

755 Klaus, A., Saga, Y., Taketo, M.M., Tzahor, E., and Birchmeier, W. (2007). Distinct roles of
756 Wnt/beta-catenin and Bmp signaling during early cardiogenesis. Proc Natl Acad Sci U S A
757 *104*, 18531-18536.

758 Lee, C.S., and Fan, C.M. (2001). Embryonic expression patterns of the mouse and chick
759 Gas1 genes. Mech Dev *101*, 293-297.

760 Lee, J.H., Protze, S.I., Laksman, Z., Backx, P.H., and Keller, G.M. (2017). Human
761 Pluripotent Stem Cell-Derived Atrial and Ventricular Cardiomyocytes Develop from Distinct
762 Mesoderm Populations. Cell Stem Cell *21*, 179-194 e174.

763 Lescroart, F., Chabab, S., Lin, X., Rulands, S., Paulissen, C., Rodolosse, A., Auer, H.,
764 Achouri, Y., Dubois, C., Bondue, A., *et al*. (2014). Early lineage restriction in temporally
765 distinct populations of Mesp1 progenitors during mammalian heart development. Nat Cell
766 Biol *16*, 829-840.

767 Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sanchez-Danes, A., Moignard,
768 V., Dubois, C., Paulissen, C., Kinston, S., *et al*. (2018). Defining the earliest step of
769 cardiovascular lineage segregation by single-cell RNA-seq. Science *359*, 1177-1181.

770 Lough, J., and Sugi, Y. (2000). Endoderm and heart development. Dev Dyn *217*, 327-342.

771 Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L.,
772 Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., *et al*. (2010). A robust and high-throughput Cre
773 reporting and characterization system for the whole mouse brain. Nat Neurosci *13*, 133-140.

774 Meilhac, S.M., and Buckingham, M.E. (2018). The deployment of cell lineages that form the
775 mammalian heart. Nat Rev Cardiol *15*, 705-724.

776 Meilhac, S.M., Esner, M., Kelly, R.G., Nicolas, J.F., and Buckingham, M.E. (2004). The
777 clonal origin of myocardial cells in different regions of the embryonic mouse heart. Dev Cell
778 *6*, 685-698.

779 Monkley, S.J., Delaney, S.J., Pennisi, D.J., Christiansen, J.H., and Wainwright, B.J. (1996).
780 Targeted disruption of the Wnt2 gene results in placentation defects. Development *122*,
781 3343-3353.

782 Morgani, S.M., Metzger, J.J., Nichols, J., Siggia, E.D., and Hadjantonakis, A.K. (2018).
783 Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo
784 regionalized cell fate patterning. Elife *7*.

785 Nam, J.S., Turcotte, T.J., and Yoon, J.K. (2007). Dynamic expression of R-spondin family
786 genes in mouse development. Gene Expr Patterns *7*, 306-312.

787 Phillips, M.D., Mukhopadhyay, M., Poscablo, C., and Westphal, H. (2011). Dkk1 and Dkk2
788 regulate epicardial specification during mouse heart development. Int J Cardiol *150*, 186-192.

789 Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J.,
790 Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., *et al*. (2019). A single-cell molecular
791 map of mouse gastrulation and early organogenesis. Nature *566*, 490-495.

792 Pimeisl, I.M., Tanriver, Y., Daza, R.A., Vauti, F., Hevner, R.F., Arnold, H.H., and Arnold,
793 S.J. (2013). Generation and characterization of a tamoxifen-inducible Eomes(CreER) mouse
794 line. Genesis *51*, 725-733.

795 Przemeck, G.K., Heinzmann, U., Beckers, J., and Hrabe de Angelis, M. (2003). Node and
796 midline defects are associated with left-right development in Delta1 mutant embryos.
797 Development *130*, 3-13.

798 Rivera-Perez, J.A., and Magnuson, T. (2005). Primitive streak formation in mice is preceded
799 by localized activation of Brachyury and Wnt3. Dev Biol *288*, 363-371.

800 Roux, M., Laforest, B., Capecchi, M., Bertrand, N., and Zaffran, S. (2015). Hoxb1 regulates
801 proliferation and differentiation of second heart field progenitors in pharyngeal mesoderm
802 and genetically interacts with Hoxa1 during cardiac outflow tract development. Dev Biol
803 *406*, 247-258.

804 Santos-Ledo, A., Washer, S., Dhanaseelan, T., Eley, L., Alqatani, A., Chrystal, P.W.,
805 Papoutsi, T., Henderson, D.J., and Chaudhry, B. (2020). Alternative splicing of jnk1a in

806 zebrafish determines first heart field ventricular cardiomyocyte numbers through modulation
807 of hand2 expression. PLoS Genet *16*, e1008782.

808 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
809 Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., *et al*. (2012). Fiji: an open-source
810 platform for biological-image analysis. Nat Methods *9*, 676-682.

811 Schultheiss, T.M., Burch, J.B., and Lassar, A.B. (1997). A role for bone morphogenetic
812 proteins in the induction of cardiac myogenesis. Genes Dev *11*, 451-462.

813 Shimono, A., and Behringer, R.R. (1999). Isolation of novel cDNAs by subtractions between
814 the anterior mesendoderm of single mouse gastrula stage embryos. Dev Biol *209*, 369-380.

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

817 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y.,
818 Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell
819 Data. Cell *177*, 1888-1902 e1821.

820 Thomas, P.Q., Brown, A., and Beddington, R.S. (1998). Hex: a homeobox gene revealing
821 peri-implantation asymmetry in the mouse embryo and an early transient marker of
822 endothelial cell precursors. Development *125*, 85-94.

823 Truett, G.E., Heeger, P., Mynatt, R.L., Truett, A.A., Walker, J.A., and Warman, M.L. (2000).
824 Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris
825 (HotSHOT). Biotechniques *29*, 52, 54.

826 Tuazon, F.B., and Mullins, M.C. (2015). Temporally coordinated signals progressively
827 pattern the anteroposterior and dorsoventral body axes. Semin Cell Dev Biol *42*, 118-133.

828 Wianny, F., Real, F.X., Mummery, C.L., Van Rooijen, M., Lahti, J., Samarut, J., and
829 Savatier, P. (1998). G1-phase regulators, cyclin D1, cyclin D2, and cyclin D3: up-regulation
830 at gastrulation and dynamic expression during neurulation. Dev Dyn *212*, 49-62.

831 Williams, R., Lendahl, U., and Lardelli, M. (1995). Complementary and combinatorial
832 patterns of Notch gene family expression during early mouse development. Mech Dev *53*,
833 357-368.

834 Xavier-Neto, J., Neville, C.M., Shapiro, M.D., Houghton, L., Wang, G.F., Nikovits, W., Jr.,
835 Stockdale, F.E., and Rosenthal, N. (1999). A retinoic acid-inducible transgenic marker of
836 sino-atrial development in the mouse heart. Development *126*, 2677-2687.

837 Yutzey, K.E., and Bader, D. (1995). Diversification of cardiomyogenic cell lineages during
838 early heart development. Circ Res *77*, 216-219.

839 Zhao, Q., Behringer, R.R., and de Crombrugghe, B. (1996). Prenatal folic acid treatment
840 suppresses acrania and meroanencephaly in mice mutant for the Cart1 homeobox gene. Nat
841 Genet *13*, 275-283.

842 **Figure legends**

843 **Fig 1. Genetic tracing of the *T+* primitive steak cells. (A)** Schematics of the $T^{nGPF\text{-}CreERT2/+}$;
844 $R26R^{tdtomato/+}$ allele (Imuta et al., 2013). Cre-ERT2 is expressed in cells expressing *T*. In the
845 presence of tamoxifen, Cre protein is translocated to the nucleus where it recombines the
846 $R26R^{tdtomato/+}$ reporter. As a result, the cell and its descendants are permanently labeled. **(B)**
847 Diagram of the experimental approach. *T*-expressing cells and their descendants are labelled,
848 from E6+21h, E7, E7+7h and E7+11h by administrating a dose of tamoxifen (Tam) to $T^{nGPF\text{-}}$
849 $^{CreERT2/+}$; $R26R^{tdtomato/+}$ mice (0.08mg/body weight via oral gavage). Cell descendants in the
850 myocardium are analysed at E12.5. **(C)** Representative hearts resulting from the
851 administration of tamoxifen at different time points in $T^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/+}$
852 immunostained with cTnnT to reveal the cardiomyoyctes (blue). Yellow arrows identify
853 small patches of tdtomato positive cardiomyocytes in the LV (ii and iv), in the RV (v) and in

854 the outflow tract and atria (vii and viii). Views are ventral. Single epicardial cells are labeled
855 in each condition. **(D-E)** Summary of all $T^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/+}$ hearts examined. The
856 contribution of the *T*-expressing cells to the different compartments of the heart is quantified
857 by measuring the proportion of tdtomato positive myocardium. Numbers in brackets in (E)
858 represent the number of litters assessed. Error bars are SD. **(F)** Stage variation quantified
859 according to Downs and Davies criteria (Downs and Davies, 1993). Number in brackets
860 represents the number of litters assessed. All the timed matings were for 2-hour periods. (**G**)
861 PCR amplicons generated from the genomic region in which Cre-mediated recombination
862 occurs, resolved on an agarose gel. Before recombination, the PCR product is 1145 bp (white
863 rectangle); after recombination it is 274 bp (black rectangle). Template gDNA was extracted
864 from either an ear clip of an adult $T^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/tdtomato}$ mouse (untreated) or $T^{nGPF\text{-}}$
865 $^{CreERT2/+}$; $R26R^{tdtomato/tdtomato}$ embryos dissected at 2, 4 and 12 hrs following oral gavage with
866 Tamoxifen, as labelled. An increase in the proportion of the recombined band can be seen
867 over time following Tamoxifen administration. (**H**) Embryos collected at E6+212h or E7+7h
868 showing variation in stage. LV: left ventricle, RV: right ventricle, OFT: outflow tract, LA:
869 left atria, RA: right atria, cTnnT: cardiac troponinin T. MS: Mid-streak LS: late-streak, OB:
870 "no bud" stage, EB: "early bud" stage, LB: "late bud stage", EHF: early head fold. cTnnT:
871 cardiac Troponin T. Scale bar: 200 $\mu$m.

