Activation of an injury-associated transient progenitor state in the epicardium is required for zebrafish heart regeneration

Yu Xia[1,2,#], Sierra Duca[1,2,#], Björn Perder[1,2], Friederike Dündar[3,4,5], Paul Zumbo[3,5], Miaoyan Qiu[1,2], Jun Yao[1,2], Yingxi Cao[1,2], Michael R. Harrison[1,2], Lior Zangi[6,7,8], Doron Betel[3,5], Jingli Cao[1,2,*]

[1]Cardiovascular Research Institute, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10021, USA

[2]Department of Cell and Developmental Biology, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10021, USA

[3]Institute for Computational Biomedicine, Department of Medicine, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10021, USA

[4]Department of Physiology and Biophysics, Weill Cornell Medical College,1300 York Ave, New York, NY 10021, USA

[5]Applied Bioinformatics Core, Weill Cornell Medical College,1300 York Ave, New York, NY 10021, USA

[6]Cardiovascular Research Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA;

[7]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[8]Black Family Stem Cell Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

*Lead Contact: jic4001@med.cornell.edu

[#]These authors contributed equally to this work.

ABSTRACT

The epicardium, a mesothelial cell tissue that encompasses vertebrate hearts, supports heart regeneration after injury through paracrine effects and as a source of multipotent progenitors. However, the progenitor state in the adult epicardium has yet to be defined. Through single-cell RNA-sequencing of isolated epicardial cells from uninjured and regenerating adult zebrafish hearts, we defined the epithelial and mesenchymal subsets of the epicardium. We further identified a transiently activated epicardial progenitor cell (aEPC) subpopulation marked by *ptx3a* and *col12a1b* expression. Upon cardiac injury, aEPCs emerge from the epithelial epicardium, migrate to enclose the wound, undergo epithelial-mesenchymal transition (EMT), and differentiate into mural cells and *pdgfra*$^{+}$*hapln1a*$^{+}$ mesenchymal epicardial cells. These EMT and differentiation processes are regulated by the Tgfβ pathway. Conditional ablation of aEPCs blocked heart regeneration through reduced Nrg1 expression and mesenchymal cell number. Our findings identify a transient progenitor population of the adult epicardium that is indispensable for heart regeneration and highlight it as a potential target for enhancing cardiac repair.

Key words: epicardium; epicardial progenitor; heart regeneration; zebrafish; scRNAseq; EMT; modRNA

INTRODUCTION

Adult zebrafish possess a remarkable capacity for scarless heart regeneration after injury, which is achieved through the proliferation of existing cardiomyocytes (CMs)[1, 2, 3, 4]. CM proliferation is aided by the cellular and molecular environment provided by non-muscle tissues, such as the epicardium, a mesothelial layer of vertebrate hearts[5, 6, 7]. The epicardium is a heterogeneous population containing stem cells or progenitors that convert into other supporting cell types, such as mural cells (i.e., smooth muscle cells and pericytes) and fibroblasts, during development and regeneration[5, 8, 9, 10, 11]. Following cardiac injury in adult zebrafish, epicardial cells are activated to turn on embryonic genes, proliferate, and migrate to repopulate the wound site. In addition to supplying supporting cell types, epicardial cells provide paracrine signals and extracellular matrix (ECM) components for CM division and coronary angiogenesis, in addition to supplying the supporting cell types[5, 6, 7]. Recent studies have highlighted the vital roles of the epicardium in zebrafish heart regeneration[12, 13, 14], and the mobilization of epicardial cells has been reported to improve mammalian heart repair[15, 16]. However, the epicardial progenitor state during heart regeneration remains largely uncharacterized due to the lack of genetic tools to label and trace epicardial subsets.

For unbiased assessment of the epicardial progenitor state in regenerating hearts, we performed single-cell RNA-sequencing (scRNA-seq) of epicardial cells isolated from hearts undergoing regeneration. Using scRNA-seq analysis and genetic approaches, we defined the epithelial and mesenchymal subpopulations and the mural lineage of the epicardium. We also identified a transiently activated epicardial progenitor cell (aEPC) population and defined their molecular features. These aEPCs are indispensable for heart regeneration as they differentiate into mural cells and mesenchymal epicardial cells and supply pro-regenerative factors during regeneration. Therefore, our study reveals the active driver of epicardium-mediated heart regeneration and provides the basis for harnessing the epicardium for heart repair.

RESULTS

Cellular heterogeneity and EMT of epicardial cells in the adult zebrafish heart

In zebrafish, *tcf21* is a widely used epicardial marker that labels both quiescent and active epicardial cells, while other epicardial markers, such as *tbx18* and *wt1*, only label part of the epicardium[17]. We first examined the *tcf21*$^+$ epicardial cell distribution in uninjured hearts, using a nuclear EGFP reporter (nucEGFP) driven by the regulatory sequence of *tcf21*[17]. As shown in Figures 1A and 1B, *tcf21*$^+$ nuclei reside in multiple layers of the uninjured ventricular wall. While the outermost layer expresses *aldh1a2* (it is expressed in the endocardium as well; see also reference[17]), the inner layers of *tcf21*$^+$ nuclei enter the compact muscle and are *aldh1a2*-negative (Figures 1B and 1C). Thus, the *tcf21* reporters label 2 distinct epicardial subsets in the uninjured adult zebrafish heart: the epithelial epicardial cells (*aldh1a2*$^+$ outermost layer) and the mesenchymal epicardial cells that enter the compact muscle.

Previous studies have suggested that zebrafish epicardial cells undergo epithelial-mesenchymal transition (EMT) during heart regeneration[18, 19]. However, EMT cannot be definitively identified without a transgenic tool specific to the epithelial layer of the epicardium. To address this, we developed an approach of pericardial sac injection of modified RNA (modRNAs)[20, 21] for transient gene expression in the epicardium. We also created a bacterial artificial chromosome (BAC) transgenic line in which the *tcf21* regulatory sequences drive a Cre-releasable floxed BFP-stop cassette followed by an mCherry-NTR cassette (*tcf21:loxP-BFP-Stop-loxP-mCherry-NTR* or *tcf21:Switch* for short, Figure 1D). By injecting Cre modRNAs into fish carrying the *tcf21:Switch* line, we labeled the epithelial layer of the epicardium with mCherry (Figures 1D and 1E, uninjured). We did not observe a single labeled mesenchymal cell in the apex half of the uninjured ventricle 10 days after injection (Figure 1F, Ctrl, 13 hearts analyzed). Although we could not rule out labels of deeper *tcf21*$^+$ cells upon injection, combined use of the *tcf21:Switch* line with the pericardiac cavity injection of Cre modRNA injection specifically limits our labeling to the epithelial epicardium at least by day 10 after injection in the adults. To monitor EMT of epicardial cells, Cre modRNAs were injected 3 days before the amputation injury, and hearts were collected at 7 dpa to assess mCherry expression. We observed 26.8% on average of mCherry$^+$ cells entering the mesenchymal layer (Figures 1E and 1F, 7 dpa), indicating an EMT process in which the epithelial epicardial cells give rise to the mesenchymal epicardial cells during heart regeneration.

scRNA-seq reveals distinct subsets of the ventricular epicardial lineage during heart regeneration

To systemically dissect epicardial subsets, we applied scRNA-seq analysis of *tcf21*$^{+}$ cells during heart regeneration. Following cardiac injury in adult zebrafish, the epicardial cells turn on embryonic genes, a process called "activation"[5]. The epicardium is activated organ-widely at 3 dpa before restricting activation to the injury site starting at 7 dpa[18]. We injured *tcf21:nucEGFP* fish by partial ventricular amputation and collected ventricles at 3 or 7 dpa together with uninjured clutchmates (Ctrl). Single live nucEGFP^{+} cells were isolated by FACS and subjected to library preparation and scRNA-seq on the 10x Genomics Chromium platform (Figures 2A and S1). Three samples were collected that comprised of 4,970, 3,743, and 6,428 cells passing the quality control for Ctrl, 3 dpa, and 7 dpa, respectively (Figure 2B). Unbiased clustering with all samples identified 12 clusters (Figure 2C and Table S1). Eight clusters are grouped as the core epicardial population (Figure 2C, clusters 0-5, 7, and 9). Three clusters, including 8, 10, and 11, appear to represent contaminating non-epicardial cells (or doublets). Cells in cluster 8 express protein tyrosine phosphatase receptor type C (*ptprc*), *cd74a*, and macrophage expressed 1 (*mpeg1*), therefore representing immune cells[22, 23, 24]. Cluster 10 exhibits high expression levels of myosin light chain 7 (*myl7*) and troponin T type 2a (*Tnnt2a*), markers of CMs, indicating cardiac muscle identity[25, 26]. Cluster 11 appears to comprise endocardial and endothelial cells enriched for Fli-1 proto-oncogene ETS transcription factor a (*fli1a*) and kinase insert domain receptor like (*kdrl*) expression (Figures 2C, 2D, and S2A-B)[27, 28].

Besides these contaminating clusters, cluster 6 was identified as mural cells expressing platelet-derived growth factor receptor beta *pdgfrb* (Figure 2C)[29, 30]. Further analysis of the mural cells demonstrated enriched transcripts for known pericyte and vascular smooth muscle cell markers, including tropomyosin 1 (*tpm1*), notch receptor 3 (*notch3*), regulator of G protein signaling 5a (*rgs5a*), myosin heavy chain 11a (*myh11a*), actin alpha 2 smooth muscle (*acta2*), transgelin (*tagln*), ATP-binding cassette sub-family C member 9 (*abcc9*), and chemokine C-X-C motif ligand 12b (*cxcl12b*), as well as two recently identified mural cell markers NDUFA4 mitochondrial complex associated like 2a

(*ndufa4l2a*) and potassium voltage-gated channel Isk-related family member 4 (*kcne4*)[31, 32] (Figure S2C). While *ndufa4l2a* is restricted to a subset of mural cells, *kcne4* is highly expressed in mural cells and also present in the core epicardial clusters at lower levels. These gene expression profiles suggest diverse mural cell types (e.g., smooth muscle cells and pericytes) in our FACS isolated samples.

We next examined the temporal dynamics of the core epicardial clusters across all samples. Cluster 5 emerges at 3 dpa and is largely reduced by 7 dpa, while cluster 9 is mainly present at 7 dpa (Figures 2E-G). Cluster 2 expanded during regeneration, while the percentages of the remaining clusters (other than 2, 5, and 9) decreased at 3 dpa, but rebounded by 7 dpa (Figures 2E-G). Our gene expression analysis demonstrates that *tcf21* is an epicardial marker with broad expression across all core epicardial clusters and different states of injury (Figures 2H and S2E). Of note, we found a relatively lower *tcf21* expression level in the 3 dpa-specific cluster 5 than in other clusters (Figures 2H, S2D, and S2E). This seems to match our recent discovery that the Tcf21 binding motifs are enriched in chromatin regions with decreased accessibility in epicardial cells at 3 dpa, which may suggest a transition in cell state[33]. In contrast to the broad expression of *tcf21*, other known epicardial markers such as *tbx18*, *wt1b*, *sema3d*, and *aldh1a2* are only enriched in specific subpopulations[17, 34]. For instance, *wt1b* expression is enriched in Cluster 1, whereas *tbx18* is relatively depleted in that cluster compared to others (Figures 2H and S2D). *sema3d* is mainly expressed in cluster 2, while *aldh1a2* is expressed in both clusters 2 and 5. In addition, vascular endothelial growth factor Aa (*vegfaa*), a pro-angiogenic factor that was reported to be expressed by the epicardium and endocardium upon heart injury[35], is enriched in the mural cells and core epicardial clusters 0 and 1 (Figures 2H and S2A). Collagen type I alpha 2 (*col1a2*), a marker of the epicardium and cardiac fibroblasts[36], is highly expressed in all clusters, including the mural cells (Figure 2H). In all, these results suggest a dynamic cellular heterogeneity of the epicardium and its derivatives during regeneration, and that clusters 5 (predominantly 3 dpa cells) and 9 (7 dpa) are likely the injury-induced pro-regenerative subsets.

The epithelial and mesenchymal epicardium

To explore the heterogeneity within the core epicardial populations, we identified marker genes that distinguish these clusters. Clusters 0 and 1 have enriched expression of the pro-angiogenic factor *vegfaa* (Figures 2H and S2A)[35], and clusterin (*clu*) is highly expressed in clusters 3, 7, and 9 (Figure 2D). *sema3d*, *aldh1a2*, and podocalyxin-like (*podxl*) define cluster 2 (Figures 2H and 3A). Podxl was reported to localize to the apical plasma membrane of epithelial or endothelial cells[37, 38], and thus is a sign of epithelial identity in the epicardium. The hyaluronic acid-organizing factors hyaluronan and proteoglycan link protein 1a (*hapln1a*), the cardiac mesenchymal stem cell and cardiac fibroblast marker platelet-derived growth factor receptor alpha (*pdgfra*), as well as the myocardial mitogen neuregulin 1 (*nrg1*) mainly label clusters other than 2 and 5 (Figures 2H and 3A)[12, 39, 40]. Interestingly, *podxl* and *aldh1a2* label clusters (clusters 2 and 5) distinct from the *hapln1a*$^{+}$*pdgfra*$^{+}$ population (0, 1, 3, 4, 7, and 9; Figures 3A and B). HCR staining results indicate that *podxl* is expressed in the epithelial layer of the epicardium (Figure 3C). By contrast, *hapln1a* is expressed by the inner layer of *tcf21*$^{+}$ epicardial cells, which represent the mesenchymal layer (Figure 3D). In agreement with this finding, a recent study showed that *hapln1a* is expressed in an epicardial subset residing in the compact muscle that mediates hyaluronic acid (HA) secretion and myocardial regeneration[12]. Thus, the *hapln1a*$^{+}$*pdgfra*$^{+}$ cells (clusters 0, 1, 3, 4, 7, and 9) are mesenchymal epicardial cells. Moreover, we noticed that a subset of the 3 dpa-specific cluster 5 expresses *podxl* (Figure 3A), while the remainder of cluster 5 cells instead express relatively high levels of *snai1a* (Figure 3E), likely representing cells undergoing EMT. A re-clustering and focused analysis on only the core epicardial clusters demonstrated the same findings (Figure S3 and Table S2). These analyses suggest that cluster 5 is a transitional population, and that a subset of these cells is likely undergoing an EMT process.

***ptx3a* and *col12a1b* label a transient, pro-regenerative epicardial subtype**

We next focused on the markers of the 3 dpa-specific cluster 5 to characterize its identity. This cluster is enriched with pro-regenerative ECM and related genes including pentraxin 3 long a (*ptx3a*), collagen type XII alpha 1b (*col12a1b*), myristoylated alanine-rich protein kinase C substrate b (*marcksb*), and fibronectin 1a (*fn1a*) (Figures 3E and S4)[41, 42, 43, 44,

[45, 46, 47]. In addition to the injury-induced epicardial expression of *fn1a* that is required for heart regeneration[41], collagen XII (Col XII) deposition was reported to be boosted in both the epicardium and wounded tissues after cryoinjury in zebrafish[42]. However, the cellular sources of Fn and Col XII within the epicardial population were unclear. Col XII is also an axon growth-promoting ECM that helps zebrafish spinal cord regeneration[43, 44]. Ptx3 is a secreted humoral innate immunity factor that orchestrates inflammation and tissue repair[45]. Besides the interactions with pathogens and complement molecules, Ptx3 also interacts with ECM components, such as fibrin and plasminogen, to promote a timely removal of fibrotic ECM for efficient tissue repair[48, 49]. Additionally, Ptx3 was reported to have a cardioprotective function after acute myocardial infarction[50]. MARCKS is a ubiquitous substrate for protein kinase C and regulates the secretion of different substances. It has been shown to be highly upregulated during optic nerve regeneration in zebrafish, lens regeneration in newts, and cardiac tissue regeneration following infarction in mice[46, 47, 51]. Other cluster 5 enriched top markers include high mobility group box 2b (*hmgb2b*), ATP synthase membrane subunit c locus 1 (*atp5mc1*), serpin peptidase inhibitor clade H member 1a (*serpinh1a*), proteasome 20S subunit beta 1 (*psmb1*), and heat shock protein 90 beta member 1 (*hsp90b1*) (Figure S5). In addition, almost all proliferating cells (*top2a*$^{+}$, a G2/M phase marker) are within the 3 dpa-specific cluster 5 (Figure 3E), suggesting that cluster 5 is likely the primary cellular driver to restore the epicardial population after amputation injury. For an overview of the biological functions of marker genes for each cluster, we performed Gene Ontology (GO) enrichment analysis with specifically enriched transcripts in each core epicardial cluster (Figure S3E and Table S3). Notable enriched GO terms in the 3 dpa-specific subset include ECM organization, regeneration, metabolic processes, and translation. These results suggest that the 3 dpa-specific cluster is a pro-regenerative subset that mediates ECM remodeling, immune responses, and epicardial cell repopulation.

We next characterized marker gene expression for the 3 dpa-specific cluster through HCR staining. *ptx3a* is undetectable in the epicardium in uninjured hearts (Figure 4A). Upon amputation injury, *ptx3a* is initially expressed in the *tcf21*$^{+}$ epithelial layer of the entire ventricular epicardium at 1 dpa (Figure 4B). By 3 dpa, the epicardial *ptx3a* transcripts are mostly restricted to the injury site in both the epithelial and mesenchymal

layers of the epicardium, labeling the leading front of the regenerating *tcf21*$^{+}$ cells in the wound region (Figures 4C and S4). Interestingly, *tcf21:nucEGFP* expression is reduced in these *tcf21*$^{+}$*ptx3a*$^{+}$ leader cells (Figure 4C), which recapitulates the scRNA-seq result that the 3 dpa-specific cells have relatively lower *tcf21* expression compared to cells of the other core clusters (Figure 2H). This suggests changes in cell state including cell proliferation. At 7 dpa, the injury site is repopulated with predominantly *tcf21*$^{+}$ cells, and *tcf21:nucEGFP* expression is comparable to the flanking regions. However, only the cells in the most newly regenerated region of the apex are *ptx3a*$^{+}$ (Figure 4D), further suggesting that *ptx3a* marks the leading front of the regenerating epicardium. The expression levels of *ptx3a* peak at 3 dpa, decrease at 7 dpa, and are minimal by 14 dpa (Figure 4E). The results of *col12a1b* HCR staining demonstrated an expression pattern similar to *ptx3a*, as suggested by our scRNA-seq analysis (Figures S4 and S6). These results indicate that the 3 dpa-specific cluster labeled by *ptx3a* and *col12a1b* represents a transient and pro-regenerative epicardial cell population.

To aid subset labeling and tracing, we generated knock-in alleles using the Crispr/Cas9 technique. A mScarlet-P2A-NTR-polyA cassette was inserted right after the start codon of the *ptx3a* gene using a double-stranded HDR template that has 463 bp and 372 bp homology arms flanking the cassette. A seamless insertion allele *ptx3a*$^{mScarlet\text{-}P2A\text{-}NTR}$ (*ptx3a*^{RNTR} for short) was recovered through genotyping and sequencing (Figure 4F). No noticeable defect in heart development was observed in fish carrying the heterozygous knock-in allele (data not shown). mScarlet signals in *ptx3a*^{RNTR} hearts emerge in *tcf21*$^{+}$ cells from 1 dpa, peak at 3 dpa in the regenerating epicardial cells flanking the injury site, decrease at 7 dpa, and are barely detectable by 14 dpa (Figures 4G and S7), recapitulating the endogenous *ptx3a* expression pattern. Of note, the lower mScarlet signals at 1 and 2 dpa is only visible after an anti-DsRed antibody staining. A *col12a1b*^{EGFP} allele was also generated using the same strategy to visualize the endogenous *col12a1b* expression (Figures 4H, 4I, and S8). We demonstrate that both *ptx3a*^{RNTR} and *col12a1b*^{GFP} drive the same expression pattern as the genes *ptx3a* and *col12a1b*: labeling the 3 dpa-specific cluster after heart injury with no apparent epicardial expression in the uninjured adult hearts. Thus, we can define the transient pro-regenerative epicardial subtype labelled by expression of *ptx3a* or *col12a1b*.

The *ptx3a*$^{+}$*col12a1b*$^{+}$ subset contains activated epicardial progenitor cells (aEPCs) that differentiate into mural cells

To infer the origin and fates of the *ptx3a*$^{+}$*col12a1b*$^{+}$ subset, we applied the Monocle3 trajectory reconstruction algorithm to all clusters[52]. We found that the 3 dpa-specific subset sits on a branching point leading to the epithelial cluster 2 (branch a), part of the mural population (cluster 6, branch b), and the mesenchymal subsets (branch c, Figures 5A and 5B). This suggests that the 3 dpa-specific cluster is a progenitor state with the potential to give rise to different types of cells, which we thus named the activated epicardial progenitor cell (aEPC) population. To define the origin of these aEPCs, zebrafish carrying the *tcf21:Switch*;*tcf21:CreER*^{t2};*col12a1b*^{EGFP} reporters were treated with 4-Hydroxytamoxifen (4-HT) at both the embryonic (1 to 5 days post fertilization (dpf)) and adult stages (from 6 to 4 days before the heart injury, Figure 5C). Hearts were collected at 3 dpa to assess colocalization of the mCherry and GFP signals. As shown in Figure 5D, all EGFP^{+} cells around the wound are mCherry^{+}, suggesting that these *col12a1b*$^{+}$ aEPCs are derived from the spared epicardial cells upon injury, particularly from the epithelial layer of *tcf21*$^{+}$ cells (i.e., cluster 2, *podxl*$^{+}$).

Interestingly, we noticed a standalone population of mural cells (branch “d”) in addition to the aEPC-derived branch “b” (Figure 5B). A recent study found that *pdgfrb*$^{+}$ cardiac mural cells are originated from the epicardium during heart development[29]. However, whether epicardial cells give rise to *pdgfrb*$^{+}$ cells during heart regeneration is still unclear. Gene expression analysis indicates that the branch “b” mural cells express *fn1a*, while the branch “d” cells are mostly negative (Figure 5E). Because *fn1a* expression is restricted to the injury site after 1 dpa[41], we hypothesized that the branch “b” (*fn1a*$^{+}$) mural cells in the wound are derived from aEPCs. We next crossed the *tcf21:H2A-mCherry* (or *tcf21:H2R* for short) line with a *pdgfrb:EGFP* reporter[29]. Upon heart injury, we observed *pdgfrb:*EGFP^{+}*tcf21*:H2R^{+} cells in the wound at 7 dpa (Figure 5F), further suggesting an epicardial origin of these mural cells in the wound. To confirm the differentiation capacity of aEPCs to mural cells, we generated a *ptx3a:CreER*^{t2} BAC line and crossed it with the *ubi:loxP-EGFP-loxP-mCherry* (*ubi:Switch*) line[53]. Adult zebrafish carrying the *ubi:Switch*;*ptx3a:CreER*^{t2} reporters were treated with 4-HT from 2 to 5 dpa,

and hearts were collected at 14 dpa for whole-mount HCR staining of *pdgfrb* (Figure 5G). We observed mCherry^{+}*pdgfrb*$^{+}$ cells in the injury site, confirming an aEPC-to-mural differentiation (Figure 5H). These results support the scRNA-seq-inferred notion that the 3 dpa-specific cluster entails activated epicardial progenitor cells that can give rise to mural cells in the regenerated hearts.

aEPCs give rise to both the epithelial and mesenchymal epicardium

To further characterize the additional aEPC differentiation potential for the mesenchymal epicardium (branch "c" in Figure 5B), we focused on cells of the core clusters for which we observed dynamic gene expression patterns along the pseudotime branches "a" and "c" (Figures S3G). The epithelial epicardium initially expresses *fn1a*, *ptx3a*, and *col12a1b* to become aEPCs, followed by a transition to *hapln1a*-expressing mesenchymal cells. To confirm these gene expression dynamics, we performed HCR staining of *hapln1a* and *ptx3a*. As shown in Figure 6A, cells expressing *hapln1a* in the mesenchymal epicardium lag behind the *ptx3a*$^{+}$ leader cells that repopulate the wound at 3 dpa. At 7 dpa, *hapln1a* expression is enriched in the regenerated epicardial cells that flank the *ptx3a*$^{+}$ cells in the wound (Figure 6B). By 14 dpa, the regenerated *tcf21*$^{+}$ cells in the wound are *hapln1a*$^{+}$ but *ptx3a*$^{-}$. These observations support the notion that aEPCs are the first responders in regenerating the epicardium, and that *hapln1a* likely marks the mature mesenchymal epicardium. To confirm the differentiation of aEPCs to *hapln1a*$^{+}$ cells and the final fate of aEPCs upon completion of regeneration, adult zebrafish carrying the *ubi:Switch*;*ptx3a:CreER*^{t2} reporters were treated with 4-HT from 2 to 5 dpa, and hearts were collected at 30 dpa for whole-mount HCR staining of *hapln1a* and *podxl* (Figure 6D). We observed prominent colocalization of mCherry with *hapln1a* or *podxl* expression at 30 dpa (Figures 6E-G). The mCherry^{+} cells reside in both the epithelial (Figures 6E-G, arrowheads) and mesenchymal layer (Figures 6E-G, arrows) of the regenerated epicardium, indicating a final fate of aEPCs to both epicardial layers (Figure 6H). In conclusion, our cell tracing experiments demonstrate that aEPCs are derived from the epithelial epicardium upon injury and serve as the cellular source to regenerate all three epicardium-derived subsets, including mural cells and epithelial and mesenchymal epicardial cells.

298

299 **aEPCs are indispensable for heart regeneration**

300 To test the requirement of aEPCs for heart regeneration, we used the *ptx3a*^{RNTR} line to
301 ablate aEPCs. Since the aEPCs emerge at 1 dpa and their number peaks at 3 dpa, we
302 applied amputation injury and then bathed fish in 5 mM metronidazole for 3 successive
303 days from 3 to 5 dpa with fish water being changed daily[14]. Hearts were collected at 7
304 dpa for analyses (Figure 7A). Blood clots and large pieces of extra tissues were observed
305 in the aEPC ablated hearts (13 of 14 hearts analyzed), while there is only a minor
306 noticeable wound in the vehicle-treated NTR-positive group (12 of 18 hearts analyzed)
307 and the Mtz-treated NTR-negative group (11 of 16 hearts analyzed, Figure 7B).
308 Remarkably, *tcf21*$^{+}$ cells failed to repopulate the wound region after aEPC ablation,
309 indicating impaired wound closure (Figures 7C and 7D). CM proliferation, a hallmark of
310 heart regeneration, is largely suppressed after aEPC ablation (~54% reduction compared
311 to both control groups; Figures 7E and 7F). By 30 dpa, Acid Fushin Orange G (AFOG)
312 staining results indicate resolved fibrin and collagen deposition in the wound of vehicle-
313 treated and Mtz-treated NTR-negative hearts. In contrast, prominent scar tissue is
314 observed in 10 of 13 aEPC-ablated hearts, indicating failed regeneration (Figures 7G and
315 H). These results suggest that aEPCs are indispensable for successful heart
316 regeneration.

317

318 **aEPCs are the primary source of pro-regenerative epicardial progenies and**
319 **paracrine factors for regeneration**

320 Recent single-cell analyses of epicardium and epicardial-derived cells have identified pro-
321 regenerative subpopulations in the adult heart. To assess how aEPCs relate to these
322 subpopulations, we re-analyzed the published scRNA-seq datasets and the defined pro-
323 regenerative genes. Kapuria et al. performed scRNA-seq of FACS-isolated *pdgfrb*$^{+}$ mural
324 cells from injured adult hearts and found that epicardium-derived *pdgfrb*$^{+}$ mural cells are
325 essential for coronary development and heart regeneration (Figures S9A and S9B)[29]. We
326 demonstrated in our current study that at least part of these *pdgfrb*$^{+}$ mural cells in the
327 injury site are derived from aEPCs in adult hearts (Figure 5). Sun et al. recently profiled
328 single *tcf21*$^{+}$ cells isolated from adult hearts upon CM ablation[12]. They found that an

hapln1a$^+$ subset providing hyaluronic acid (HA) is required for heart regeneration and compact muscle development. We processed their dataset and checked the expression of aEPC markers. As shown in Figures S9C and S9D, *ptx3a* and *col12a1b* are expressed in the injured 7-day sample but are merely detectable in the uninjured one. *Ptx3a* is primarily expressed in their cluster 2, which was suggested by Sun et al. to give rise to the adjacent clusters, including the *hapln1a*-enriched clusters[12]. This analysis and our results indicate the progenitor property of *ptx3a*$^+$ epicardial cells in different heart injury models, and that the pro-regenerative *hapln1a*$^+$ epicardial cells are progenies of *ptx3a*$^+$ aEPCs. In addition, we previously reported a transcriptomic profile of 31 isolated *tcf21*$^+$ cells from uninjured adult hearts[54]. We found that *caveolin-1* (*cav1*), a pan-epicardial marker expressed in all 3 clusters, is required for heart regeneration. In our current dataset, *cav1* expression is broadly observed across clusters, although it is relatively higher in the epithelial and aEPC subpopulations (Figure S4). Similarly, De Bakkers et al. found that deletion of another epicardial gene paired related homeobox 1b (*prrx1b*), blocked heart regeneration with increased fibrosis[13]. In agreement with De Bakkers et al.'s result, we found that *prrx1a* is expressed in both the outermost and inner layers of the epicardium in our dataset, while *prrx1b* expression is much lower and only detected in a few cells (Figure S4). Our re-analyses indicate that *ptx3a*$^+$ cells are slightly reduced in the *prrx1b* mutant (Figures S9E-G). Thus, the pro-regenerative *cav1*$^+$ or *prrx1*$^+$ epicardial cells are broader populations that include the aEPCs and their progenies in the regenerating heart. These analyses define aEPCs as the primary cellular source of essential epicardial cell progenies for heart regeneration in zebrafish.