872 **Fig 2. Independent *Foxa2+* progenitors contribute to the ventricles and outflow tract.**
873 **A)** Schematics of the $Foxa2^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/+}$ allele (Imuta et al., 2013). **(B)** Pulse-
874 labelling of the *Foxa2*-expressing cells and their descendants, from stage (i) E6+21h, (ii)
875 E7+7h by administration of tamoxifen to $Foxa2^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/+}$ mice (0.08mg/body

876 weight via oral gavage). Cells descendants in the heart analysed at E12.5. **(C)** Representative
877 hearts resulting from the administration of tamoxifen at different time points in *Foxa2*$^{nGPF-}$
878 $^{CreERT2/+}$; *R26R*$^{tdtomato/+}$ immunostained for cTnnT to reveal the cardiomyocytes (blue). Views
879 are ventral. Number in brackets represents the number of litters assessed. Single epicardial
880 cells are also labelled in iii. **(D)** Summary of all hearts analysed. Scale bar: 200 $\mu$m.

881 **Fig 3. T and Foxa2 colocalise in primitive streak cells**. Representative embryos of E6+21h
882 MS (A-B) and LS (C-D) and E7+7h "early bud" EB (E-F). Embryos are immunostained for T
883 (red) and Foxa2 (green). Images are Z-maximum projections. Views are lateral/slightly
884 posterior to show the whole width of the primitive streak. Insets in Bi-iii, Di-iii and Fi-iii
885 show magnified views and white arrows point to T+/Foxa2+ double positive cells. Scale bar:
886 100 $\mu$m.

887 **Fig 4. Ventricular precursors originate from distal regions of the primitive streak. (A)**
888 Localisation of the tdtomato+ cell from the *Foxa2* lineage are assessed at (i) E7+7h. The
889 embryos are immunostained for T (red) and Foxa2 (green). Image is a z-maximum
890 projection. (ii) The same embryo (i), but with Foxa2+ cells masked to reveal the
891 tdtomato+/Foxa2- cells. Insets (iii) and (iv) show magnified views in single optical sections
892 of (i). Red arrows point to tdtomato+/Foxa2-/T- mesodermal cells. Yellow arrows point to
893 tdtomato+/Foxa2-/T+ primitive streak cells. Green arrows point to tdtomato+/Foxa2+/T-
894 endodermal cells. No tdtomato+/Foxa2+/T+ cells are identifiable. **(B)** (i) A common
895 progenitor between the RV and the outflow tract in the distal primitive streak. *Foxa2* is
896 downregulated as it switches its contribution toward outflow tract myocardium. (ii) Proximal
897 PS contribute to the atria and are descended from cells that had not expressed *Foxa2* in their

898 past. PS primitive streak, pro. PS: proximal primitive streak. Ant: anterior, Post: posterior.
899 Scale bar: 100 $\mu$m

900 **Fig 5. Subclusters corresponding to the right ventricle, outflow tract and atrial**
901 **progenitors can be defined within the AHF and pSHF.** (**A**) Tdtomato localisation in *T*$^{nGPF-}$
902 $^{CreERT2/+}$*; R26R*$^{tdtomato/tdtomato}$*, Bre-Cerulean* embryos at the EHF stage and following tamoxifen
903 administration at E6+21h (i-ii) and at E7+7h (iii-v). (**B**) The four E7+7 h *T*$^{nGPF-CreERT2/+}$*;*
904 *R26R*$^{tdtomato/tdtomato}$ embryos analysed in the single cell transcriptomic assay. (**C**) Diagram
905 showing the repartition of the tdtomato positive cells from the FHF, AHF and pSHF clusters
906 into the left and right ventricle, head mesenchyme, outflow and Atria clusters. (**D**) UMAP
907 plot of the FHF, AHF and pSHF and subpopulations corresponding to progenitors
908 contributing to the left and right ventricle, Head Mesenchyme, outflow tract and atria in
909 addition to the Pharyngeal arches. (**E**) Expression heat map of marker genes (source data
910 Figure 5) and tdtomato. Dotted square on the left handside indicates genes associated with the
911 FHF, AHF or pSHF (Figure 5-Figure supplement 2). Scale indicates z-scored expression
912 values. (**F**) UMAP showing the log normalised counts of selected genes. n.: node, PS:
913 Primitive streak, FHF: first heart field, SHF: Second heart field, Hm; Head mesenchyme,
914 OFT: outflow tract, Pha: Pahryngeal arches, AHF: anterior heart field, pSHF: posterior
915 second heart field. Scale bar: 100 $\mu$m.

916 **Fig 6. The molecular signature of the primitive streak cells contributing to the ventricles**
917 **or outflow/atria are distinct.** (**A**) Mid-late streak and OB/EB *T*$^{nGPF-CreERT2/+}$*; R26R*$^{tdtomato/tdtomato}$
918 embryos analysed in the sc-RNA seq. (**B-C**) Expression heat map of marker genes (source
919 data Figure 6). Scale indicates z-scored expression values. (**D**) UMAP plot showing the

920 integrated data from the two scRNA-seq mid-late streak and OB/EB datasets (see also Figure
921 6–figure supplement 1). Colour code correspond to the embryonic stage of collection or
922 population identity. (**E**) Expression heat map of marker genes comparing the anterior
923 primitive streak and primitive streak cells at the mid-late streak stages and primitive streak
924 cells at the OB-EB stages (source data Figure 6). Scale indicates z-scored expression values.
925 (**F**) UMAP plot of the anterior primitive streak, (aPS), mid-late streak (MS-LS) primitive
926 streak and OB-EB primitive streak cells color coded. (**G-H**) Violin plots showing the
927 normalized log2 expression value of *Eomes* and *T* (G) and *Foxa2*, *Lhx1*, *Sp8 Hes1*, *Ccnd2*
928 and *Notch2* (H) in the anterior primitive streak (aPS) and primitive streak at mid-late streak
929 and OB-EB stages. DE: definitive endoderm, aPS: anterior primitive streak, PS: primitive
930 streak, Nascent meso: nascent mesoderm, LPM/Ex-meso: lateral plate mesoderm and Extra-
931 embryonic mesoderm, mesenchyme. PGC: primordial germs cells. MS: mid-streak, LS: late
932 streak, OB: no bud, EB: early bud.

933 **Fig 7. Model of early cardiac development** (**A**) Cells located in distal regions of the
934 primitive streak contribute first to the left ventricle (mid-streak stage), then to the right
935 ventricle (late-streak stage) and finally to the outflow tract (OB-EB stages). Although the
936 outflow and atria leave the primitive streak at similar stages they arise from different regions.
937 The outflow tract originates from distal locations in the primitive streak while atrial
938 progenitors are positioned more proximally. The distal primitive streak cells express *Foxa2*
939 when they contribute to the ventricles. They stop expressing *Foxa2* when they contribute to
940 the outflow tract. (**B**) Proposed location of the FHF, AHF and pSHF in comparison to the
941 hypothetical fate of these different cardiac regions at the pre-headfold stage in (A). (**C**)

942 Lineage tree of the primitive streak. Within the distal end of the primitive streak, independent
943 primitive streak cells contribute to the left ventricle, the right ventricle, the (at the mid-late
944 streak stage) and to the outflow tract (OB-EB stages). A common progenitor between the
945 right ventricle and the outflow tract exists. By contrast, atrial progenitors are located in the
946 proximal primitive streak. LV: left ventricle, RV: right ventricle, OFT: outflow tract, Pha:
947 Pharyngeal arches, hm: head mesenchyme, FHF: first heart field, AHF: anterior heart field,
948 pSHF: posterior second heart field. OB: no bud stage. EB: early bud stage.

949 **Supporting information**

950 **S1 Fig. Tamoxifen activity persists for at least 24 hours when administrated at a high**
951 **dose by oral gavage. (A)** The administration of a high dose of tamoxifen (0.08mg/bw by oral
952 gavage) at E5 in $T^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/+}$ mice leads to the presence of tdtomato positive-
953 cells in mesoderm derivatives including cardiomyocytes (see yellow arrow in inset), head
954 mesenchyme (red arrow), endothelium (green arrow) and allantois. hm: head mesoderm, ht:
955 heart tube, a: allantois. Mouse were mated for a two-hour period. Scale bar: 100 $\mu$m

956 **S2 Fig. *Foxa2*-expressing cells contribute to the outflow tract myocardium but not to**
957 **the atria. A)** Heart resulting from the administration of tamoxifen at E6+21h. View is
958 ventral. Tdtomato-positive cardiomyocytes are absent from the myocardium in the atria,
959 however, contribution to the epicardium (yellow arrow) and myocardium (yellow arrows) in
960 the ventricle and outflow tract is visible. (**B**) E8 embryo resulting from the administration of
961 tamoxifen at E6+21h in $Foxa2^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/+}$ mouse and immunostained for
962 Foxa2 (green). Tdtomato-positive cells are localised in the pericardium, cardiomyocytes and

963 endoderm but not in the endocardium. OFT: outflow tract, LA: left atria, RA: right atria, LV:
964 left ventricle, RV: right ventricle, CC: cardiac crescent, cardio: cardiomyocyte, endo:
965 endoderm, Scale bars: 200 $\mu$m in (A) and 100 $\mu$m in (B).

966 **S3 Fig. Outflow tract and atrial progenitors are located away from regions with high**
967 **BMP signalling activity. (A)** Tdtomato localisation in $T^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/tdtomato}$
968 embryos immunostained against P-Smad1/5/8 following tamoxifen administration at E7+7h.
969 Scale bar: 100 $\mu$m.

970 **S4 Fig. Identification of the heart fields by scRNA-seq. (A)** UMAP plot showing the
971 ectodermal, epiblast and mesodermal cell populations coloured by cluster identity. (**B**)
972 UMAP showing the log normalised counts of selected genes. (**C**) Expression heat map of
973 marker genes (source data Figure 5) and tdtomato. Scale indicate z-scored expression values.
974 PS: primitive streak, ecto: surface ectoderm, inter-m: intermediate mesoderm, caudal-m:
975 caudal mesoderm, pSHF: posterior second heart field, AHF: anterior heart field, FHF: first
976 heart field, ex-m; extra embryonic mesoderm and mesenchyme, al.: allantois.

977 **S5 Fig. Bre-cerulean line report BMP signalling activity in the mesoderm.** (**A-A''**)
978 Colocalisation of the Cerulean signal and P-Smad1/5/8 in Bre-cerulean embryos at the
979 cardiac crescent stage. (A) z-max proj. (A') Single optical projection. (A'') Magnified view
980 form insets in A'. cc.: cardiac crescent. e: endoderm. p: pericardium. Scale bar: 100 $\mu$m

981 **S6 Fig. scRNA-seq analysis of the mid-late streak and OB/EB stages. (A** and **C)** UMAP
982 plot coloured by cluster identity from sc-RNA seq analysis of $T^{nGPF\text{-}CreERT2/+}$; $R26R^{tdtomato/tdtomato}$
983 embryos at the E6+21h, MS-LS stages (A) and at the E7+3h, OB-EB stages (C). (**B** and **D**)

UMAP showing the log normalised counts of selected genes. Colour intensity is proportional to the expression level of a given gene. **(E)** Violin plots showing the normalized log2 expression value of *Radlh2*, *Cdx1* and *Cdx4*. DE: definitive endoderm, aPS: anterior primitive streak, PS: primitive streak, Nascent meso: nascent mesoderm, LPM/Ex-meso: lateral plate mesoderm and Extra-embryonic mesoderm, mesenchyme. PGC: primordial germs cells. MS: mid-streak, LS: late streak, OB: no bud, EB: early bud.