We have shown that aEPCs express published pro-regenerative factors such as *aldh1a2*, *fn1*, *fstl1*, *tmsb4x*, and *col12a1*[15, 33, 41, 42, 55, 56] (Figure S4). Expression of these factors is likely reduced upon aEPC ablation, which may contribute to regeneration defects. To further assess how aEPCs support myocardium regeneration, we checked the epicardium-derived mitogenic factor *nrg1*[40]. scRNA-seq data shows that *nrg1* expression is enriched in the mesenchymal epicardial cells (Figure 3A). We found that aEPC ablation significantly reduced *nrg1* expressing cells in the wound at 7 dpa (41% reduction compared to vehicle treatment and 43% reduction compared to the Mtz-treated NTR-negative group: Figures 8A and 8B). In addition, Sun et al. demonstrated that the

hapln1a$^+$ epicardial cells mediate HA deposition for myocardial regeneration[12]. As expected, aEPC ablation significantly reduced the number of *hapln1a*$^+$ mesenchymal epicardial cells in the wound by ~42% compared to the control groups at 7 dpa (Figures 8C and 8D). Thus, reduced Nrg1 signals and HA deposition in the aEPC-depleted heart likely contribute to reduced CM proliferation and observed regeneration defects. In summary, our results suggest that aEPCs are the primary cellular source of the essential epicardial cell progenies and paracrine factors required for successful heart regeneration. Nrg1, HA, and *hapln1a*$^+$ cells are among the downstream effectors of aEPC activation in supporting heart regeneration.

Tgfβ signaling regulates aEPC EMT and differentiation

We next asked how epicardial EMT regulates heart regeneration. The Tgfβ pathway is known as a regulator of EMT and has been reported to play important roles in zebrafish heart regeneration in epicardial cells after heart injury[57]. However, the underlying mechanism is not fully understood. Our scRNA-seq data indicate an injury-induced upregulation of *tgfb1a* in part of the aEPCs, which matches the expression pattern of *snai1a* at 3 dpa (Figure 9A). HCR staining showed co-expression of *tgfb1a* and *snail1a* in *col12a1b*$^+$ aEPCs in the wound at 3 dpa (Figure 9B). To further test the function of Tgfβ in aEPC EMT, we treated fish with SB431542 (a Tgfβ pathway inhibitor) after amputation injury and assessed heart regeneration at 7 dpa (Figure 9C)[42]. This treatment led to large blood clots in all hearts (7 of 7 hearts), while DMSO-treated hearts are largely normal with a minor noticeable wound (5 of 6 hearts; Figure 9D). We observed 55% and 53% reductions of *tcf21*$^+$ cells and *ptx3a*$^+$ cells, respectively, in the wound after SB431542 treatment. The thickness of the epicardial cell cap that covers the wound is also reduced by 71% on average (Figures 9E-H). Further HCR staining demonstrated a 68% reduction of *hapln1a*$^+$ mesenchymal epicardial cells entering the wound (Figures 9I and 9J), suggesting defects of epicardial differentiation and EMT. Thus, Tgfβ inhibition largely mimicked the aEPC depletion phenotypes. These results suggest that Tgfβ regulates EMT and mesenchymal cell differentiation of aEPCs, which are essential processes for heart regeneration.

Comparison with mouse epicardial cells upon myocardial infarction

To assess the similarities and differences between zebrafish and mouse epicardium, we analyzed a published scRNA-seq dataset of adult mouse epicardial cells (Figure S10). In the injured adult mouse heart, epicardial cells form a multi-cell layer in the wound[58]. Hesse et al. named these cells as epicardial stromal cells (EpiSC) and performed scRNA-seq of FACS-isolated EpiSC 5 days after myocardial infarction (MI)[10]. The dataset comprises 11 clusters that are separated into 3 groups: I, II, and III (Figure S10A). Cells in group I (expressing *Wt1*) are located in the outermost layer of the epicardium. Expression of group III markers are present throughout the activated epicardium but mostly in the inner layers of the epicardium. Group II has both epithelial clusters (expressing *Wt1*) and inner layer clusters and is enriched with ECM-related pathways[10]. We examined the expression of zebrafish cluster markers in the mouse dataset. As shown in Figures S10B-F, the homologs of zebrafish epithelial epicardium markers *Podxl*, *Sema3d*, and *Aldh1a2* are enriched in mouse EpiSC group I. The homologs of zebrafish aEPC makers *Ptx3*, *Col12a1*, Marcks, *Lox*, *Hop90b1*, *Serpinh1*, and *Tmsb4x* are primarily expressed in mouse EpiSC group II. The zebrafish mesenchymal or mural epicardium markers *Hapln1*, *Pdgfra*, and *Pdgfrb* are enriched in mouse EpiSC group III. In addition, zebrafish *fn1a*, *psmb1*, and *atp5mc1* are makers for aEPCs but are also expressed in part of the epithelial epicardium cluster (Figure 3E and S5). Similarly, the mouse homologs *Fn1*, *Psmb1*, and *Atp5g1* are highly expressed in EpiSC groups I and II (Figures S10B-F). Thus, mouse EpiSC groups I, II, and III are comparable to zebrafish epithelial, aEPC, and mesenchymal/mural subsets, respectively.

However, unlike in zebrafish, mouse epithelial epicardial cells do not give rise to mesenchymal EpiSCs in the infarct[10]. This observation was also supported by other studies[6, 59]. Although group II of the mouse EpiSC does express makers of zebrafish aEPCs, Hesse et al. study showed no differentiation trajectory from group II to cells in other groups. These differentiation deficiencies may contribute to the limited regenerative capacity of the adult mouse heart, which warrant further genetic studies in mice. In all, this comparison demonstrates both similarities and differences in epicardial populations between zebrafish and mice. It also implies that activating such a progenitor state in mice has the potential to promote cardiac repair.

DISCUSSION

Here we have defined the epithelial, mesenchymal, and mural subsets of the epicardial lineage in adult zebrafish. We identified the *ptx3a*$^{+}$*col12a1b*$^{+}$ epicardial cells as the adult progenitors - aEPCs. These aEPCs undergo EMT and orchestrate ECM remodeling and cell differentiation to mediate heart regeneration (Figure 9K). Thus, augmenting the aEPC activation after heart injury is of potential value for enhancing heart regeneration.

Our modRNA-assisted genetic tracing demonstrated an active epicardial EMT process, for the first time, during heart regeneration in adult zebrafish. It was reported that epicardial EMT occurs prior to fate specification in a chick heart development model[60]. Hampering epicardial EMT in mice abolishes epicardial lineages and leads to severe heart development defects[61, 62, 63]. In contrast to zebrafish, adult mammalian epicardial cells have no or limited EMT upon heart injuries[6, 59]. Our study further highlights that deficiency in epicardial EMT may contribute to the limited regenerative capacity of the adult mammalian heart. In addition, the pericardial sac injection of modRNA is a novel approach for mechanistic studies of the epicardium. It can be applied in multiple contexts in both zebrafish and mammals (e.g., gain- and loss-of-function assays for epicardial activation and EMT).

Previous work with the *tcf21:CreER*^{t2} transgenic line discovered that epicardial cells contribute to a large amount of perivascular (or mural) cells during zebrafish heart regeneration[17]. However, since the *tcf21* reporters also marks part of the *pdgfrb*$^{+}$ mural cells during regeneration, it was unclear whether these perivascular contributions originated from the existing *tcf21:*CreER^{+}*pdgfrb*$^{+}$ mural cells or epicardial progenitors. Our pseudotime analysis uncovers the existence of both trajectories, and we confirm experimentally that at least a subset of the mural cells in the injury site is derived from aEPCs. The *hapln1a*$^{+}$ mesenchymal subset is divided into multiple clusters, which indicates further heterogeneity. These subsets may have additional functions in supporting regeneration even if their relative proportion remains unchanged. Gene expression profiles suggest that some mesenchymal epicardial cells may have a fibroblast identity, which warrants further investigation. Our GO term analysis suggests a number of unique contributions that each core epicardial cluster makes during

453 regeneration. How these subsets coordinate with each other and with other cardiac cell
454 types to exert efficient regeneration warrants further investigation. A recent study by Sun
455 et al. demonstrated that the *hapln1a*$^{+}$ epicardial cells mediate HA secretion and
456 myocardial development and regeneration[12]. Thus, both studies suggest that the entire
457 epicardial population actively participated in the regeneration process with diverse cellular
458 and paracrine contributions.

459 We showed that genetic ablation of *col12a1b*-expressing aEPCs upon heart injury
460 led to the formation of collagen-enriched scar tissues at 30 dpa. This suggests that the
461 initial deposition of the pro-regenerative collagen XII[43, 44] may be necessary for
462 regeneration. Notably, it was reported that transient collagen deposition is required for
463 zebrafish heart regeneration in a cryoinjury model[36]. The pro-regenerative collagen
464 remodeling, in terms of the composition of collagen components and the associated
465 deposition timing warrants further investigation. Although it has been shown that Tgfβ
466 inhibition blocks heart regeneration[57], our results provide further cellular insights that
467 Tgfβ-regulated EMT and differentiation of aEPCs contribute to heart regeneration.
468 Spatiotemporal activity of the Tgfβ signaling may regulate dynamic ECM deposition.
469 Moreover, the function of *ptx3a* implies a major role of aEPCs in mediating inflammation
470 during regeneration[45]. Thus, our discovery suggests aEPCs as a molecular hub
471 connecting ECM remodeling and immune responses during regeneration. Our findings
472 also open new research avenues to precisely manipulate the regeneration program.

473 In mammals, the adult epicardium shows analogous activation upon heart injury
474 (such as re-activation of embryonic gene expression, proliferation, and secretion), but this
475 activation is limited in term of mitogen secretion, EMT, and differentiation capacity[6, 10, 59].
476 The similarities between zebrafish aEPCs and mouse epicardial-derived cells suggest
477 that awakening the progenitor potential in the adult mammalian epicardium could promote
478 cardiac repair after myocardial infarction. In all, our study has revealed the plasticity of
479 adult epicardial cells and highlighted the aEPCs as a target for enhancing cardiac
480 regeneration.
481
482

METHODS

Animal Maintenance and Procedures

Animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at Weill Cornell Medical College. Adult zebrafish of the Ekkwill and Ekkwill/AB strains were maintained as described[2, 64]. Water temperature was maintained at 28°C, and fish were kept on a 14/10 light/dark cycle at a density of 5-10 fish per liter. Animals between 3 and 12 months of both sexes were used for adult experiments. Heart resection injury was done as described previously[2]. For aEPC ablation, we applied amputation injury and then bathed fish in 5 mM metronidazole (Mtz, Sigma-Aldrich, M1547) for 3 successive days from 3 to 5 dpa with fish water being changed daily[14]. For lineage tracing, embryos and larvae were treated with 10 μM 4-Hydroxytamoxifen (4-HT, Sigma-Aldrich, H7904) in fish water for time periods as mentioned in the figures. Adult fish were placed in a mating tank of aquarium water containing 5 or 10 μM 4-HT as noted in the figures. Fish were maintained for 16 h, rinsed with fresh aquarium water, and returned to a recirculating aquatic system for 8 h, before repeating this incubation, as noted in the figures. For Tgfβ pathway inhibition, adult fish were incubated with 20 μM SB431542 as noted in the Figures with daily water changes[42]. modRNAs of Cre were synthesized as described previously[65]. Up to 2 μl of 10 mg/ml modRNA was injected into the pericardial sac of each fish without poking the heart under anesthesia as described previously[21] except using a 10 μl hamilton syringe.

The following previously published lines were used: *Tg(tcf21:nucEGFP)*^{pd41} (Ref. [17]), *Tg(tcf21:H2A-mCherry)*^{pd252} (Ref. [66]), *Tg(pdgfrb:EGFP)*^{ncv22} (Ref. [67]), and *Tg(ubi:loxP-EGFP-loxP-mCherry)*^{cz1701} (*ubi:Switch*) (Ref. [53]). Newly generated lines are described below. All reporters were analyzed as hemizygotes. Transgenic lines generated in this study are readily available on request.

Generation of knock-in lines

We applied a Crispr/Cas9 knock-in strategy to insert EGFP-poly A, or mScarlet-P2A-NTR-poly A cassette right after the start codon following a published protocol[68]. Briefly, 2~3 sgRNAs close to the ATG site were selected by using a design tool from Integrated DNA Technologies, Inc. HDR (Homology-directed repair) templates were designed to

include a mutated sgRNA target site, homology arms (HAs) flanking the ATG start codon, and a cassette encoding EGFP, or mScarlet-P2A-NTR with polyA. Gene-specific Alt-R crRNAs were synthesized by IDT. Bipartite synthetic sgRNAs were heteroduplexed by using crRNAs and a tracrRNA according to manufacturer recommendations. HDR sequences were synthesized by GENEWIZ and ligated to a pUC57 vector. HDR templates were digested from the pUC57 vector with flanked blunt-ends restriction enzymes and were column-purified. 250 ng/μl sgRNA, 250 ng/μl rCas9 (PNA Bio), and 50 ng/μl HDR templates were injected into one-cell stage embryos. Stable transgenic lines with seamless insertion alleles were identified by genotyping and sequencing of F1s. The following knock-in alleles were generated for this study:

Tg(ptx3a:mScarlet-P2A-NTR)

A mScarlet-P2A-NTR-poly A cassette was inserted right after the start codon using the following gRNAs and HAs. sgRNA1: AACAAGGAGGACCATCCAAG; sgRNA2: GTCTTGCTCATATTAACACA. 5'HA: 5'-agtACTTGCATTTAATACAGATATTCCCAAAA TGTTTCCAGAGCACTATTTCAAAATTCATTGCTATTTCCAGGTTTCCTTACATCTTTC CCAACATCCTCCCCCAAAATCTCTCTCTTTCTCTCTATGCCTGTGCTGCAGGAGCT GTACAGCAATATGACTACCCAAACTACCTCCAACGAAGAATTTCTCACTGGAGAAA CTCACTCTCTCCCCCACCACTCTCTCTCACTCTCTCTCTCTCTCTCTCACTCTCT CTCTCTCTCTCTCTCTCTCTCTCTCTCTCAGCGCAATGTTTAAATAGACCACAGA GCTGCTCTATCAGGCTCAGTTGCCTTTCAGCTGCAGTTCACCAAACAAGCAGGAGT CCAGCCAACAAATTGCTCAGCATCTA**tCAtGGtGGtCCAaCCAtGAcG**ACAAAGACAA TTATACACAAAAGCTGA**GTgTTcCaCAaATatACtC-3';** 3'HA: 5'-agtACTTATTGCATAC TCAAAACACAATAGTCAAATACAGTAATCAATCAGTTATTCTAATTTACATCATATCA TGCATCCTGTTAAACTACAAGAAATATTCCAGCATTTAACTCTTAACTTTTAACTTTT AACATAACTCCTTTCTGAAAGGTAGTGGAATTATACATCTCTTTATCTCACCTCTTAC TCTGATCACAGTTGAGTATTTCTTTCTGGTTACTCACCTTCCAACTGGTCCTCATCA GGTATTTCATTAAAATAATTGTGTCCAAAGCCCACTTCGATATCATCGTCATAATTGT CAGATACCGCCATGGACAATGAAGCCACCAGACACATGGTCTGGAGGAATGTCCG CATGGACTTGCTGGC-3'. The underlined sequences are the mutated gRNA binding

sites with the mutations shown in lower case letters. The full name of this line is *Tg(ptx3a:mScarlet-P2A-NTR)*^{wcm107} (*ptx3a*^{RNTR} for short).

Tg(col12a1b:EGFP)

An EGFP-poly A cassette were inserted right after the start codon using the following gRNAs and HAs. sgRNA1: CCTGACCGACATCTTCACCC; sgRNA2: CTGACATCAGTGAGAGCCCA, sgRNA3: AATTATTAGAGTCTTAACCA; 5'HA: 5'-cccGGGTTTTACAGTGCATGTAAAGTCTGAATTTGTCCAGGCACAACAGGGGGCGC AGAAGCCCTGCAGATGTTTTTCCATCCCTACCATGCAGGATCCTAAACCAGGCAGC TTTCCAGACAATCCCATATAGTCTATGAACTCACAGTTTCCAATGTTAAACTGATGA ATCAGCAGGGTTTGCGATGACATAAAGCATATGCAGCCTATAATTGTTTTTACAGTT GAACTAAAATGTTCTTTCTAATAACCGCACAGTGAATTATTTTAATTAATCCCCCTTT TATTAATATTTTCCAA**AAaTAcTtGtGaCTTtACgAAcG**AAAAAAAAAACTTTTCTTCTC TGACCAAAGAAGCAGCTGGGATTTTGTATGAAAATATAATTAAAATCAGATTTAATT GAGATCTCTCAATGCTGGAGTGGAGCATGTATTTTCCCCTTGATTTGTGAGCTGTG TGTGTGTGTGTGTGTGTGTGTGTTTGTGTGTTTGTGTGTCTGTGTGTAATCTGCT TGAATTTTCAGATTTGATTGATGTTCTGCTGATAC**CGGGGTGAAG**-3'; 3'HA: 5'-gatATCAATGCACATGAGGAATTATTTGGCTCCAGTCCTTTTGTTCATTCCTTGAACA AAGGCAATGAAAAATAATGTAAGCACTTAATATAATGGGCATCTATTACAGTAAAGT CGATATTTATAGAGTTTAAAACAGAAATCAGACTATGATCTAAACTGTTGGAAGAAA GTGGTTTAATTTACTGTAATTGGATGAATGGAAAGGCATTTGTTGGCTTTTTTTTTTT TTTTTA**CgTTGtGCcCgaACgGAaGTCAG**TGCTGCGAGTGTAAAGAGAGCCACCGCT GCCAAATG**CCgcACtGA-3.** The underlined sequences are the mutated gRNA binding sites with the mutations shown in lower case letters. The lines in use were generated by using sgRNA1. The full name of these lines are *Tg(col12a1b:EGFP)*^{wcm108} (*col12a1b*^{EGFP} for short).

Generation of *Tg(ptx3a:CreER*^{t2}*)* zebrafish

The translational start codon of *ptx3a* in the BAC clone CH211-272F3 was replaced with a *CreER*^{t2}*-polyA* cassette in SW105 cells using recombineering[69, 70]. The 5' and 3' HAs for recombination are 300~500-base pair (bp) fragments upstream and downstream of

the start codon and were PCR amplified with restriction enzyme site and inserted into a vector to flank the *CreER^{t2}-polyA* cassette. The primer sequences were: ptx3a-5HA-forward: gcggccgcAGTACTTGC ATTTAATACAGAT; ptx3a-5HA-reverse: gaattcGTGTTAATATGAGCAAGACTCA; ptx3a-3HA-forward: gggcccAGtACTTATTGCATACTCAAAAC; and ptx3a-3HA-reverse: GCCAGCAAGTCCATGCG. The same technology was used to replace the loxP site in the BAC vector with a cassette containing the Tol2 fragments, as well as a lens-specific crystallin promoter upstream of mCherry (iTol2Amp-γ-crystallin:RFP, a gift from Nadia Mercader Huber; Addgene plasmid # 108455)[71]. The final BAC was purified with Nucleobond BAC 100 kit (Clontech) and co-injected with 50 ng/μl Tol2 RNA into one-cell-stage zebrafish embryos. Stable transgenic lines were selected. The full name of this line is *Tg(ptx3a:CreER^{t2})^{wcm109}*.

Generation of *Tg(tcf21:loxP-BFP-pA-loxP-mCherry-NTR)* zebrafish

To make a BAC construct, the translational start codon of *tcf21* in the BAC clone DKEYP-79F12 was replaced with the *loxP-BFP-polyA-loxP-mCherry-NTR-polyA* cassette by Red/ET recombineering technology (Gene Bridges) [72]. The 5' and 3' homologous arms for recombination were a 50-base pair (bp) fragment upstream and downstream of the start codon and were included in PCR primers to flank the *loxP-BFP-polyA-loxP-mCherry-NTR-polyA* cassette. To avoid aberrant recombination between the insertion cassette and the endogenous *loxP* site in the BAC vector, we replaced the vector-derived *loxP* site with an I-Sce I site using the same technology. The final BAC was purified with Nucleobond BAC 100 kit (Clontech) and co-injected with I-Sce I into one-cell-stage zebrafish embryos. Stable transgenic lines with bright fluorescence were selected. The full name of this line is *Tg(tcf21:loxP-BFP-pA-loxP-mCherry-NTR)^{wcm110}* (*tcf21:BsRNTR* for short).

Single cell-RNA sequencing

Ventricles were collected from adult *tcf21:nucEGFP* fish at 6 months of age. Ventricular nucEGFP^{+} epicardial cells were isolated as described previously[54]. Briefly, ventricles were collected on ice and washed several times to remove blood cells. Ventricles were digested in an Eppendorf tube with 0.5 ml HBSS plus 0.13 U/ml Liberase DH (Roche) at

37°C, while stirring gently with a Spinbar® magnetic stirring bar (Bel-Art Products). Supernatants were collected every 5 min and neutralized with sheep serum. Dissociated cells were spun down and re-suspended in DMEM plus 10% fetal bovine serum (FBS) medium with 1.5 μg/ml propidium iodide (PI) and sorted using a Becton-Dickinson Aria II sorter for EGFP-positive and PI-negative cells. The isolated cells were sent to the Epigenomics Core Facility of Weill Cornell Medicine for single cell RNA-seq library preparation using the 10x Genomics Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3, and Chromium Single Cell B Chip Kit. The libraries were sequenced on a pair-end flow cell with a 2 x 50 cycles kit on Illumina HiSeq4000.

scRNA-seq analysis

The raw reads were aligned and processed with the CellRanger pipeline (v. 3.0.2) using the zebrafish transcriptome version GRCz10. Subsequent analyses were performed in R following the recommendations of Amezquita et al. (https://osca.bioconductor.org/)[73] using numerous functions provided in the R packages scater and scran[74, 75] as well as Seurat following the tutorials of the Satija Lab (https://satijalab.org/seurat/)[76]. We first removed low-quality droplets and rarely covered genes from all samples: cells were required to have a minimum of $10^{2.5}$ genes and maximum of 5% mitochondrial reads; and genes were removed if they were detected in either fewer than 0.2% of the cells or in fewer than 5 cells per sample. Read counts of 3 samples were normalized using *SCTransform* as implemented in Seurat v3.1 correcting for batch effects between the samples[77]. For visualizations and additional downstream analyses, the SCTransform-normalized (log-transformed) expression values were used unless noted otherwise.

For identifying clusters of cells with similar global transcriptomes, a shared nearest neighbor graph was constructed using Seurat's *FindNeighbors* function with default settings (e.g. k = 20) and using the first 20 principal components following PCA. Clusters were identified with Seurat's *FindClusters* function with the resolution parameter set to 0.3 (Ref. [78]). In addition, UMAP coordinates were calculated[79]. We assessed the identity of the cells within the resulting clusters using marker genes detected by Seurat's *FindAllMarkers* function with default settings as well as marker genes identified by SC3 (Ref. [80]). Gene Ontology (GO) analysis was performed by using *clusterProfiler*[81]. By

focusing on the core clusters, we re-did all processing steps including removal of genes that were expressed in fewer than 5 cells, calculation of normalized expression values correcting for the batch effect of the different conditions, PCA, clustering, and UMAP calculation.

To infer the developmental order of certain subpopulations within the regenerating samples, we applied the trajectory reconstruction algorithm Monocle 3 (Ref. [52]). We converted the read counts into a monocle object and re-processed the data using *preprocess_cds* with the number of dimensions set to 100. Cells were clustered with *cluster_cells* using the UMAP dimensions. To identify the trajectories of individual cells through the UMAP space, *learn_graph* was used. To determine pseudotime values, root nodes were identified for each partition (as determined in the previous step) and pseudotime values were calculated based on each cell's projection on the principal graph.

Histology and Microscopy

Freshly collected hearts were fixed with 4% paraformaldehyde (PFA) for 2 h at room temperature or overnight at 4C. Fixed hearts were mounted with Fluoromount G (Southern Biotechnology, cat#0100-01) between two coverslips for imaging of both ventricular surfaces, embedded in low-melting point agarose for whole-mount imaging, or applied to cryosection at a 10 μm thickness. Hybridization Chain Reaction (HCR 3.0) staining of whole-mounted hearts or cryosections was done following the published protocols[82]. HCR probes for *ptx3a*, *col12a1b*, *pdgfrb*, *hapln1a*, *podxl*, *atp5mc1*, *hmgb2b*, *hsp90b1*, *loxa*, *psmb1*, *serpinh1a*, *nrg1*, *snai1a*, and *tgfb1a* were synthesized by Molecular Instruments Inc. Immunostaining of whole-mounted hearts or heart sections was done as described previously[14, 66]. Primary antibodies used in this study include rabbit anti-Mef2 (this study), mouse anti-PCNA (Sigma, P8825), rabbit anti-Aldh1a2 (GeneTex, GTX124302), rabbit anti-DsRed (Takara, 632496), and mouse anti-Tnnt (ThermoFisher, MS-295-PABX). The Mef2 antibody was generated using a Mef2aa peptide (amino acid 314-512 of XP_021323249.1) as an antigen and is now commercially available at Boster Bio (DZ01398-1). Secondary antibodies (ThermoFisher) used in this study were Alexa Fluor 488 goat anti-rabbit and goat anti-mouse, Alexa Fluor 546 goat

anti-rabbit and goat anti-mouse, and Alexa Fluor 633 goat anti-mouse. Acid Fuchsin-Orange G staining was performed as described[2].

Bright-field images of whole-mounted hearts were captured using a Zeiss Axiozoom V16 microscope. Fluorescent images of whole-mounted and sectioned heart tissues were imaged used a Zeiss 800 confocal microscope. AFOG staining images were captured on a Leica Dmi8 compound microscope. Analyses of CM proliferation were performed as previously described by counting Mef2 and PCNA nuclei in wound sites[55].

Data collection and statistics

Clutchmates, or hearts collected from clutchmates, were randomized into different groups for each treatment. No animal or sample was excluded from the analysis unless the animal died during the procedure. All experiments were performed with at least 2 biological replicates. Sample sizes were chosen based on previous publications and experiment types and are indicated in each figure legend. All measurements were taken from distinct samples. All statistical values are displayed as Mean +/- Standard Deviation (s.d.). Sample sizes, statistical tests, and *P* values are indicated in the figures or the legends. All box plots show 5 elements: the minimum, lower quartile, median, upper quartile, and maximum values. Student's *t*-tests (two-tailed) were applied when normality and equal variance tests were passed. The Mann-Whitney Rank Sum test was used when these failed. Fisher's exact test was used where appropriate.

DATA AVAILABILITY

The scRNA-seq datasets have been deposited at NCBI's Gene Expression Omnibus under accession numbers GSE202836.

CODE AVAILABILITY

All scripts as well as the code and cell labels used for generating the scRNA-seq based figures can be found at https://github.com/abcwcm/Cao_Epicardium.

ACKNOWLEDGMENTS

We thank Adedeji A. Afolalu, Chaim Shapiro, Soji Hosten, and Chelsea Quaies for fish care, Naoki Mochizuki for the *Tg(pdgfrb:EGFP)*^{ncv22} line, Geoffrey Pitt, Junsu Kang, and Todd Evans for comments on the manuscript. This work was supported by Rudin Foundation fellowships to Y.X. and J.Y., a predoctoral training grant position (T32-HD060600) and a predoctoral fellowship (F31-HL158168) from National Institutes of Health (NIH) to S.D., a predoctoral training grant position in Stem Cell Biology and Regenerative Medicine from New York State Stem Cell Science program (NYSTEM) to B.P., American Heart Association (AHA) Career Development Award (AHA941434) to M.R.H., NIH grants (R01HL142768 and R01HL149137) to L.Z., AHA Career Development Award (18CDA34110108), Weill Cornell Start-up fund, and NIH grant (R01HL155607) to J.C.

AUTHOR CONTRIBUTIONS

Conceptualization, J.C.; Methodology, F.D., P.Z., M.R.H., L.Z., D.B., and J.C.; Investigation, Y.X., S.D., B.P., M.Q, J.Y., Y.C., F.D., and P.Z.; Resources, D.B., L.Z., and J.C.; Writing and editing, Y.X., S.D., and J.C.; Funding Acquisition, Y.X., S.D., L.Z., and J.C.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1. Kikuchi K*, et al.* Primary contribution to zebrafish heart regeneration by gata4(+) cardiomyocytes. *Nature* **464**, 601-605 (2010).

2. Poss KD, Wilson LG, Keating MT. Heart regeneration in zebrafish. *Science* **298**, 2188-2190 (2002).

3. Jopling C, Sleep E, Raya M, Marti M, Raya A, Izpisua Belmonte JC. Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **464**, 606-609 (2010).

4. Bertozzi A*, et al.* Is zebrafish heart regeneration "complete"? Lineage-restricted cardiomyocytes proliferate to pre-injury numbers but some fail to differentiate in fibrotic hearts. *Developmental biology* **471**, 106-118 (2021).

5. Cao J, Poss KD. The epicardium as a hub for heart regeneration. *Nat Rev Cardiol* **15**, 631-647 (2018).

6. Quijada P, Trembley MA, Small EM. The Role of the Epicardium During Heart Development and Repair. *Circ Res* **126**, 377-394 (2020).

7. Simoes FC, Riley PR. The ontogeny, activation and function of the epicardium during heart development and regeneration. *Development* **145**, (2018).

8. Bollini S*, et al.* Re-activated adult epicardial progenitor cells are a heterogeneous population molecularly distinct from their embryonic counterparts. *Stem cells and development* **23**, 1719-1730 (2014).

9. Lupu IE, Redpath AN, Smart N. Spatiotemporal Analysis Reveals Overlap of Key Proepicardial Markers in the Developing Murine Heart. *Stem Cell Reports* **14**, 770-787 (2020).