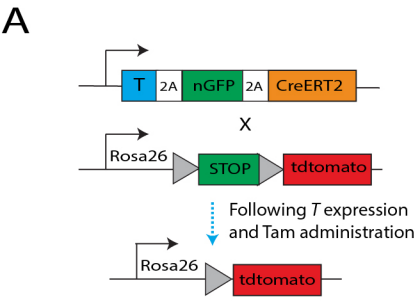
**S1_raw image.** Related to Fig 1G

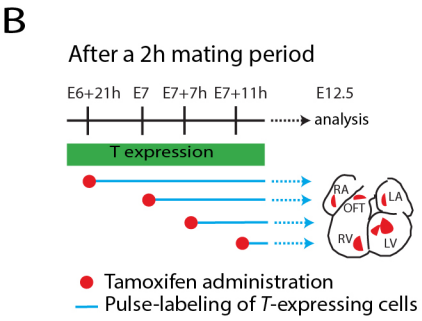
**S2_ source data.** Quantification of labelled surface area related to Fig1E

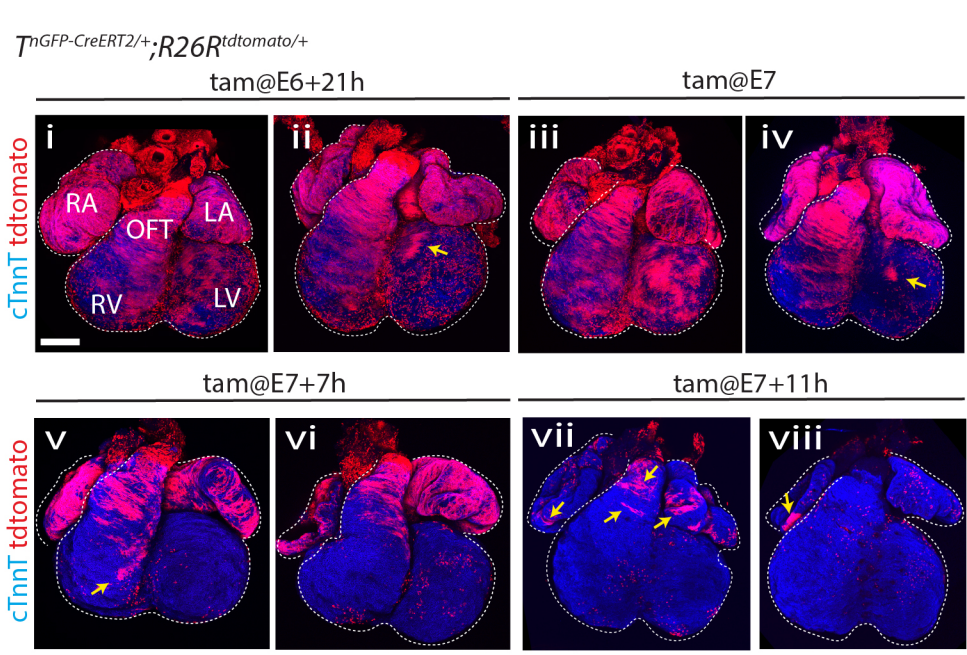
**S3_source data.** Single cell RNA sequencing data related to Fig. 5 and S4 Fig.

**S4_source data.** Single cell RNA sequencing data related to Fig 6 and S6 Fig.

# Figure 1



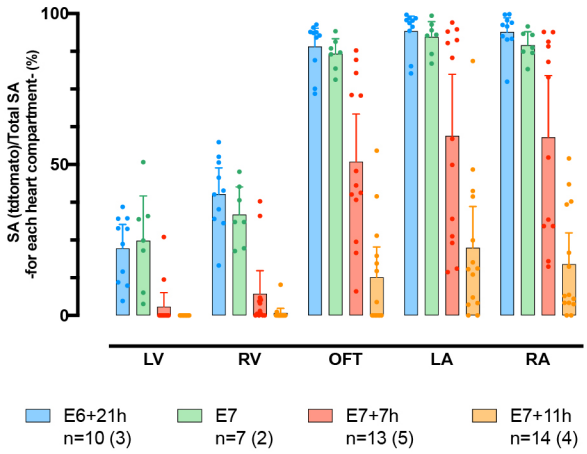




D

| | Embryo ID | Myocardium | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | LV>10% | LV<10% | RV>10% | RV<10% | OFT | LA | RA |
| Tam E6+21h | BRJH74.2c-1 | X | | X | | X | X | X |
| | BRJH74.2c-2 | X | | X | | X | X | X |
| | BRJH74.2c-3 | X | | X | | X | X | X |
| | BRJH74.3f-1 | X | | X | | X | X | X |
| | BRJH74.3f-2 | X | | X | | X | X | X |
| | BRJH74.3f-3 | | X | X | | X | X | X |
| | BRJH74.1d-1 | X | | X | | X | X | X |
| | BRJH74.1d-2 | X | | X | | X | X | X |
| | BRJH74.1d-3 | X | | X | | X | X | X |
| | BRJH74.1d-4 | | X | X | | X | X | X |
| Tam E7 | BRJH70.1d-1 | X | | X | | X | X | X |
| | BRJH70.1d-2 | X | | X | | X | X | X |
| | BRJH70.1d-3 | | X | X | | X | X | X |
| | BRJH70.1d-4 | X | | X | | X | X | X |
| | JSAJ1b-1 | X | | X | | X | X | X |
| | JSAJ1b-2 | X | | X | | X | X | X |
| | JSAJ1b-3 | | X | X | | X | X | X |
| Tam E7+7h | BRJH72.1-1 | X | | X | | X | X | X |
| | BRJH72.1-1 | | | | X | X | X | X |
| | BRJH60.2c-1 | | | | X | X | X | X |
| | BRJH60.2c-1 | | | | X | X | X | X |
| | BRJH60.2c-1 | | | | X | X | X | X |
| | BRJN38.54e-1 | | | | X | X | X | X |
| | BRJN38.54e-1 | | | | | X | X | X |
| | BRJN38.54e-1 | | | | | X | X | X |
| | BRJN55.1i | | | | | X | X | X |
| | BRJN53.3b-1 | | | | | X | X | X |
| | BRJN53.3b-2 | | | | | X | X | X |
| | BRJN53.3b-3 | | | | | X | X | X |
| | BRJN53.3b-4 | X | | X | | X | X | X |
| Tam E7+11h | BRJH75.2f-1 | | | | | | X | X |
| | BRJH75.2f-2 | | | | | X | X | X |
| | BRJH75.2f-3 | | | | | X | X | X |
| | BRJH75.2f-4 | | | X | | X | X | X |
| | BRJH75.2f-5 | | | | | | X | X |
| | BRJH78.1e-1 | | | | | X | X | X |
| | BRJH78.1e-2 | | | | | | X | X |
| | BRJH78.1e-3 | | | | | X | X | X |
| | BRJH78.1e-4 | | | | | | X | X |
| | BRJH78.1e-5 | | | | | | | X |
| | BRJH71.4e | | | | | | | |
| | BRJN57.3c-1 | | | | | | X | X |
| | BRJN57.3c-2 | | | | X | X | X | X |
| | BRJN57.3c-3 | | | | | X | X | X |

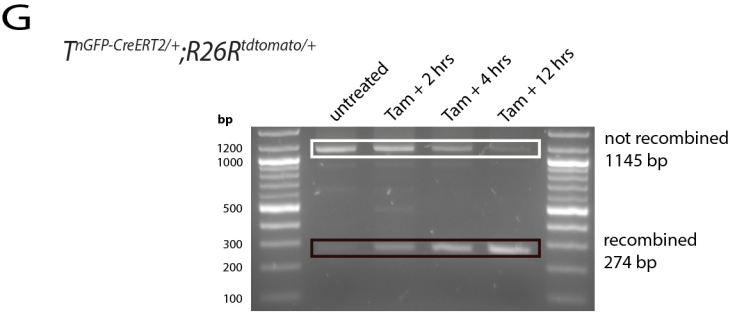
E



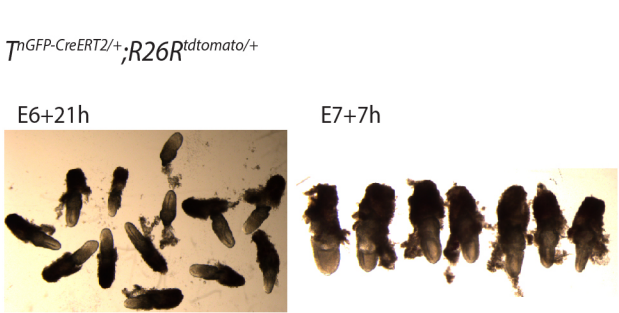

F

| TnGFP-CreERT2/+;R26Rtdtomato/+ | E6+21h (3) | E7+7h (3) | E7+11h (1) |
|---|---|---|---|
| ES | 3/18 (17%) | | |
| MS | 10/18 (56%) | | |
| LS | 5/18 (27%) | 2/19 (10%) | |
| OB | | 7/19 (37%) | 1/8 (10%) |
| EB | | 6/19 (32%) | 2/8 (37%) |
| LB | | 4/19 (21%) | 3/8 (32%) |
| EHF | | | 1/8 (21%) |