10. Hesse J*, et al.* Single-cell transcriptomics defines heterogeneity of epicardial cells and fibroblasts within the infarcted murine heart. *eLife* **10**, (2021).

11. Ma H*, et al.* Functional coordination of non-myocytes plays a key role in adult zebrafish heart regeneration. *EMBO Rep* **22**, e52901 (2021).

12. Sun J*, et al.* hapln1 Defines an Epicardial Cell Subpopulation Required for Cardiomyocyte Expansion During Heart Morphogenesis and Regeneration. *Circulation*, 101161CIRCULATIONAHA121055468 (2022).

13. de Bakker DEM*, et al.* Prrx1b restricts fibrosis and promotes Nrg1-dependent cardiomyocyte proliferation during zebrafish heart regeneration. *Development* **148**, (2021).

14. Wang J, Cao J, Dickson AL, Poss KD. Epicardial regeneration is guided by cardiac outflow tract and Hedgehog signalling. *Nature* **522**, 226-230 (2015).

15. Wei K*, et al.* Epicardial FSTL1 reconstitution regenerates the adult mammalian heart. *Nature* **525**, 479-485 (2015).

16. Bargehr J*, et al.* Epicardial cells derived from human embryonic stem cells augment cardiomyocyte-driven heart regeneration. *Nature biotechnology* **37**, 895-906 (2019).

17. Kikuchi K*, et al.* tcf21+ epicardial cells adopt non-myocardial fates during zebrafish heart development and regeneration. *Development* **138**, 2895-2902 (2011).

18. Lepilina A*, et al.* A dynamic epicardial injury response supports progenitor cell activity during zebrafish heart regeneration. *Cell* **127**, 607-619 (2006).

19. Missinato MA, Tobita K, Romano N, Carroll JA, Tsang M. Extracellular component hyaluronic acid and its receptor Hmmr are required for epicardial EMT during heart regeneration. *Cardiovascular research* **107**, 487-498 (2015).

20. Hadas Y, Katz MG, Bridges CR, Zangi L. Modified mRNA as a therapeutic tool to induce cardiac regeneration in ischemic heart disease. *Wiley Interdiscip Rev Syst Biol Med* **9**, (2017).

21. Bise T, Jazwinska A. Intrathoracic Injection for the Study of Adult Zebrafish Heart. *J Vis Exp*, (2019).

22. Ellett F, Pase L, Hayman JW, Andrianopoulos A, Lieschke GJ. mpeg1 promoter transgenes direct macrophage-lineage expression in zebrafish. *Blood* **117**, e49-56 (2011).

23. Shen YC*, et al.* The cytokine macrophage migration inhibitory factor (MIF) acts as a neurotrophin in the developing inner ear of the zebrafish, Danio rerio. *Developmental biology* **363**, 84-94 (2012).

24. Bertrand JY, Kim AD, Teng S, Traver D. CD41+ cmyb+ precursors colonize the zebrafish pronephros by a novel migration route to initiate adult hematopoiesis. *Development* **135**, 1853-1862 (2008).

25. Huang CJ, Tu CT, Hsiao CD, Hsieh FJ, Tsai HJ. Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* **228**, 30-40 (2003).

26. Sehnert AJ, Huq A, Weinstein BM, Walker C, Fishman M, Stainier DY. Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. *Nat Genet* **31**, 106-110 (2002).

27. Lawson ND, Weinstein BM. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Developmental biology* **248**, 307-318 (2002).

28. Jin SW, Beis D, Mitchell T, Chen JN, Stainier DY. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132**, 5199-5209 (2005).

29. Kapuria S*, et al.* Heterogeneous pdgfrb+ cells regulate coronary vessel development and revascularization during heart regeneration. *Development* **149**, (2022).

30. Ando K*, et al.* Conserved and context-dependent roles for pdgfrb signaling during zebrafish vascular mural cell development. *Developmental biology* **479**, 11-22 (2021).

31. Shih YH, Portman D, Idrizi F, Grosse A, Lawson ND. Integrated molecular analysis identifies a conserved pericyte gene signature in zebrafish. *Development* **148**, (2021).

32. Ando K, Ishii T, Fukuhara S. Zebrafish Vascular Mural Cell Biology: Recent Advances, Development, and Functions. *Life (Basel)* **11**, (2021).

33. Cao Y*, et al.* Identification of enhancer regulatory elements that direct epicardial gene expression during zebrafish heart regeneration. *Development* **149**, (2022).

34. Weinberger M, Simoes FC, Patient R, Sauka-Spengler T, Riley PR. Functional Heterogeneity within the Developing Zebrafish Epicardium. *Dev Cell* **52**, 574-590 e576 (2020).

35. Karra R, Foglia MJ, Choi WY, Belliveau C, DeBenedittis P, Poss KD. Vegfaa instructs cardiac muscle hyperplasia in adult zebrafish. *Proc Natl Acad Sci U S A* **115**, 8805-8810 (2018).

36. Sanchez-Iranzo H*, et al.* Transient fibrosis resolves via fibroblast inactivation in the regenerating zebrafish heart. *Proc Natl Acad Sci U S A* **115**, 4188-4193 (2018).

37. Navis A, Marjoram L, Bagnat M. Cftr controls lumen expansion and function of Kupffer's vesicle in zebrafish. *Development* **140**, 1703-1712 (2013).

38. Meder D, Shevchenko A, Simons K, Fullekrug J. Gp135/podocalyxin and NHERF-2 participate in the formation of a preapical domain during polarization of MDCK cells. *J Cell Biol* **168**, 303-313 (2005).

39. Ivey MJ, Kuwabara JT, Riggsbee KL, Tallquist MD. Platelet-derived growth factor receptor-alpha is essential for cardiac fibroblast survival. *American journal of physiology Heart and circulatory physiology* **317**, H330-H344 (2019).

40. Gemberling M, Karra R, Dickson AL, Poss KD. Nrg1 is an injury-induced cardiomyocyte mitogen for the endogenous heart regeneration program in zebrafish. *eLife* **4**, (2015).

41. Wang J, Karra R, Dickson AL, Poss KD. Fibronectin is deposited by injury-activated epicardial cells and is necessary for zebrafish heart regeneration. *Developmental biology* **382**, 427-435 (2013).

42. Marro J, Pfefferli C, de Preux Charles AS, Bise T, Jazwinska A. Collagen XII Contributes to Epicardial and Connective Tissues in the Zebrafish Heart during Ontogenesis and Regeneration. *PLoS One* **11**, e0165497 (2016).

43. Wehner D*, et al.* Wnt signaling controls pro-regenerative Collagen XII in functional spinal cord regeneration in zebrafish. *Nat Commun* **8**, 126 (2017).

44. Tsata V*, et al.* A switch in pdgfrb(+) cell-derived ECM composition prevents inhibitory scarring and promotes axon regeneration in the zebrafish spinal cord. *Dev Cell* **56**, 509-524 e509 (2021).

45. Doni A, Mantovani A, Bottazzi B, Russo RC. PTX3 Regulation of Inflammation, Hemostatic Response, Tissue Repair, and Resolution of Fibrosis Favors a Role in Limiting Idiopathic Pulmonary Fibrosis. *Front Immunol* **12**, 676702 (2021).

46. Veldman MB, Bemben MA, Thompson RC, Goldman D. Gene expression analysis of zebrafish retinal ganglion cells during optic nerve regeneration identifies KLF6a and KLF7a as important regulators of axon regeneration. *Developmental biology* **312**, 596-612 (2007).

47. Sousounis K*, et al.* Molecular signatures that correlate with induction of lens regeneration in newts: lessons from proteomic analysis. *Hum Genomics* **8**, 22 (2014).

48. Cappuzzello C*, et al.* Mesenchymal Stromal Cell-Derived PTX3 Promotes Wound Healing via Fibrin Remodeling. *J Invest Dermatol* **136**, 293-300 (2016).

49. Doni A*, et al.* An acidic microenvironment sets the humoral pattern recognition molecule PTX3 in a tissue repair mode. *The Journal of experimental medicine* **212**, 905-925 (2015).

50. Salio M*, et al.* Cardioprotective function of the long pentraxin PTX3 in acute myocardial infarction. *Circulation* **117**, 1055-1064 (2008).

51. Bock-Marquette I*, et al.* Thymosin beta4 mediated PKC activation is essential to initiate the embryonic coronary developmental program and epicardial progenitor cell activation in adult mice in vivo. *Journal of molecular and cellular cardiology* **46**, 728-738 (2009).

52. Cao J*, et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496-502 (2019).

53. Mosimann C, Kaufman CK, Li P, Pugach EK, Tamplin OJ, Zon LI. Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Development* **138**, 169-177 (2011).

54. Cao J*, et al.* Single epicardial cell transcriptome sequencing identifies Caveolin 1 as an essential factor in zebrafish heart regeneration. *Development* **143**, 232-243 (2016).

55. Kikuchi K*, et al.* Retinoic acid production by endocardium and epicardium is an injury response essential for zebrafish heart regeneration. *Dev Cell* **20**, 397-404 (2011).

56. Bollini S, Riley PR, Smart N. Thymosin beta4: multiple functions in protection, repair and regeneration of the mammalian heart. *Expert Opin Biol Ther* **15 Suppl 1**, S163-174 (2015).

57. Chablais F, Jazwinska A. The regenerative capacity of the zebrafish heart is dependent on TGFbeta signaling. *Development* **139**, 1921-1930 (2012).

58. van Wijk B, Gunst QD, Moorman AF, van den Hoff MJ. Cardiac regeneration from activated epicardium. *PLoS One* **7**, e44692 (2012).

59. Zhou B*, et al.* Adult mouse epicardium modulates myocardial injury by secreting paracrine factors. *The Journal of clinical investigation* **121**, 1894-1904 (2011).

60. Mantri M*, et al.* Spatiotemporal single-cell RNA sequencing of developing chicken hearts identifies interplay between cellular differentiation and morphogenesis. *Nat Commun* **12**, 1771 (2021).

61. Jackson-Weaver O*, et al.* PRMT1-p53 Pathway Controls Epicardial EMT and Invasion. *Cell Rep* **31**, 107739 (2020).

62. Liu X*, et al.* Wdpcp promotes epicardial EMT and epicardium-derived cell migration to facilitate coronary artery remodeling. *Sci Signal* **11**, (2018).

63. Singh A*, et al.* Hippo Signaling Mediators Yap and Taz Are Required in the Epicardium for Coronary Vasculature Development. *Cell Rep* **15**, 1384-1393 (2016).

64. Wang J*, et al.* The regenerative capacity of zebrafish reverses cardiac failure caused by genetic cardiomyocyte depletion. *Development* **138**, 3421-3430 (2011).

65. Sultana N, Sharkar MTK, Hadas Y, Chepurko E, Zangi L. In Vitro Synthesis of Modified RNA for Cardiac Gene Therapy. *Methods in molecular biology* **2158**, 281-294 (2021).

66. Cao J*, et al.* Tension Creates an Endoreplication Wavefront that Leads Regeneration of Epicardial Tissue. *Dev Cell* **42**, 600-615 e604 (2017).

67. Ando K*, et al.* Clarification of mural cell coverage of vascular endothelial cells by live imaging of zebrafish. *Development* **143**, 1328-1339 (2016).

68. DiNapoli SE*, et al.* Synthetic CRISPR/Cas9 reagents facilitate genome editing and homology directed repair. *Nucleic acids research* **48**, e38 (2020).

69. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. Simple and highly efficient BAC recombineering using galK selection. *Nucleic acids research* **33**, e36 (2005).

70. Suster ML, Sumiyama K, Kawakami K. Transposon-mediated BAC transgenesis in zebrafish and mice. *BMC genomics* **10**, 477 (2009).

71. Sanchez-Iranzo H*, et al.* Tbx5a lineage tracing shows cardiomyocyte plasticity during zebrafish heart regeneration. *Nat Commun* **9**, 428 (2018).

72. Singh SP, Holdway JE, Poss KD. Regeneration of amputated zebrafish fin rays from de novo osteoblasts. *Dev Cell* **22**, 879-886 (2012).

73. Amezquita RA*, et al.* Orchestrating single-cell analysis with Bioconductor. *Nature methods* **17**, 137-145 (2020).

74. McCarthy DJ, Campbell KR, Lun AT, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179-1186 (2017).

75. Lun AT, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res* **5**, 2122 (2016).

76. Stuart T*, et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).

77. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, 296 (2019).

78. Waltman L, Eck NJv. A smart local moving algorithm for large-scale modularity-based community detection. *European Physical Journal B*, (2013).

79. Becht E*, et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nature biotechnology*, (2018).

80. Kiselev VY*, et al.* SC3: consensus clustering of single-cell RNA-seq data. *Nature methods* **14**, 483-486 (2017).

81. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).

82. Choi HMT*, et al.* Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, (2018).

FIGURE LEGENDS

Figure 1. Cellular heterogeneity and EMT of *tcf21*$^+$ cells in the zebrafish heart.

(**A**) Cartoon of an adult zebrafish heart showing the ventricle and outflow tract (OFT). The frames indicate representative regions for cryosection-section images in B, C and E. (**B**) A cryosection image showing *tcf21:*nucEGFP in magenta, antibody staining against aldh1a2 in green. Nuclei were stained with DAPI (blue). A single-channel image of aldh1a2 signals is shown in grayscale on the right. Arrowheads indicate EGFP$^+$adh1a2$^+$ cells. Scale bar, 50 μm. (**C**) A cryosection image showing *tcf21:*nucEGFP in magenta, antibody staining against Tnnt in green. Nuclei were stained with DAPI (blue). A single-channel image of EGFP is shown in grayscale on the right. Arrows indicate EGFP$^+$ mesenchymal epicardial cells. Scale bar, 50 μm. (**D**) Schematic of experimental design for modRNA injection. (**E**) Section images of uninjured (left) and 7 dpa (right) hearts carrying the *tcf21:Switch* reporter at 10 days post Cre modRNA injection. Arrows and arrowheads indicate representative mCherry$^+$ mesenchymal and epithelial cells, respectively. Scale bar, 50 μm. (**F**) Quantification of mCherry$^+$ mesenchymal epicardial cells in the experiment of E. The largest cryosection of each heart was quantified for mCherry$^+$ cells in the apex half of the ventricle. n = 13 (Ctrl) and 6 (7 dpa), respectively. Student's *t* test.

Figure 2. ScRNA-seq reveals distinct subsets of the ventricular epicardial lineage during heart regeneration.

(**A**) Experimental design for single epicardial cell isolation and transcriptome sequencing. (**B**) UMAP of three samples combined. After the removal of droplets with very few genes as well as genes that could not be detected in at least 5 cells, the final dataset comprised 4,970, 3,743, and 6,428 cells for Ctrl, 3 dpa, and 7 dpa, respectively. (**C**) UMAP of cell clusters with inferred cellular identities. (**D**) Expression of top 5 cluster marker genes across different clusters. This dot plot depicts the abundance and expression magnitude of individual genes across cells of given clusters. The size of the dot represents the fraction of cells with at least one UMI of the specified gene. (**E**) UMAPs showing changes of clusters across samples. (**F**) Proportions of the samples per cluster. (**G**) Cluster

proportions per sample. (**H**) Normalized expression of top marker genes on UMAPs (3 samples combined).