G

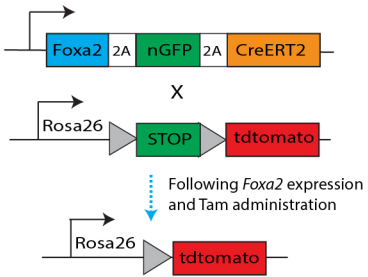



H

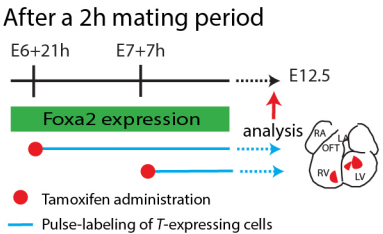


# Figure 2

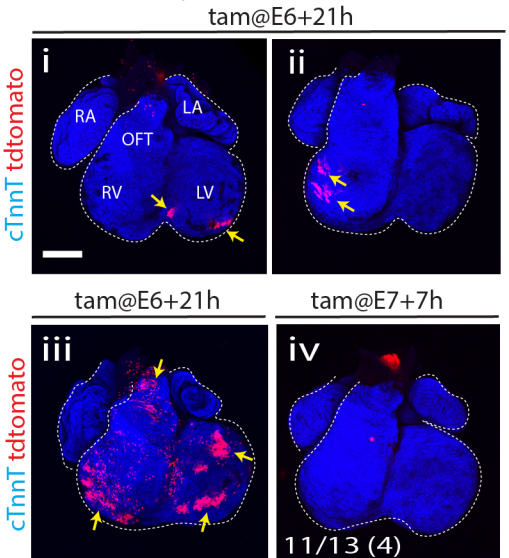
A




B




C




D

| | Embryo ID | Myocardium | | | | |
|---|---|---|---|---|---|---|
| | | LV | RV | OFT | LA | RA |
| Tam E6+21h (0.08mg/bw) | BRJH70.4b-1 | | X | | | |
| | BRJH70.4b-2 | X | X | X | | |
| | BRJH70.4b-3 | X | X | | | |
| | BRJH70.4b-4 | X | | | | |
| | BRJH70.4b-5 | | | | | |
| | BRJH70.4b-6 | X | X | X | | |
| | BRJH82.1g-1 | X | X | X | | |
| | BRJH75.4f-1 | X | | | | |
| | BRJH75.4f-2 | X | X | X | | |
| | BRJN58.1-1 | X | X | X | | |
| | BRJN58.1-2 | | | | | |
| | BRJN58.1-3 | X | X | X | | |
| | BRJN59.2d-1 | X | X | X | | |
| | BRJN59.2d-2 | | | | | |
| | BRJN59.2d-3 | X | X | X | | |
| | BRJH85.1e-1 | X | X | X | | |
| | BRJH85.1e-2 | X | X | X | | |

**Figure 3**

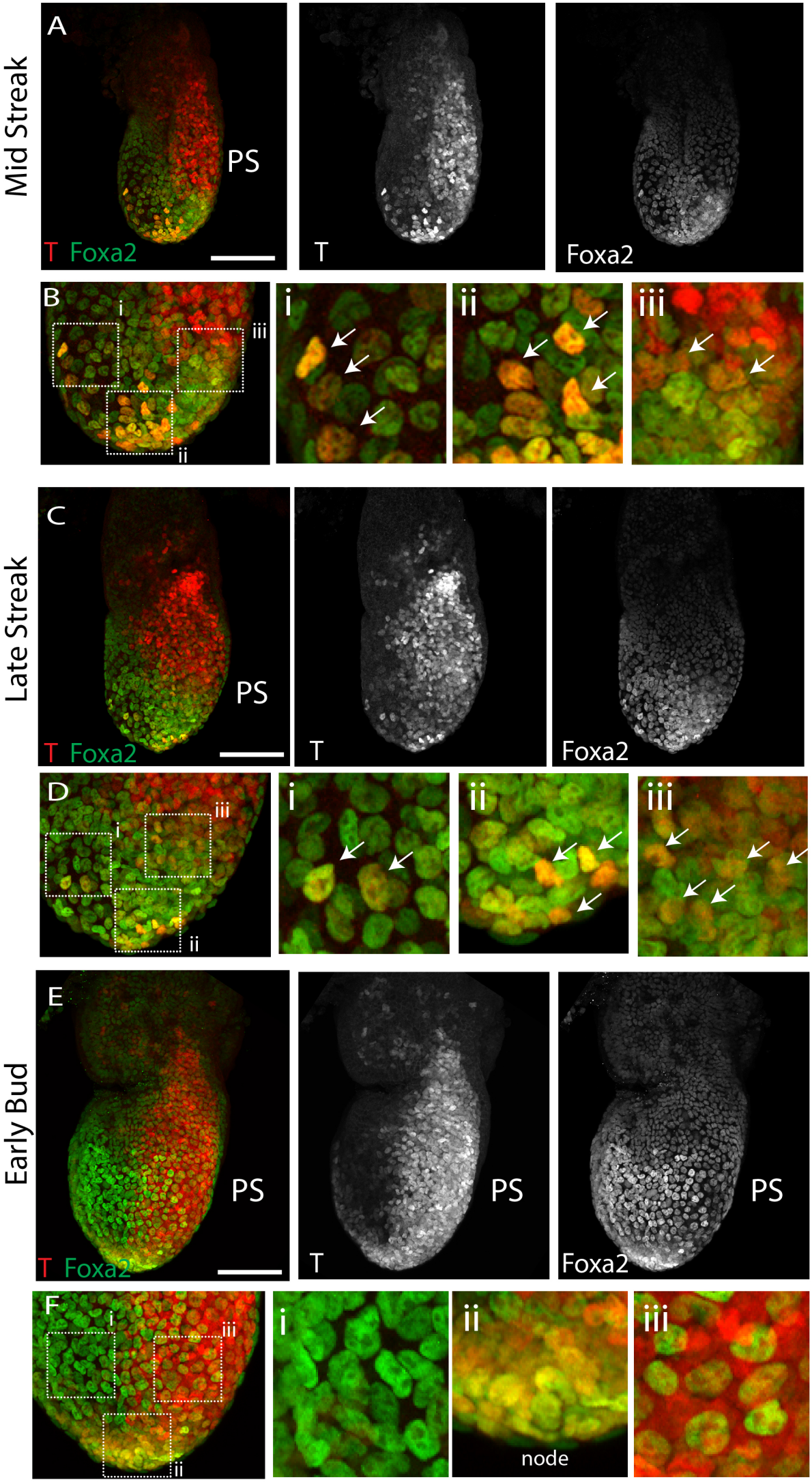
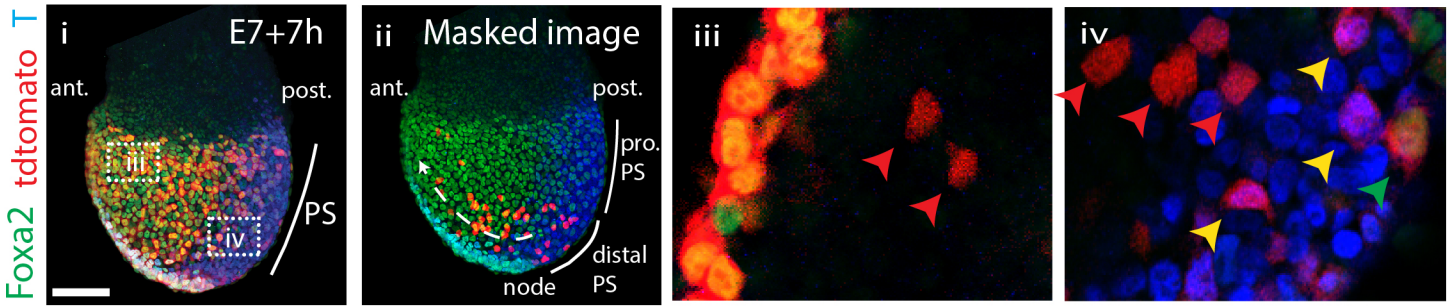

Mid Streak
A
PS
T Foxa2
T
Foxa2
B
i
ii
iii
Late Streak
C
PS
T Foxa2
T
Foxa2
D
i
ii
iii
Early Bud
E
PS
T Foxa2
T
PS
Foxa2
PS
F
i
ii
iii
node

**Figure 4**

A

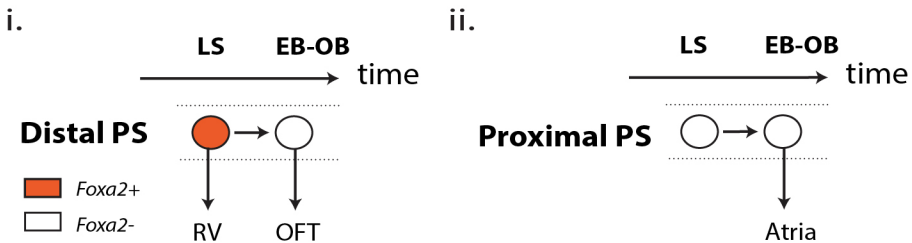
*Foxa2*$^{nGFP\text{-}CreERT2/+}$;*R26R*$^{tdtomato/+}$ tam@E6+21h



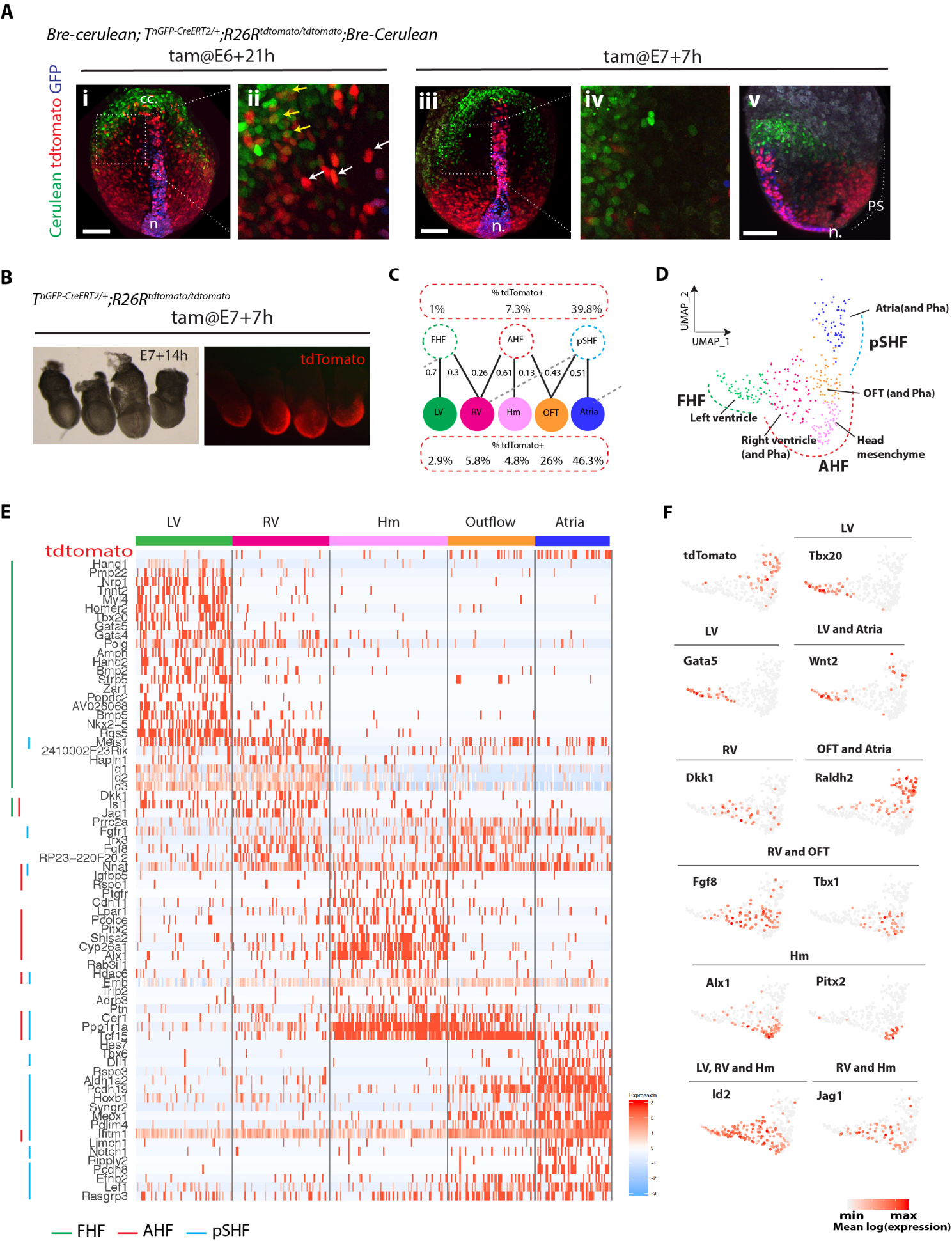

max z-projection

Foxa2- Tdtomato+

B



# Figure 5


A
Bre-cerulean; TnGFP-CreERT2/+;R26Rtdtomato/tdtomato;Bre-Cerulean
tam@E6+21h
tam@E7+7h
Cerulean tdtomato GFP
i
ii
iii
iv
v
cc
n
PS
B
TnGFP-CreERT2/+;R26Rtdtomato/tdtomato
tam@E7+7h
E7+14h
tdTomato
C
% tdTomato+
1%
7.3%
39.8%
FHF
AHF
pSHF
0.7
0.3
0.26
0.61
0.13
0.43
0.51
LV
RV
Hm
OFT
Atria
% tdTomato+
2.9%
5.8%
4.8%
26%
46.3%
D
UMAP_2
UMAP_1
Atria(and Pha)
pSHF
OFT (and Pha)
FHF
Left ventricle
Right ventricle (and Pha)
Head mesenchyme
AHF
E
LV
RV
Hm
Outflow
Atria
tdtomato
Hand1
Pmp22
Nrp1
Tnnt2
Myl4
Homer2
Tbx20
Gata5
Gata4
Polg
Ampn
Hand2
Bmp2
Strp5
Zar1
Popdc2
AV026068
Bmp5
Nkx2-5
Rgs5
Meis1
2410002F23Rik
Hapln1
Id1
Id2
Id3
Dkk1
Isl1
Jag1
Prrc2a
Fgfr1
Irx3
Fgf8
RP23-220F20.2
Nnat
Igfbp5
Rspo1
Ptgfr
Cdh11
Lpar1
Pcolce
Pitx2
Shisa2
Cyp26a1
Alx1
Rab3il1
Hdac6
Emb
Trib2
Adrb3
Ptn
Cer1
Ppp1r1a
Tcf15
Hes7
Tbx6
Dll1
Rspo3
Aldh1a2
Pcdh19
Hoxb1
Syngr2
Meox1
Pdlim4
Ifitm1
Limch1
Notch1
Ripply2
Pcdh8
Efnb2
Lef1
Rasgrp3
Expression
FHF
AHF
pSHF
F
LV
tdTomato
Tbx20
LV
LV and Atria
Gata5
Wnt2
RV
OFT and Atria
Dkk1
Raldh2
RV and OFT
Fgf8
Tbx1
Hm
Alx1
Pitx2
LV, RV and Hm
RV and Hm
Id2
Jag1
min
max
Mean log(expression)

# Figure 6

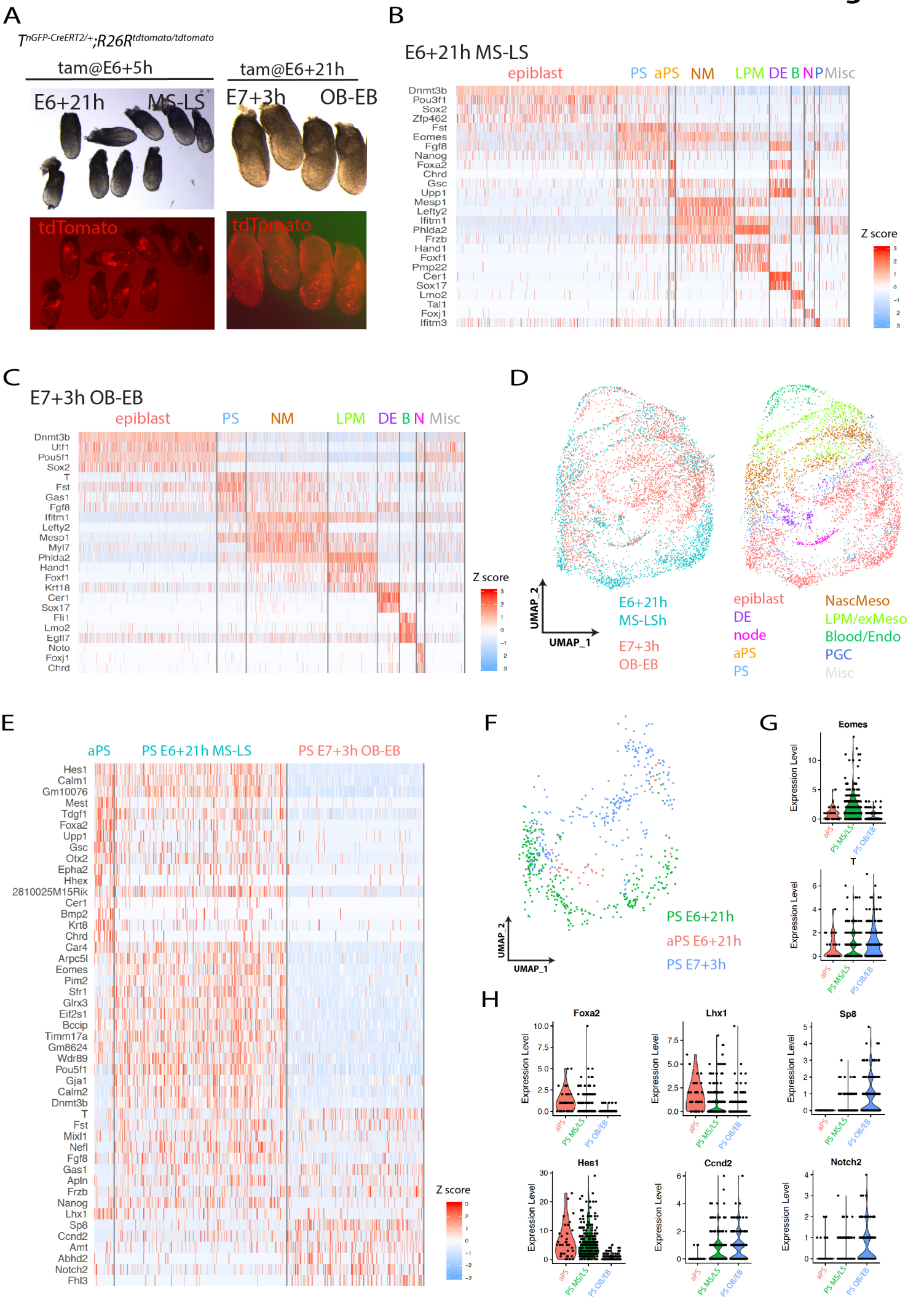
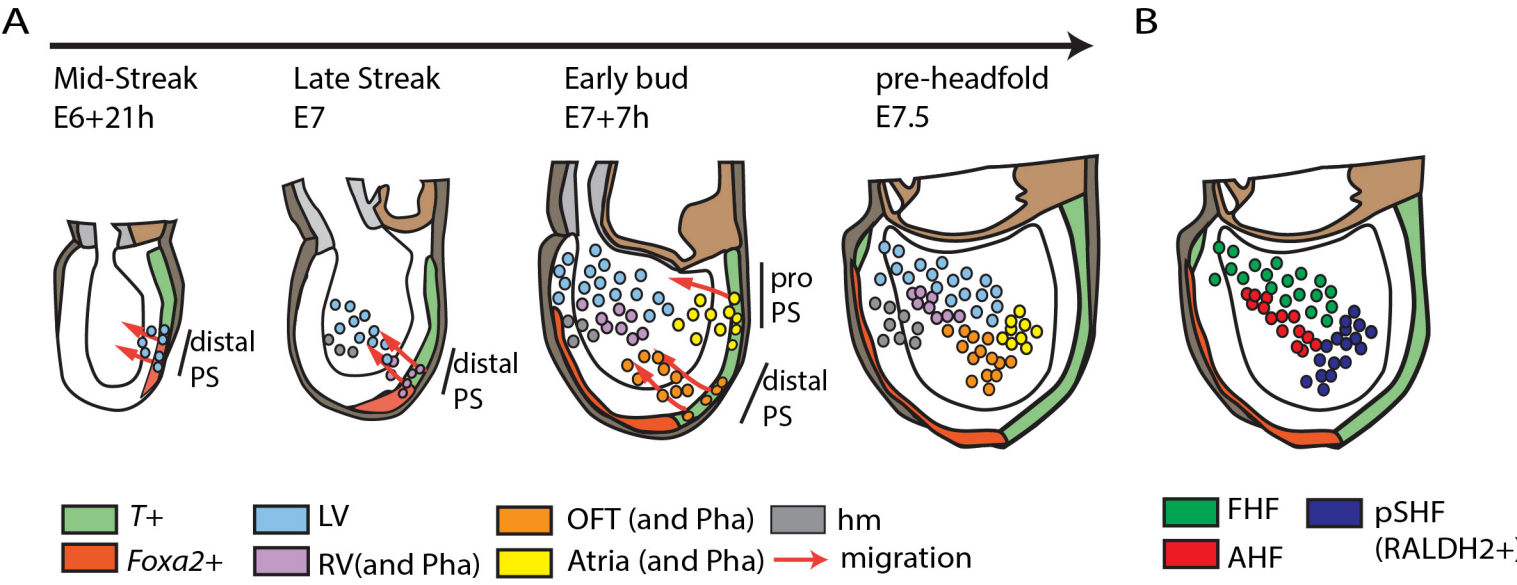
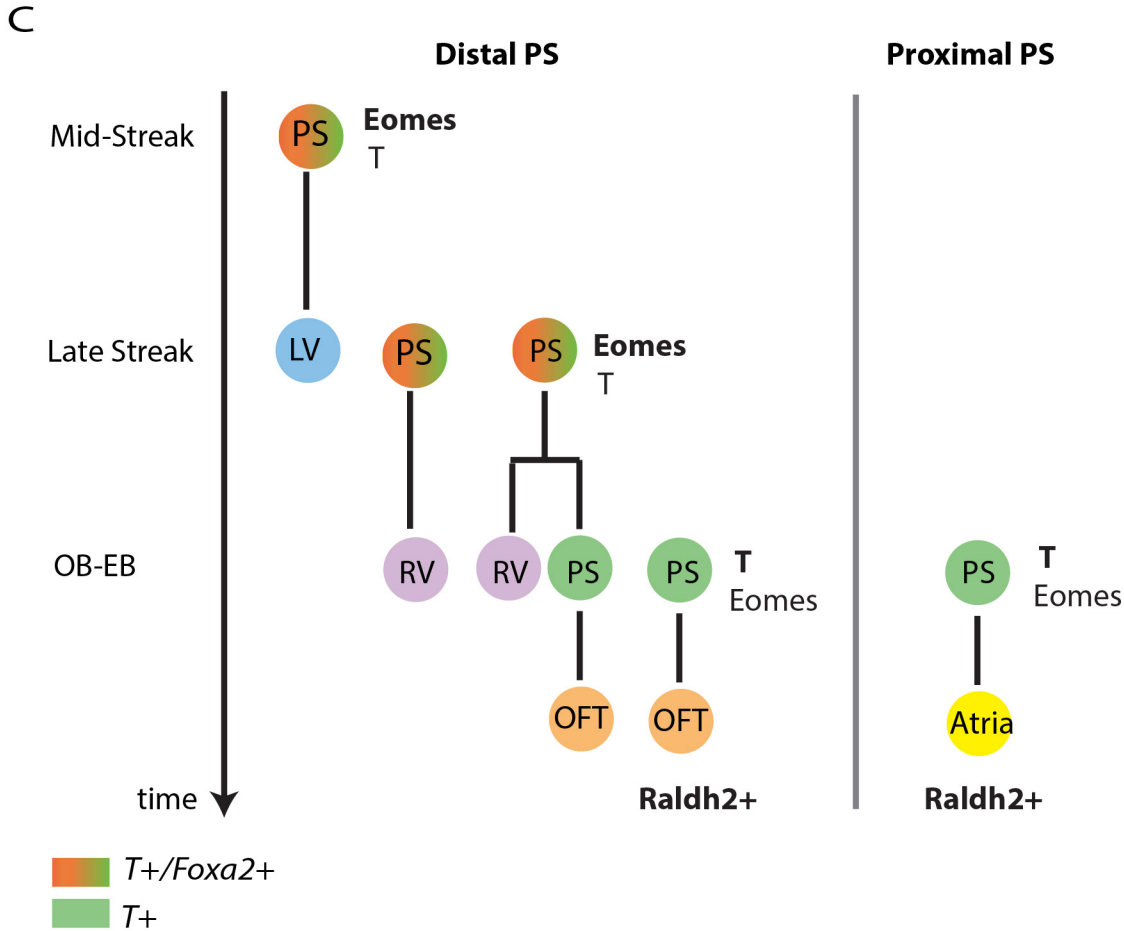