Figure 3. The epithelial and mesenchymal subsets of the epicardium.

(**A**) Normalized expression of marker genes on UMAPs (3 samples combined). (**B**) Heatmap of 3 marker genes: *tcf21*, *aldh1a2*, and *hapln1a*. Normalized expression values are shown; cells were sorted by cluster membership. Clusters 2 and 5 are highlighted in red frames. (**C**) Images of heart sections showing *tcf21:*nucEGFP in magenta, HCR staining of *podxl* in green. Nuclei were stained with DAPI (blue). A single-channel image of *podxl* is shown in grayscale on the right. Arrows indicate representative EGFP$^+$*podxl*$^+$ cells. Scale bar, 50 μm. (**D**) Images of heart sections showing *tcf21:*nucEGFP in magenta, HCR staining of *hapln1a* in green. Nuclei were stained with DAPI (blue). A single-channel image of *hapln1a* is shown in grayscale on the right. Arrows indicate representative EGFP$^+$*hapln1a*$^+$ cells. Scale bar, 50 μm. (**E**) Normalized expression levels of marker genes for the 3 dpa-specific cluster on UMAPs (3 samples combined).

Figure 4. *ptx3a* and *col12a1b* label a transient pro-regenerative epicardial subtype.

(**A-E**) HCR staining results of *ptx3a* (green) on heart sections collected at 1 (B), 3 (C), 7 (D), and 14 dpa (E) together with the uninjured control (Ctrl, A). *tcf21*:nucEGFP (magenta) labels the epicardial cells. Nuclei were stained with DAPI (blue). Single-channel images show signals of *ptx3a* or nucEGFP. White dashed lines indicate the injury sites. The framed regions are enlarged to show details on the right of each panel. Arrows denote representative *ptx3a*$^+$EGFP$^+$ cells. Scale bars, 100 μm. (**F**) Schematic for generating the knock-in alleles for *ptx3a*. The gRNA binding site is marked with a cyan arrow. (**G**) The *ptx3a*^{RNTR} reporter recapitulates expression of *ptx3a* in the injured heart (3 dpa). No conclusive epicardial expression was observed in the uninjured heart (Ctrl). *tcf21*:nucEGFP (magenta) labels the epicardial cells. Nuclei were stained with DAPI (blue). Single-channel images show signals of *ptx3a*^{RNTR} (with anti-DsRed antibody staining) at the bottom. White dashed lines indicate the injury sites. The framed regions are enlarged to show details on the right. Arrows denote representative RNTR$^+$EGFP$^+$ cells. Images of additional timepoints are in Figure S7. Scale bar, 100 μm. (**H**) Schematic

for generating the knock-in alleles for *col12a1b*. The gRNA binding site is marked with a cyan arrow. (**I**) Section images showing *col12a1b*^{EGFP} reporter expression in green and *ptx3a*^{RNTR} expression in magenta (with anti-DsRed antibody staining). No epicardial EGFP expression was observed in the uninjured heart (left). Nuclei were stained with DAPI (blue). Single-channel images show signals of *col12a1b*^{EGFP} or *ptx3a*^{RNTR}. White dashed lines indicate the injury sites. The framed regions are enlarged to show details on the right. Arrows denote representative EGFP$^+$mCherry$^+$ cells. Scale bar, 100 μm.

Figure 5. aEPCs give rise to *pdgfrb*$^+$ mural cells during regeneration.

(**A, B**) Cell trajectories suggested by pseudotime analysis with Monocle 3. Shown on a UMAP, the starting point of each trajectory was labeled with a number. The pseudotime is shown as a heatmap in (A). Red arrows and letters highlight different branches in (B). Clusters are labeled in the same number and color as in Figure 2C. (**C**) Schematic of transgenic lines and experimental design to define the origin of aEPCs. 4HT was used at 10 μM. (**D**) Section images of 3 dpa hearts carrying the *tcf21:Switch; tcf21:CreER*^{t2}*; col12a1b*^{EGFP} reporters. 4HT treatment was performed as indicated in (C). mCherry and EGFP are shown in magenta and green, respectively in the merged image. Single-channel images are shown in grayscale. Arrows indicate representative EGFP$^+$mCherry$^+$ cells. White dashed lines indicate the injury sites. Scale bar, 100 μm. (**E**) *fn1a* expression shown on the pseudotime UMAP. The mural cell trajectories are circled with red dashed lines. (**F**) Whole-mount images of the ventricular surface showing expression of *pdgfrb:EGFP* (green) and *tcf21:H2R* (magenta) in transgenic lines at 7 dpa. The framed regions are enlarged to show details on the right, with single-channel images shown in grayscale. Arrows indicate representative EGFP$^+$mCherry$^+$ cells. Scale bar, 100 μm. (**G**) Schematic of transgenic lines and experimental design to define the fate of aEPCs. 4HT was used at 5 μM. (**H**) Whole-mount images of the ventricular surface from hearts carrying the *ubi:Switch; ptx3a:CreER*^{t2} reporters. *pdgfrb* expression is detected by HCR staining (green). mCherry is shown in magenta, and nuclei were stained with DAPI (blue). Arrows indicate representative *pdgfrb*$^+$mCherry$^+$ cells. A maximum projection image is shown on the left. Z-stack images of the lettered frames are shown on the right. Scale bar, 50 μm.

Figure 6. aEPCs give rise to the epithelial and mesenchymal epicardial cells.

(**A-C**) Section images showing HCR staining results of *hapln1a* (green) and *ptx3a* (magenta) at 3 (A), 7 (B), and 14 dpa (C). Epicardial cells are labeled with *tcf21*:nucEGFP (blue). Arrows in A-C indicate representative *hapln1a*$^{+}$ cells. The brackets in B outline the regenerated regions that express *hapln1a*. White dashed lines indicate the injury sites. Scale bars, 50 μm. (**D**) Schematic of transgenic lines and experimental design to define the fate of aEPCs. 4HT was used at 5 μM. (**E**) Orthogonal view of a z-stack image showing the ventricular surface layers from hearts carrying the *ubi:Switch;ptx3a:CreER*^{t2} reporters. A maximum projection image of the x-y plane is shown with *hapln1a* expression (HCR staining) in green and mCherry in magenta. White lines and numbers indicate positions for views of the y-z planes (right) and the x-z planes (top), respectively. Nuclei were stained with DAPI (blue) and was omitted from the maximum projection image to keep image clarity. Arrows indicate representative *hapln1a*$^{+}$mCherry^{+} cells. Arrowheads indicate mCherry^{+} cells in the epithelial layer. Scale bar, 50 μm. The framed regions are enlarged to show details in F. (**F**) Optical section (z-stack) images of the framed regions in E. Arrows indicate *hapln1a*$^{+}$mCherry^{+} cells. Scale bar, 20 μm. (**G**) Orthogonal view of a z-stack image showing the ventricular surface layers from hearts carrying the *ubi:Switch;ptx3a:CreER*^{t2} reporters. An optical section image (x-y plane) is shown at the bottom, and the x-z plane of z-stacks in shown on top. mCherry and HCR staining signals of *podxl* are shown in magenta and green, respectively. The white line indicates position for the view of the x-z plane. Arrowheads indicate *podxl*$^{+}$mCherry^{+} epithelial epicardial cells. (**H**) A UMAP highlights pseudotime trajectories a and c.

Figure 7. aEPC ablation blocks heart regeneration.

(**A**) Experimental design. Siblings carrying the *ptx3a*^{RNTR} allele and /or the *tcf21:nucEGFP* reporter were treated with 5 mM Mtz or vehicle (Ctrl) from 3 dpa to 5 dpa. (**B**) Whole mount images of hearts collected at 7 dpa. Dash lines denote the injury sites. Large blood clot and extra tissue were observed in hearts of Mtz treated NTR^{+} animals (13 of 14) but not in those from the vehicle-treated NTR^{+} fish (12 of 18) or Mtz treated NTR^{-} fish (11 of 16). Scale bar, 200 μm. (**C**) Section images of injured ventricles from 3 treatment groups at 7 dpa. The epicardial cells are labeled with *tcf21:*nucEGFP (green). Nuclei were stained

with DAPI (magenta). White dashed lines indicate the injury sites. Scale bar, 100 μm. (**D**) Quantification of EGFP$^+$ cells in the wound region from experiments in (C). From left to right, n = 8, 8, and 7, respectively. NS, not significant. Student's *t* test. (**E**) Section images of injured ventricles from 3 treatment groups at 7 dpa. Ventricular CM proliferation was assessed by anti-PCNA (green) and Mef2 (magenta) staining. Nuclei were stained with DAPI (blue). The framed regions are enlarged to show proliferating CMs (some denoted with arrowheads). (**F**) Quantified PCNA$^+$ CM indices in injury sites in experiments from (E). From left to right, n = 16, 16, and 14, respectively. NS, not significant. Student's *t* test. (**G**) Section images of ventricles at 30 dpa stained with Acid Fuchsin-Orange G to characterize non-muscle components in the injuries (blue for collagen, red for fibrin). (**H**) Semiquantitative assessment of cardiac injuries based on muscle and scar morphology (robust, partial, or blocked regeneration). Data were analyzed using Fisher's exact test. n = 13 for each treatment group. NS, not significant.

Figure 8. aEPC ablation reduced *nrg1* and *hapln1a* expression in the wound.

(**A**) Section images of the injury site showing HCR staining of *nrg1* in green and *tcf21*:nucEGFP in magenta. The framed regions are enlarged to show details with arrowheads denoting representative *nrg1*$^+$EGFP$^+$ cells. Scale bar, 50 μm. (**B**) Quantification of *nrg1*$^+$ epicardial cells in the wound regions shown in (A). From left to right, n = 8, 8, and 9, respectively. NS, not significant. Student's *t* test. (**C**) Section images of the injury site showing HCR staining of *hapln1a* in green and *tcf21*:nucEGFP in magenta. Arrowheads indicate representative *hapln1a*$^+$EGFP$^+$ cells. Scale bar, 50 μm. (**D**) Quantification of *hapln1a*$^+$EGFP$^+$ cells in the wound regions shown in (C). From left to right, n = 9, 10, and 13, respectively. NS, not significant. Student's *t* test.

Figure 9. Tgfβ signaling regulates aEPC EMT and differentiation.

(**A**) UMAPs showing *tgfb1a* and *snai1a* expression across samples. The aEPC population (cluster 5) is highlighted in frames. (**B**) Section images of the injury site showing HCR staining signals of *tgfb1a* and *snai1a* at 3 dpa in green and magenta, respectively. *col12a1b*^{EGFP} is shown in blue. The white dash line denotes the injury site. The framed region is enlarged to show details on the right with different channel combinations. DAPI

staining is shown in white in the last panel. Arrowheads indicate representative *tgfb1a*$^{+}$*snai1a*$^{+}$EGFP^{+} cells. (**C**) Experimental design for SB431542 treatment. (**D**) Whole-mount images of hearts collected at 7 dpa. Dash lines denote the injury sites. Large blood clot and extra tissue were observed in SB431542-treated hearts (7 of 7), but not in those from the DMSO-treated fish (5 of 6). Scale bar, 200 μm. (**E**) Section images of injured ventricles at 7 dpa. The epicardial cells are labeled with *tcf21:*nucEGFP (magenta) and HCR staining signals of *ptx3a* in green. Nuclei were stained with DAPI (blue). White dashed lines indicate the injury sites. The framed regions are enlarged to show details with arrowheads indicating representative *ptx3a*$^{+}$EGFP^{+} cells. Scale bar, 100 μm. (**F-H**) Quantifications of t*cf21*:nucEGFP^{+} nuclei in the wound region (F), average thickness of the epicardial cap covering the wound (G), and number of *ptx3a*$^{+}$ cells (H) from experiments in (E). n = 6 (DMSO) and 7 (SB431542), respectively, for each quantification. Student's *t* test. (**I**) Section images of the injury site showing HCR staining of *hapln1a* in green and *tcf21*:nucEGFP in magenta. The framed regions are enlarged to show details with arrowheads indicating representative *hapln1a*$^{+}$EGFP^{+} cells. Scale bar, 100 μm. (**J**) Quantification of *hapln1a*$^{+}$EGFP^{+} cells in the wound regions shown in (I). n = 6 (DMSO) and 7 (SB431542), respectively. Student's *t* test. (**K**) The working model. For simplicity, mesenchymal epicardial cells away from the injury site are omitted. The dashed lines indicate predicted mechanisms.

Figure 1

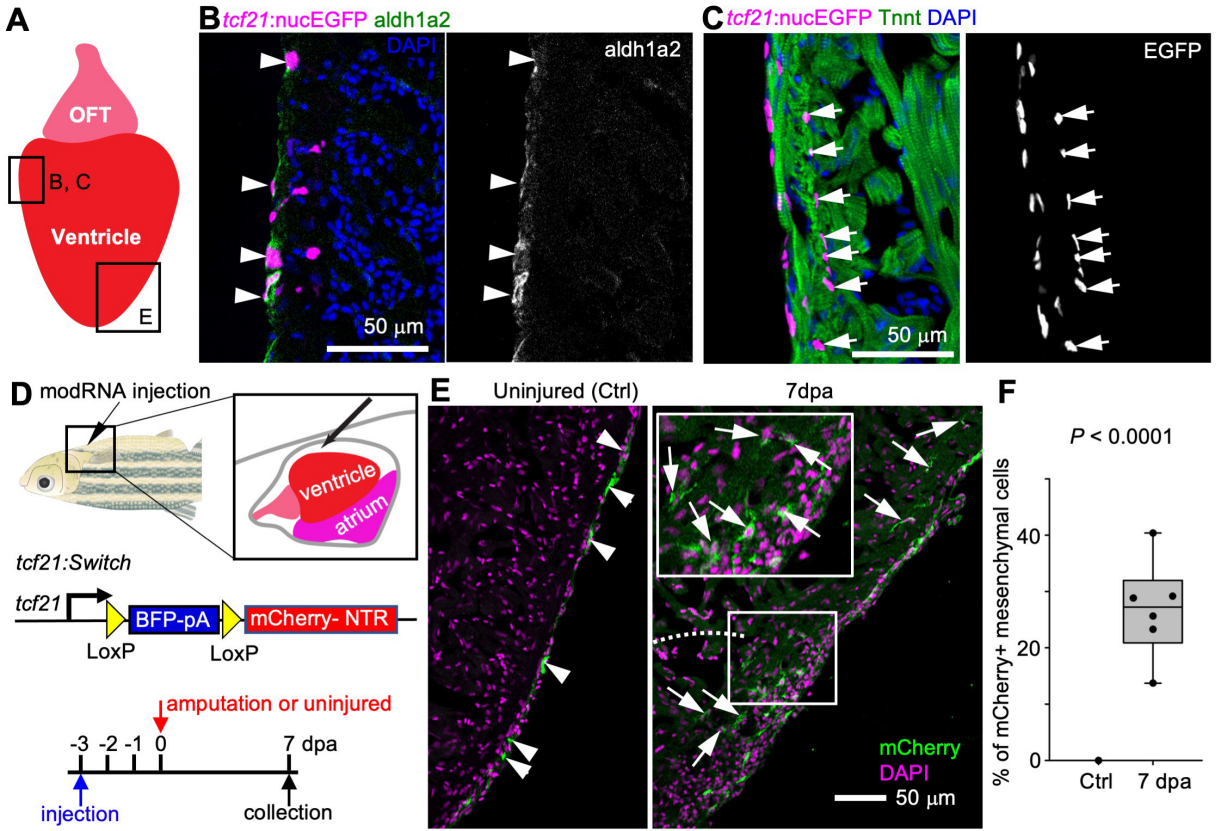

A
OFT
B, C
Ventricle
E
B tcf21:nucEGFP aldh1a2
DAPI
aldh1a2
50 μm
C tcf21:nucEGFP Tnnt DAPI
EGFP
50 μm
D modRNA injection
ventricle
atrium
tcf21:Switch
tcf21
BFP-pA
mCherry- NTR
LoxP
LoxP
amputation or uninjured
-3 -2 -1 0
7 dpa
injection
collection
E
Uninjured (Ctrl)
7dpa
mCherry
DAPI
50 μm
F
P < 0.0001
% of mCherry+ mesenchymal cells
40
20
0
Ctrl
7 dpa