A
TnGFP-CreERT2/+;R26Rtdtomato/tdtomato
tam@E6+5h
tam@E6+21h
E6+21h
MS-LS
E7+3h
OB-EB
tdTomato
tdTomato
B
E6+21h MS-LS
epiblast
PS
aPS
NM
LPM
DE
B
N
P
Misc
Dnmt3b
Pou3f1
Sox2
Zfp462
Fst
Eomes
Fgf8
Nanog
Foxa2
Chrd
Gsc
Upp1
Mesp1
Lefty2
Ifitm1
Phlda2
Frzb
Hand1
Foxf1
Pmp22
Cer1
Sox17
Lmo2
Tal1
Foxj1
Ifitm3
Z score
C
E7+3h OB-EB
epiblast
PS
NM
LPM
DE
B
N
Misc
Dnmt3b
Utf1
Pou5f1
Sox2
T
Fst
Gas1
Fgf8
Ifitm1
Lefty2
Mesp1
Myl7
Phlda2
Hand1
Foxf1
Krt18
Cer1
Sox17
Fli1
Lmo2
Egfl7
Noto
Foxj1
Chrd
Z score
D
UMAP_2
UMAP_1
E6+21h
MS-LSh
E7+3h
OB-EB
epiblast
DE
node
aPS
PS
NascMeso
LPM/exMeso
Blood/Endo
PGC
Misc
E
aPS
PS E6+21h MS-LS
PS E7+3h OB-EB
Hes1
Calm1
Gm10076
Mest
Tdgf1
Foxa2
Upp1
Gsc
Otx2
Epha2
Hhex
2810025M15Rik
Cer1
Bmp2
Krt8
Chrd
Car4
Arpc5l
Eomes
Pim2
Sfr1
Glrx3
Eif2s1
Bccip
Timm17a
Gm8624
Wdr89
Pou5f1
Gja1
Calm2
Dnmt3b
T
Fst
Mixl1
Nefl
Fgf8
Gas1
Apln
Frzb
Nanog
Lhx1
Sp8
Ccnd2
Amt
Abhd2
Notch2
Fhl3
Z score
F
UMAP_2
UMAP_1
PS E6+21h
aPS E6+21h
PS E7+3h
G
Eomes
T
Expression Level
aPS
PS MS/LS
PS OB/EB
H
Foxa2
Lhx1
Sp8
Hes1
Ccnd2
Notch2
Expression Level
aPS
PS MS/LS
PS OB/EB

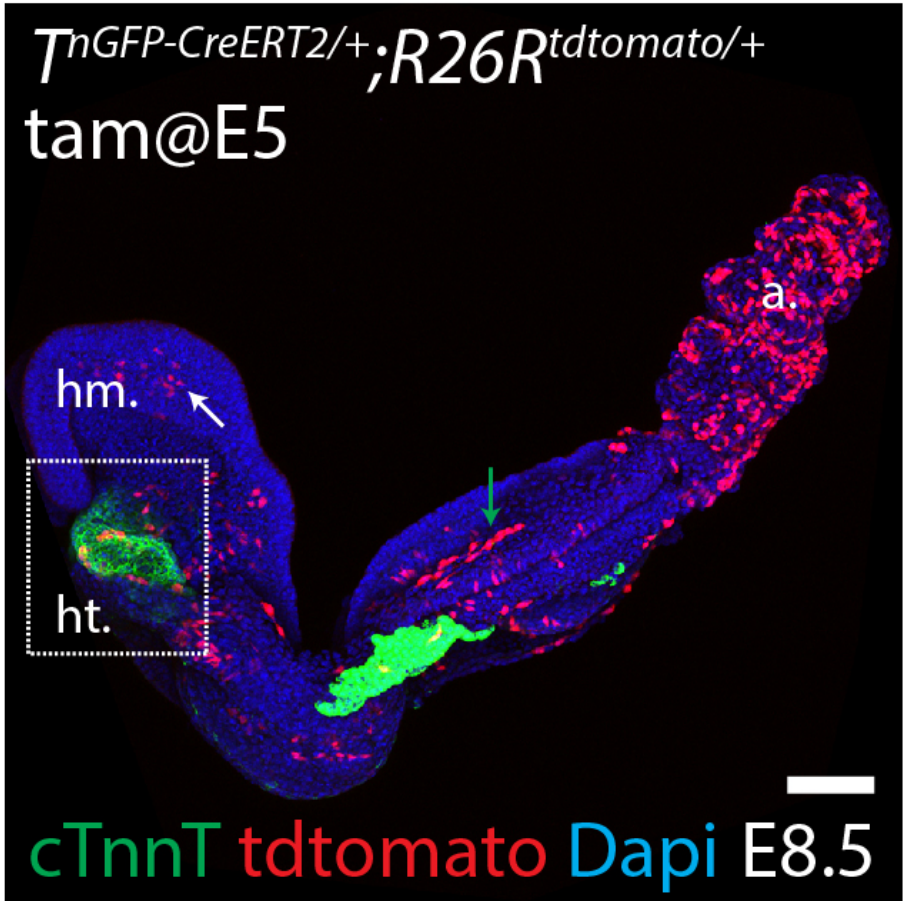
**Figure 7**

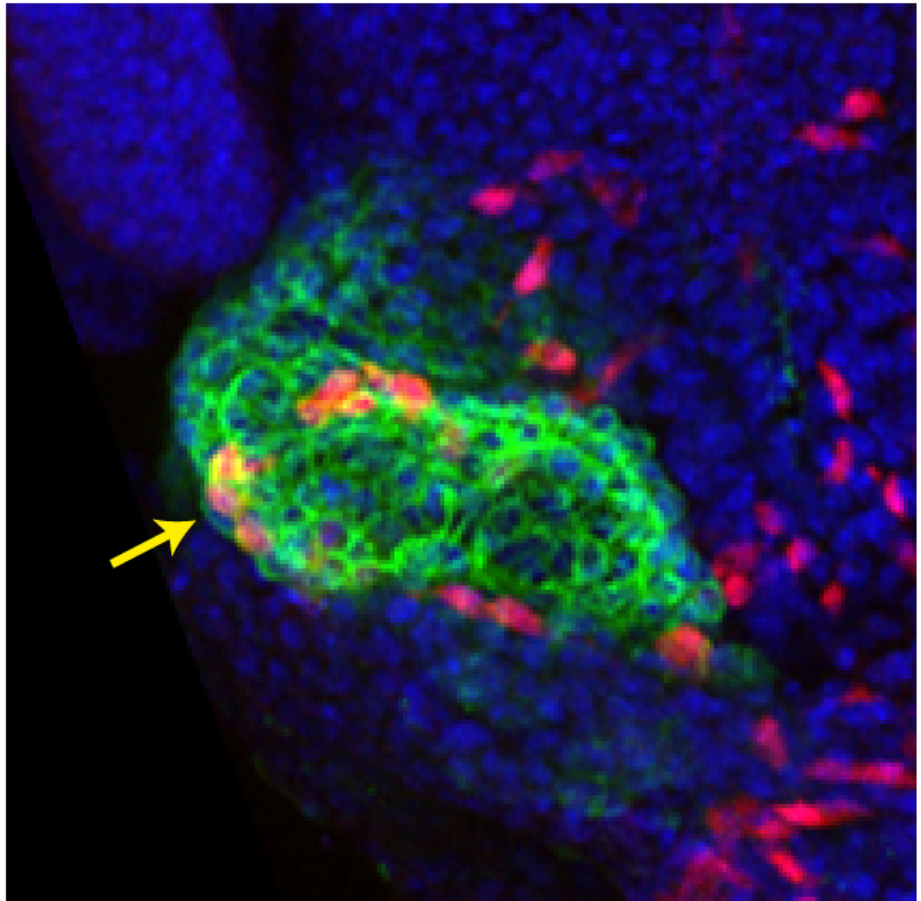

A
Mid-Streak
E6+21h
Late Streak
E7
Early bud
E7+7h
pre-headfold
E7.5
distal
PS
distal
PS
pro
PS
distal
PS
B
T+
Foxa2+
LV
RV(and Pha)
OFT (and Pha)
Atria (and Pha)
hm
migration
FHF
AHF
pSHF
(RALDH2+)


C
Distal PS
Proximal PS
Mid-Streak
Late Streak
OB-EB
time
PS
Eomes
T
LV
PS
PS
Eomes
T
RV
RV
PS
PS
T
Eomes
OFT
OFT
Raldh2+
PS
T
Eomes
Atria
Raldh2+
T+/Foxa2+
T+

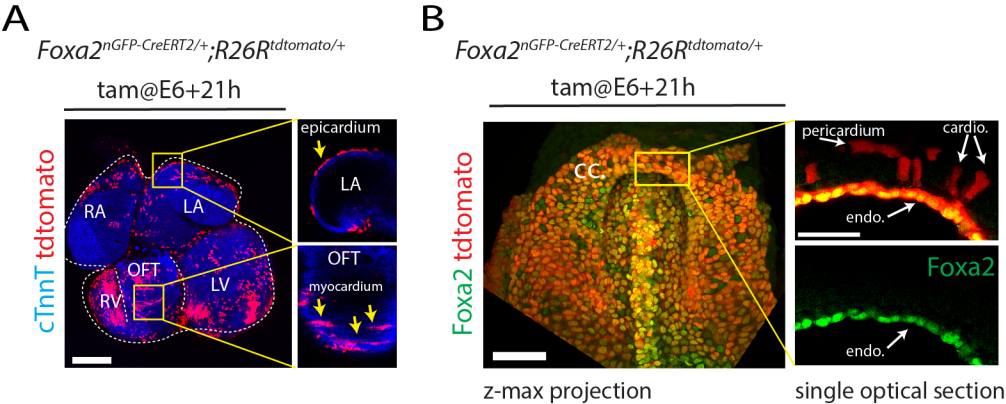
# S1 Fig.

A


T^nGFP-CreERT2/+;R26R^tdtomato/+
tam@E5
a.
hm.
ht.
cTnnT tdtomato Dapi E8.5

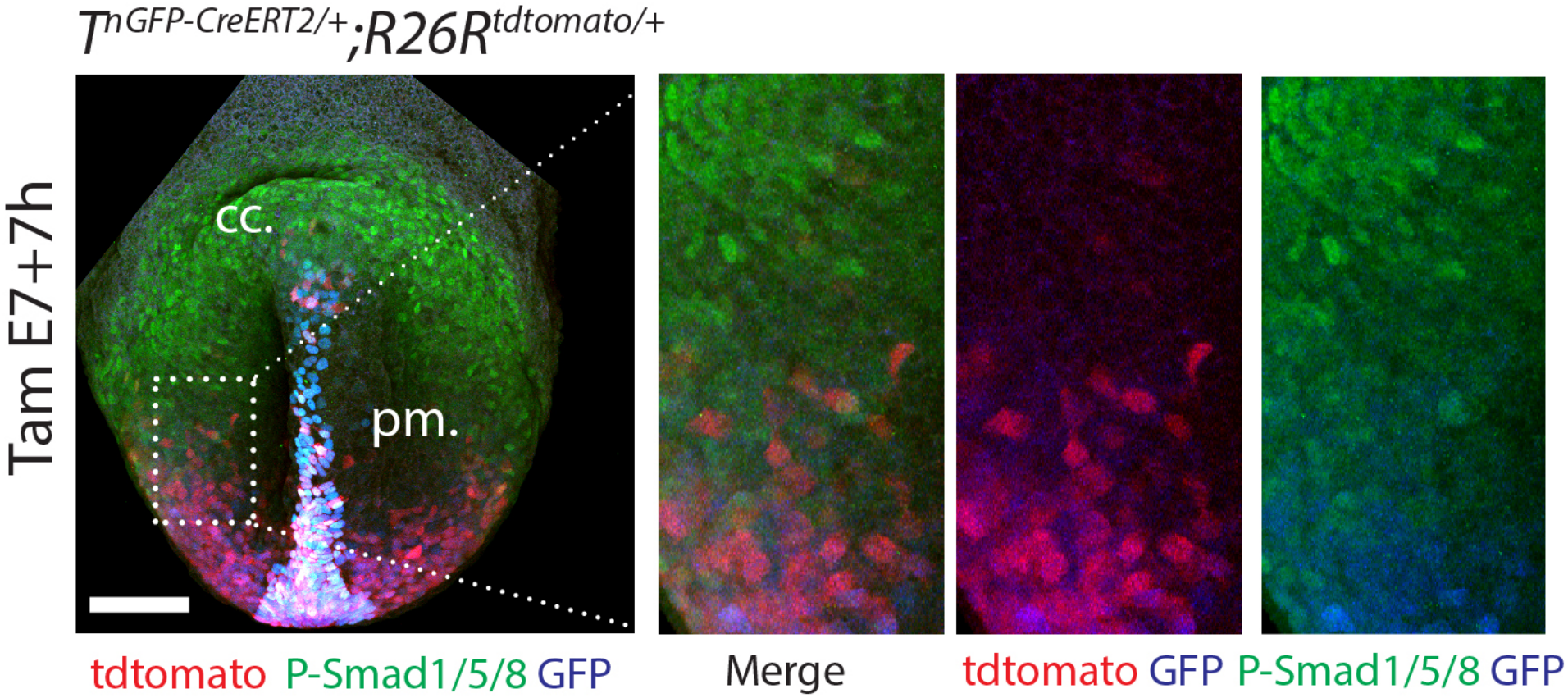


**S2 Fig.**

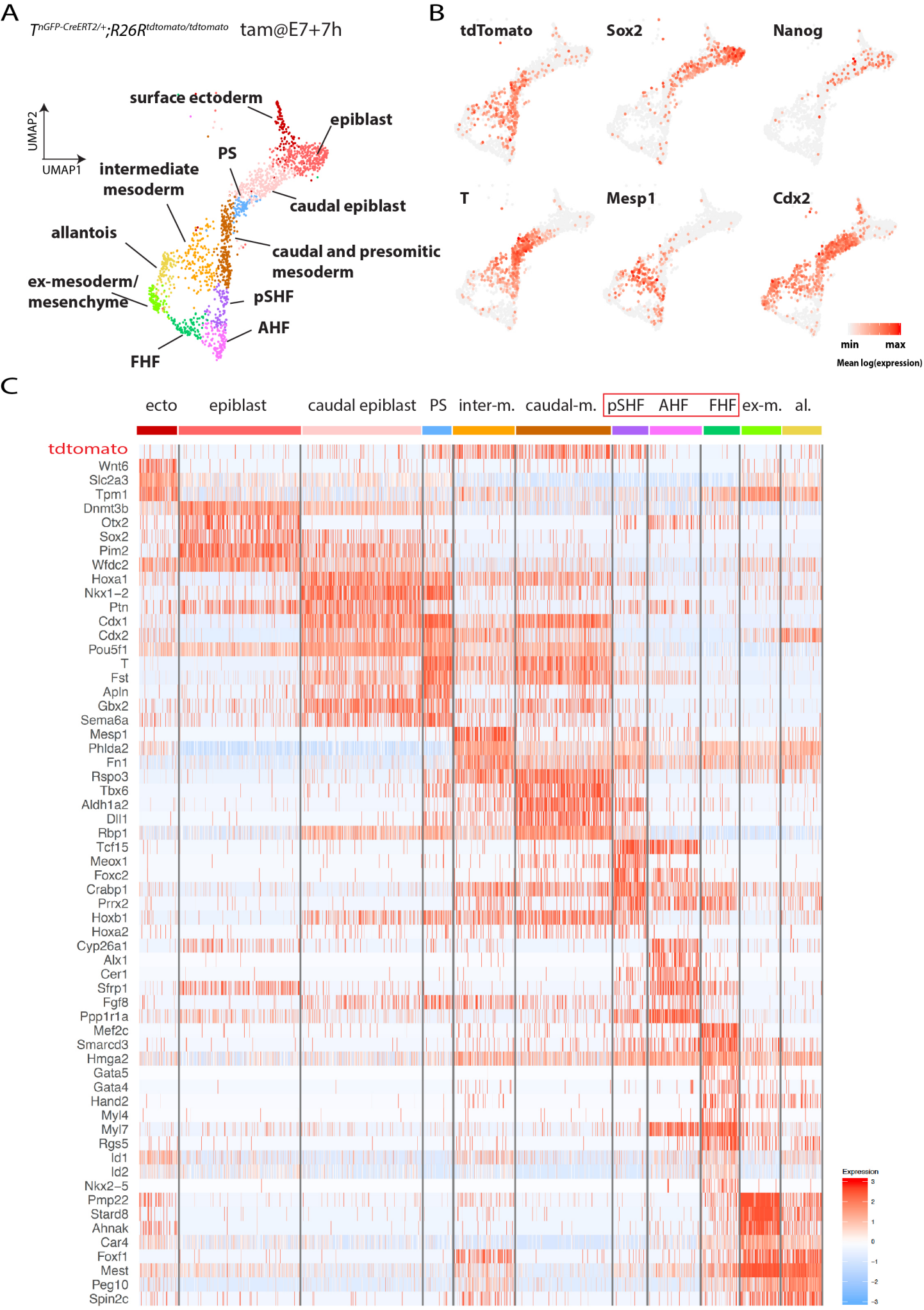

A
Foxa2nGFP-CreERT2/+;R26Rtdtomato/+
tam@E6+21h
cTnnT tdtomato
RA
LA
OFT
RV
LV
epicardium
LA
OFT
myocardium
B
Foxa2nGFP-CreERT2/+;R26Rtdtomato/+
tam@E6+21h
Foxa2 tdtomato
CC.
pericardium
cardio.
endo.
Foxa2
endo.
z-max projection
single optical section