Figure 2

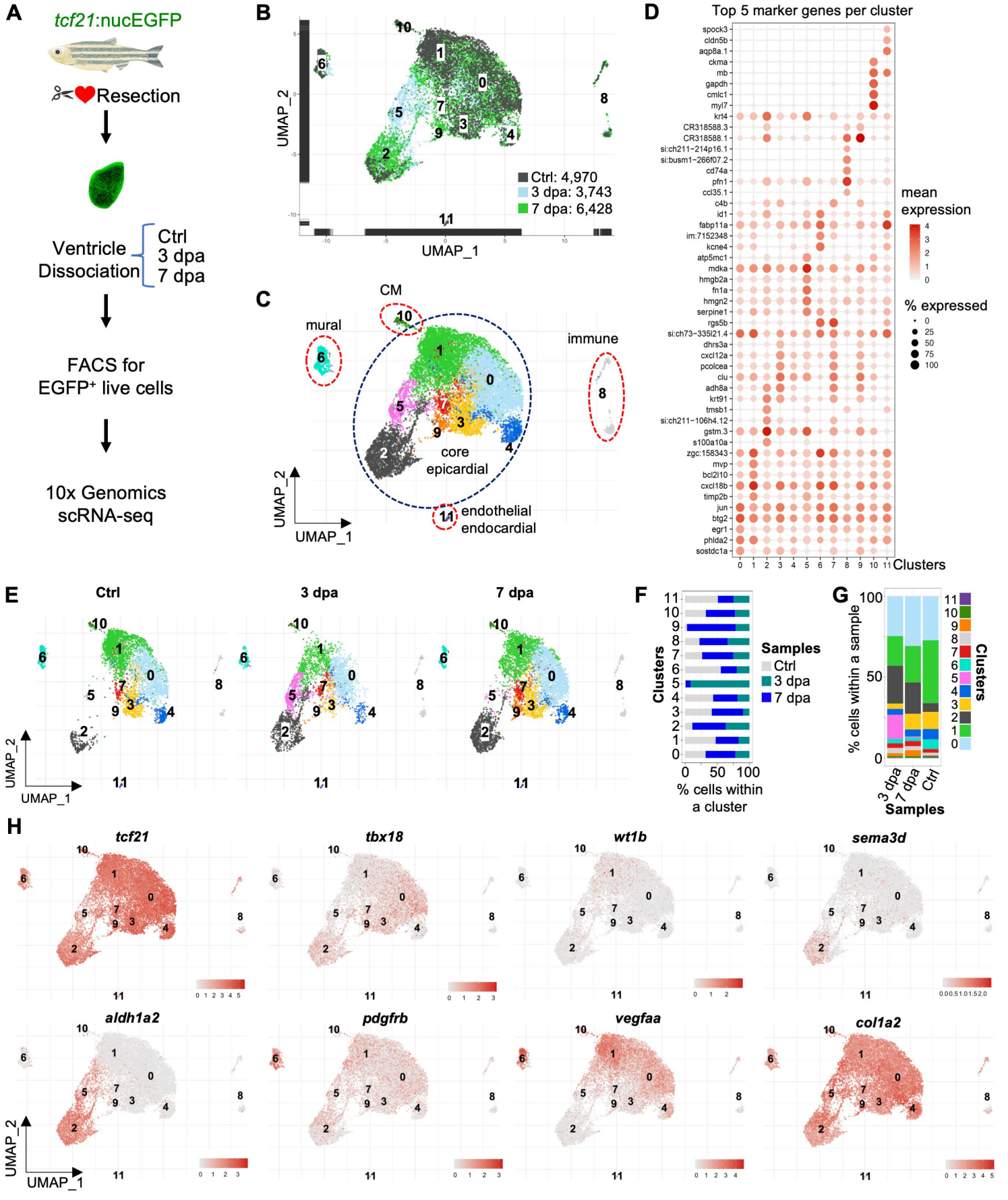

A
tcf21:nucEGFP
Resection
Ventricle Dissociation
Ctrl
3 dpa
7 dpa
FACS for EGFP+ live cells
10x Genomics scRNA-seq
B
UMAP_2
UMAP_1
Ctrl: 4,970
3 dpa: 3,743
7 dpa: 6,428
C
CM
mural
immune
core epicardial
endothelial endocardial
D
Top 5 marker genes per cluster
mean expression
% expressed
Clusters
E
Ctrl
3 dpa
7 dpa
F
Clusters
Samples
Ctrl
3 dpa
7 dpa
% cells within a cluster
G
% cells within a sample
Samples
H
tcf21
tbx18
wt1b
sema3d
aldh1a2
pdgfrb
vegfaa
col1a2

Figure 3

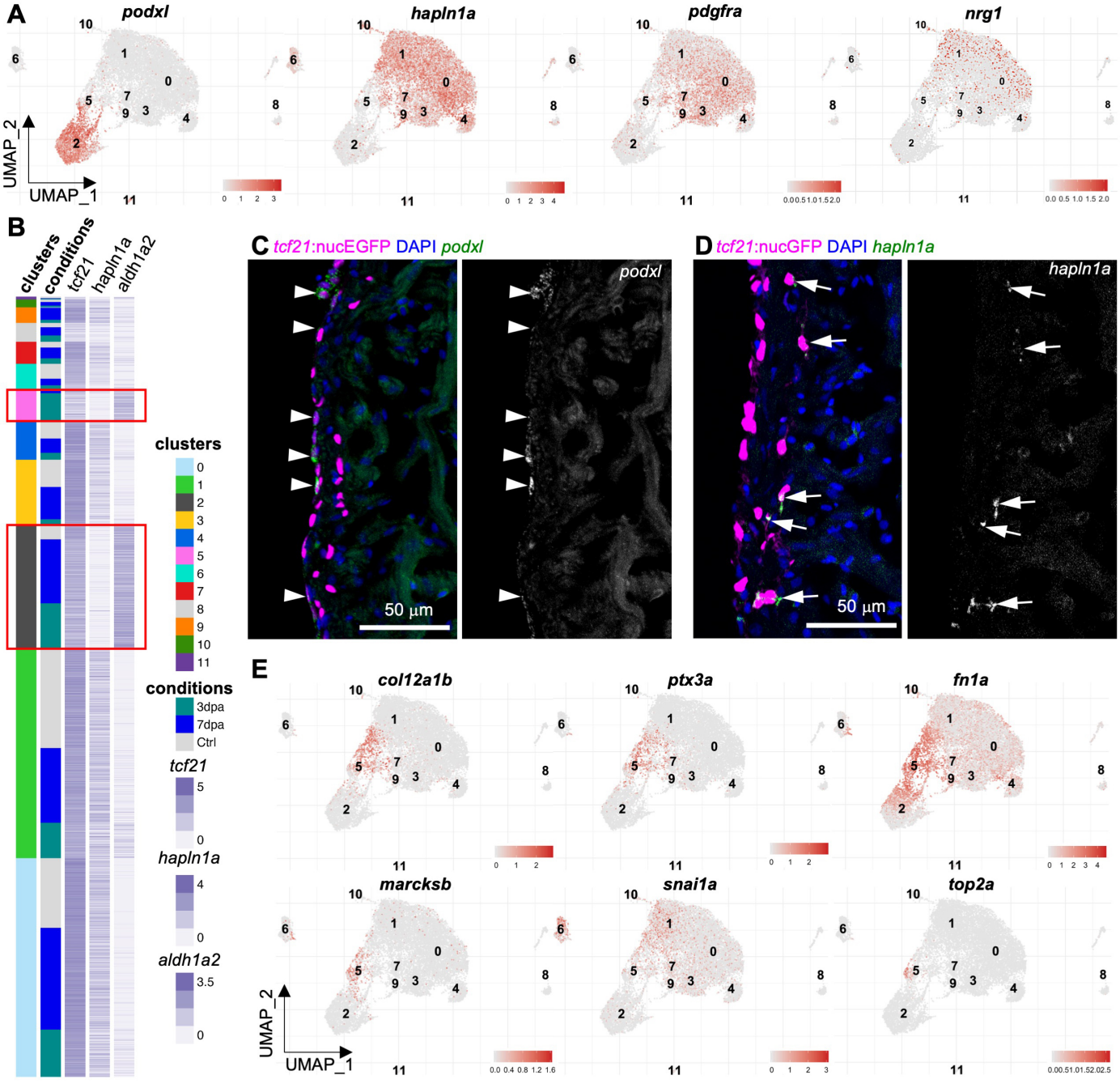

A
podxl
hapln1a
pdgfra
nrg1
UMAP_2
UMAP_1
B
clusters
conditions
tcf21
hapln1a
aldh1a2
clusters
0
1
2
3
4
5
6
7
8
9
10
11
conditions
3dpa
7dpa
Ctrl
tcf21
5
0
hapln1a
4
0
aldh1a2
3.5
0
C tcf21:nucEGFP DAPI podxl
podxl
50 µm
D tcf21:nucGFP DAPI hapln1a
hapln1a
50 µm
E
col12a1b
ptx3a
fn1a
marcksb
snai1a
top2a
UMAP_2
UMAP_1

Figure 4

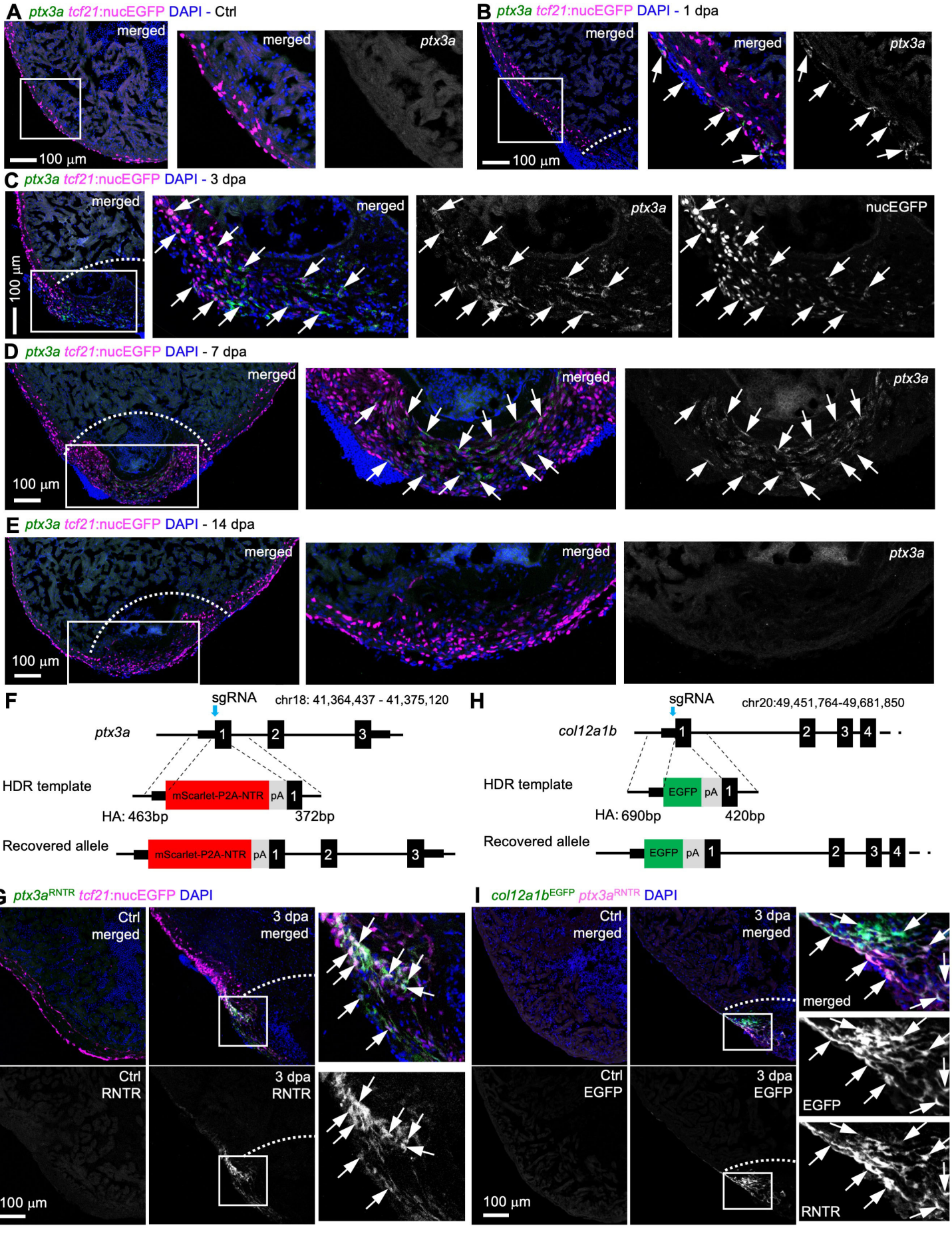

A ptx3a tcf21:nucEGFP DAPI - Ctrl
merged
merged
ptx3a
100 µm
B ptx3a tcf21:nucEGFP DAPI - 1 dpa
merged
merged
ptx3a
100 µm
C ptx3a tcf21:nucEGFP DAPI - 3 dpa
merged
merged
ptx3a
nucEGFP
100 µm
D ptx3a tcf21:nucEGFP DAPI - 7 dpa
merged
merged
ptx3a
100 µm
E ptx3a tcf21:nucEGFP DAPI - 14 dpa
merged
merged
ptx3a
100 µm
F
sgRNA
chr18: 41,364,437 - 41,375,120
ptx3a
1
2
3
HDR template
mScarlet-P2A-NTR
pA
1
HA: 463bp
372bp
Recovered allele
mScarlet-P2A-NTR
pA
1
2
3
H
sgRNA
chr20:49,451,764-49,681,850
col12a1b
1
2
3
4
HDR template
EGFP
pA
1
HA: 690bp
420bp
Recovered allele
EGFP
pA
1
2
3
4
G ptx3aRNTR tcf21:nucEGFP DAPI
Ctrl
merged
3 dpa
merged
Ctrl
RNTR
3 dpa
RNTR
100 µm
I col12a1bEGFP ptx3aRNTR DAPI
Ctrl
merged
3 dpa
merged
merged
Ctrl
EGFP
3 dpa
EGFP
EGFP
100 µm
RNTR