**S3 Fig.**

A


$T^{nGFP\text{-}CreERT2/+}$;$R26R^{tdtomato/+}$
Tam E7+7h
cc.
pm.
tdtomato P-Smad1/5/8 GFP
Merge
tdtomato GFP P-Smad1/5/8 GFP

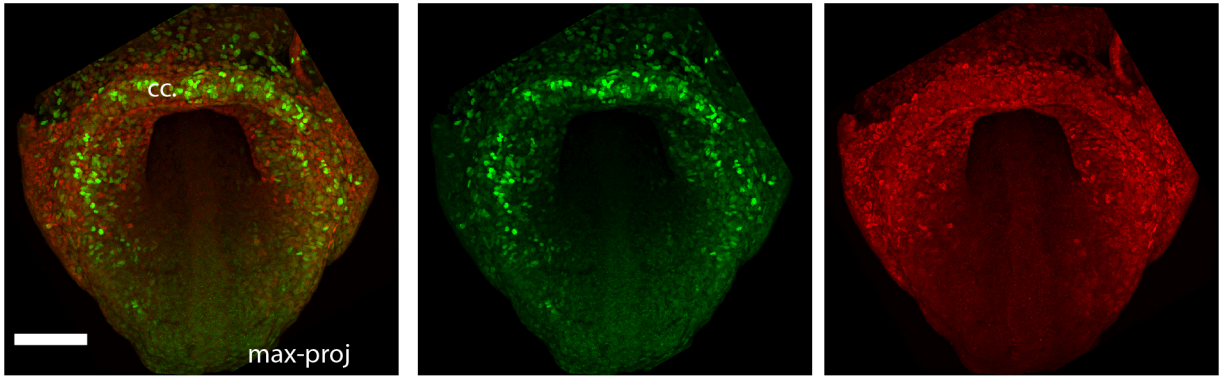
**S4 Fig.**


A
TnGFP-CreERT2/+;R26Rtdtomato/tdtomato tam@E7+7h
UMAP2
UMAP1
surface ectoderm
epiblast
PS
intermediate mesoderm
caudal epiblast
allantois
caudal and presomitic mesoderm
ex-mesoderm/ mesenchyme
pSHF
AHF
FHF
B
tdTomato
Sox2
Nanog
T
Mesp1
Cdx2
min
max
Mean log(expression)
C
ecto
epiblast
caudal epiblast
PS
inter-m.
caudal-m.
pSHF
AHF
FHF
ex-m.
al.
tdtomato
Wnt6
Slc2a3
Tpm1
Dnmt3b
Otx2
Sox2
Pim2
Wfdc2
Hoxa1
Nkx1−2
Ptn
Cdx1
Cdx2
Pou5f1
T
Fst
Apln
Gbx2
Sema6a
Mesp1
Phlda2
Fn1
Rspo3
Tbx6
Aldh1a2
Dll1
Rbp1
Tcf15
Meox1
Foxc2
Crabp1
Prrx2
Hoxb1
Hoxa2
Cyp26a1
Alx1
Cer1
Sfrp1
Fgf8
Ppp1r1a
Mef2c
Smarcd3
Hmga2
Gata5
Gata4
Hand2
Myl4
Myl7
Rgs5
Id1
Id2
Nkx2−5
Pmp22
Stard8
Ahnak
Car4
Foxf1
Mest
Peg10
Spin2c
Expression
3
2
1
0
−1
−2
−3

**S5 Fig.**

A

*Bre-Cerulean*



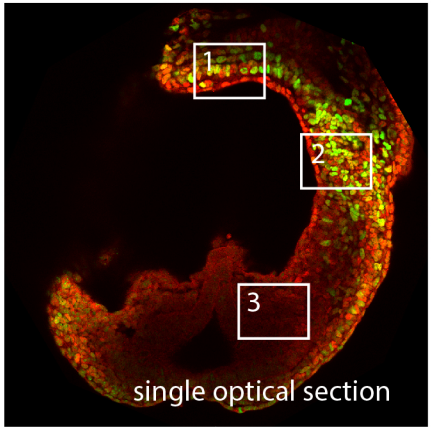

P-Smad1/5/8 Cerulean

Cerulean

P-Smad1/5/8

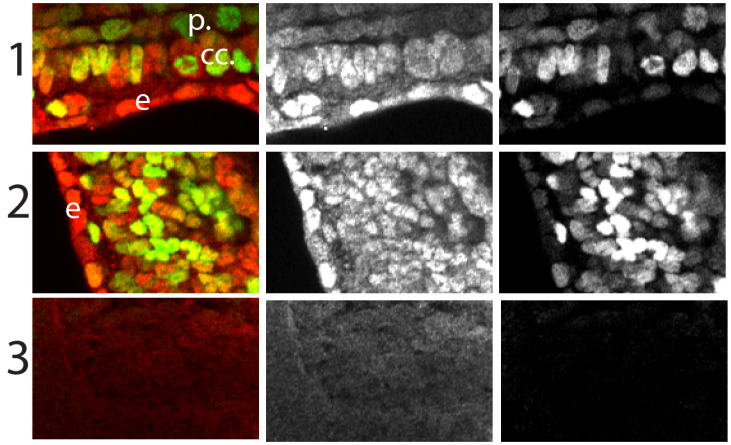
A′

*Bre-Cerulean*

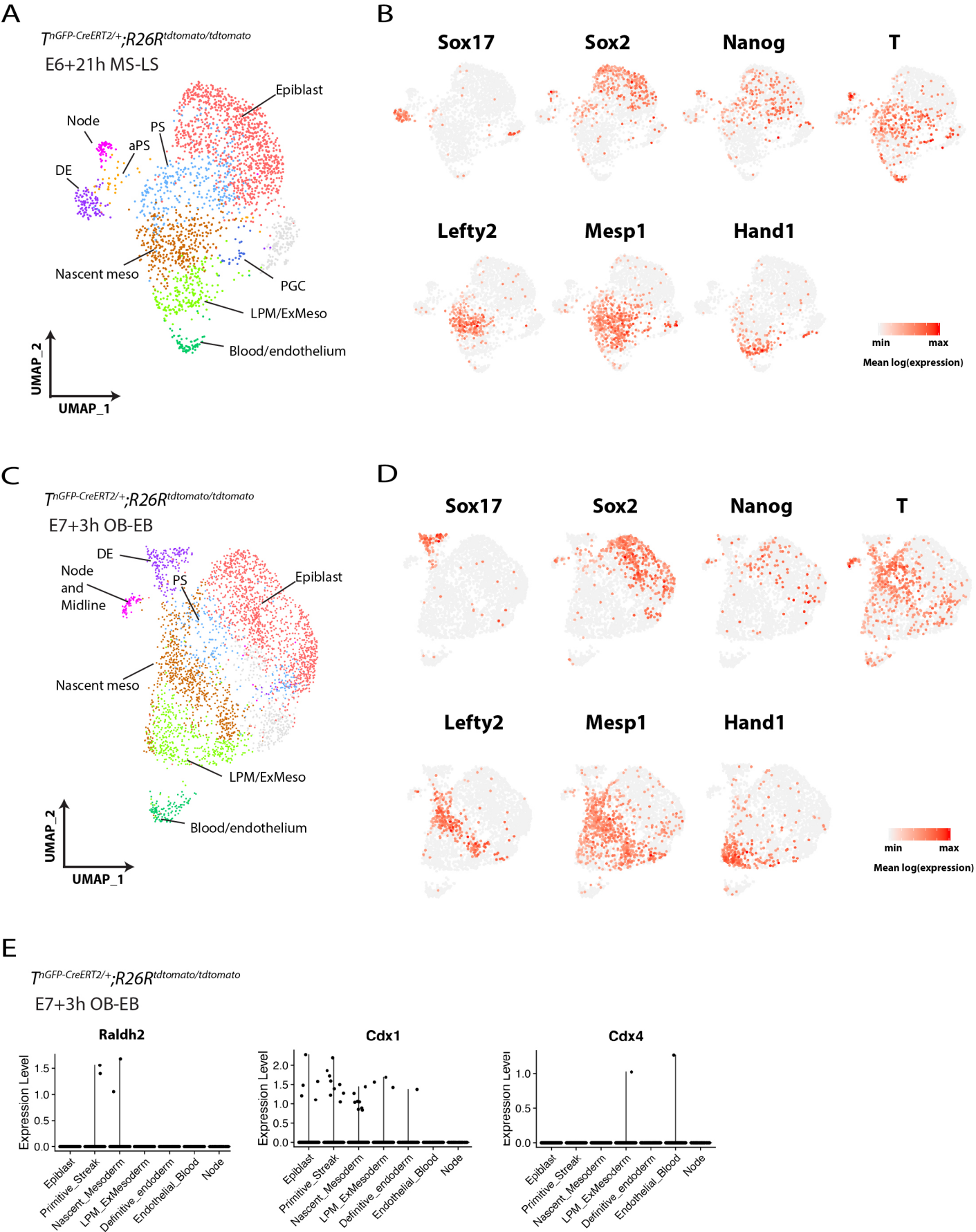



P-Smad1/5/8 Cerulean

A″




Merge P-Smad1/5/8 Cerulean


A
TnGFP-CreERT2/+;R26Rtdtomato/tdtomato
E6+21h MS-LS
Epiblast
Node
PS
aPS
DE
Nascent meso
PGC
LPM/ExMeso
Blood/endothelium
UMAP_2
UMAP_1
B
Sox17
Sox2
Nanog
T
Lefty2
Mesp1
Hand1
min
max
Mean log(expression)
C
TnGFP-CreERT2/+;R26Rtdtomato/tdtomato
E7+3h OB-EB
DE
Node and Midline
PS
Epiblast
Nascent meso
LPM/ExMeso
Blood/endothelium
UMAP_2
UMAP_1
D
Sox17
Sox2
Nanog
T
Lefty2
Mesp1
Hand1
min
max
Mean log(expression)
E
TnGFP-CreERT2/+;R26Rtdtomato/tdtomato
E7+3h OB-EB
Raldh2
Cdx1
Cdx4
Expression Level
Epiblast
Primitive_Streak
Nascent_Mesoderm
LPM_ExMesoderm
Definitive_endoderm
Endothelial_Blood
Node