Figure 5

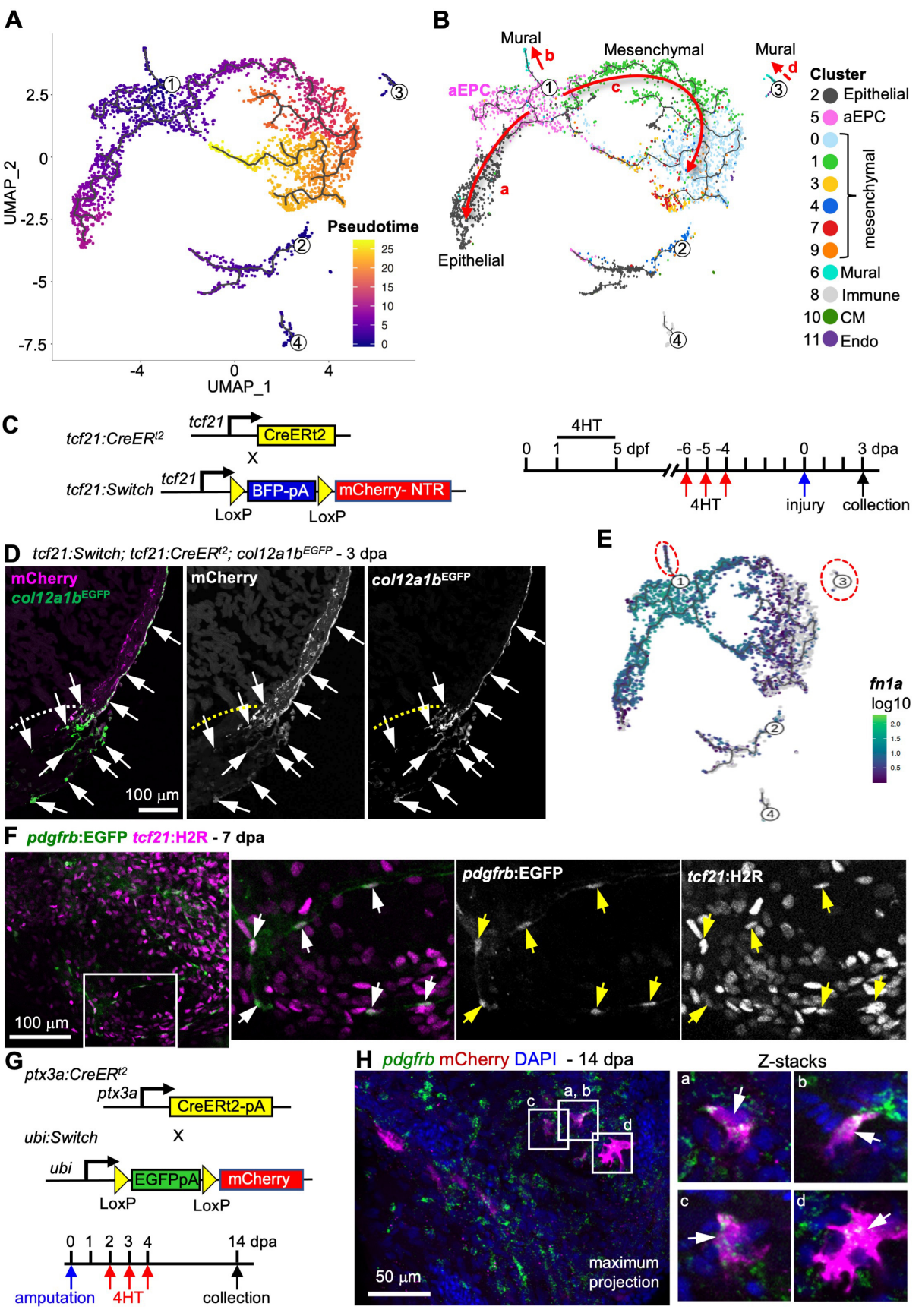

A
UMAP_2
UMAP_1
Pseudotime
25
20
15
10
5
0
B
Mural
Mesenchymal
aEPC
Epithelial
Cluster
2 Epithelial
5 aEPC
0
1
3
4
7
9
mesenchymal
6 Mural
8 Immune
10 CM
11 Endo
C
tcf21:CreERt2
tcf21
CreERt2
X
tcf21:Switch
BFP-pA
mCherry- NTR
LoxP
LoxP
4HT
0 1 5 dpf -6 -5 -4 0 3 dpa
injury
collection
D
tcf21:Switch; tcf21:CreERt2; col12a1bEGFP - 3 dpa
mCherry
col12a1bEGFP
100 µm
E
fn1a
log10
F
pdgfrb:EGFP tcf21:H2R - 7 dpa
100 µm
pdgfrb:EGFP
tcf21:H2R
G
ptx3a:CreERt2
ptx3a
CreERt2-pA
ubi:Switch
X
ubi
EGFPpA
mCherry
LoxP
LoxP
0 1 2 3 4 14 dpa
amputation
4HT
collection
H
pdgfrb mCherry DAPI - 14 dpa
Z-stacks
50 µm
maximum projection

Figure 6

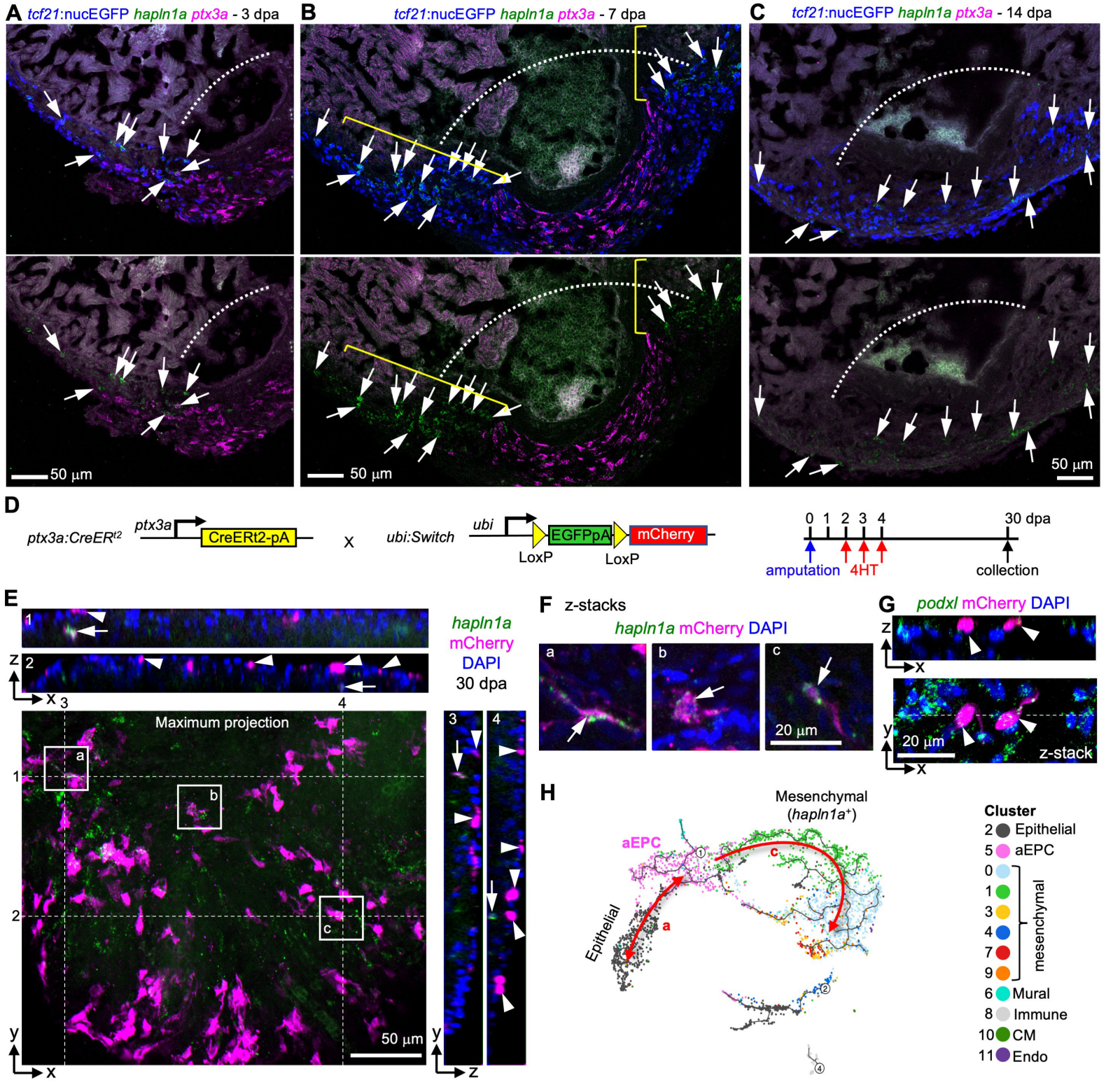

A tcf21:nucEGFP hapln1a ptx3a - 3 dpa
B tcf21:nucEGFP hapln1a ptx3a - 7 dpa
C tcf21:nucEGFP hapln1a ptx3a - 14 dpa
50 μm
50 μm
50 μm
D
ptx3a:CreER^t2
ptx3a
CreERt2-pA
X
ubi:Switch
ubi
EGFPpA
mCherry
LoxP
LoxP
0 1 2 3 4
30 dpa
amputation
4HT
collection
E
hapln1a
mCherry
DAPI
30 dpa
Maximum projection
a
b
c
50 μm
F z-stacks
hapln1a mCherry DAPI
20 μm
G podxl mCherry DAPI
20 μm
z-stack
H
Mesenchymal
(hapln1a+)
aEPC
Epithelial
Cluster
2 Epithelial
5 aEPC
0
1
3
4
7
9
mesenchymal
6 Mural
8 Immune
10 CM
11 Endo

Figure 7

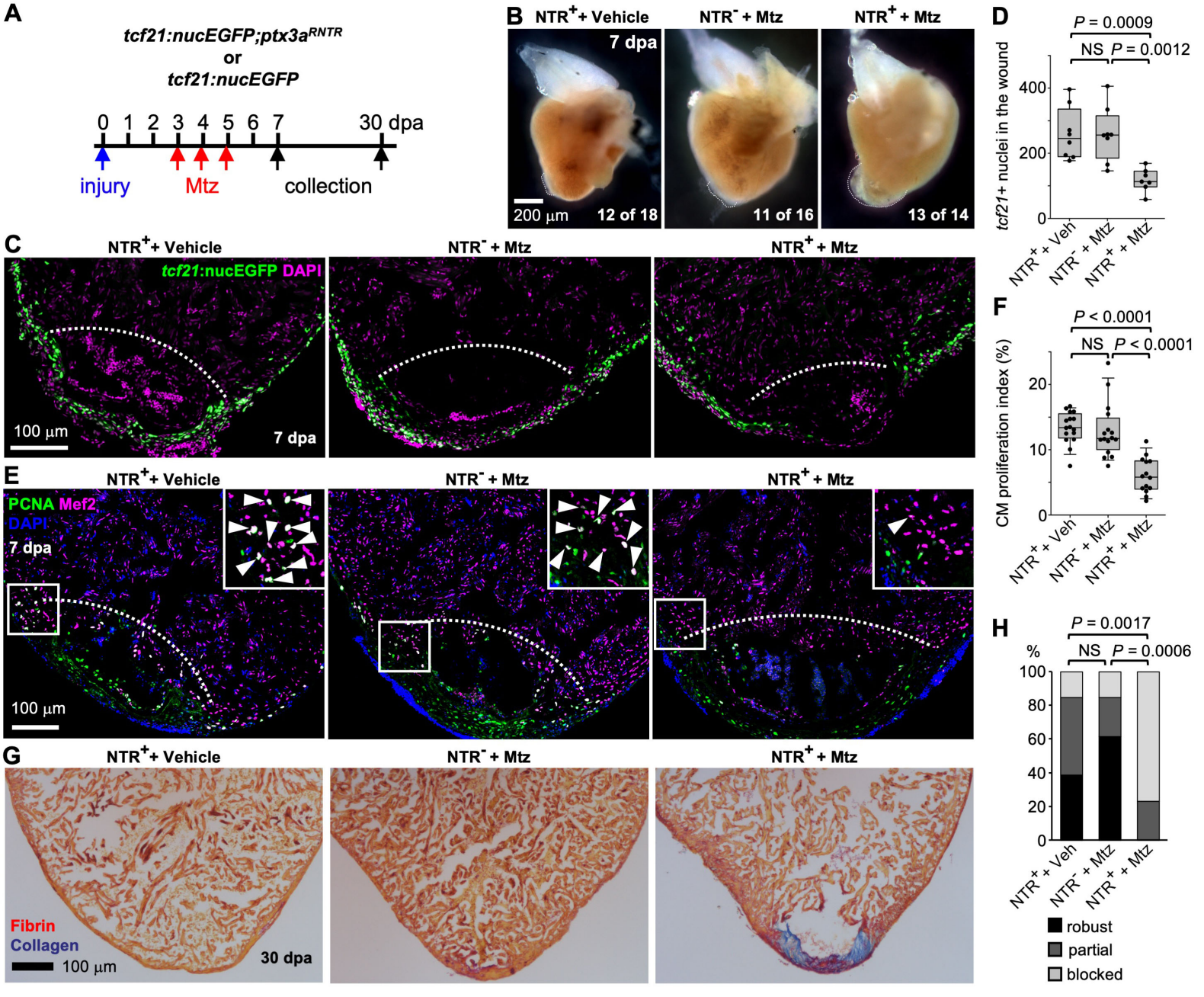

A
tcf21:nucEGFP;ptx3a^RNTR
or
tcf21:nucEGFP
0 1 2 3 4 5 6 7 30 dpa
injury
Mtz
collection
B
NTR+ + Vehicle
NTR- + Mtz
NTR+ + Mtz
7 dpa
200 μm
12 of 18
11 of 16
13 of 14
C
NTR+ + Vehicle
NTR- + Mtz
NTR+ + Mtz
tcf21:nucEGFP DAPI
100 μm
7 dpa
D
P = 0.0009
NS P = 0.0012
tcf21+ nuclei in the wound
400
200
0
NTR+ + Veh
NTR- + Mtz
NTR+ + Mtz
E
NTR+ + Vehicle
NTR- + Mtz
NTR+ + Mtz
PCNA Mef2
DAPI
7 dpa
100 μm
F
P < 0.0001
NS P < 0.0001
CM proliferation index (%)
20
10
0
NTR+ + Veh
NTR- + Mtz
NTR+ + Mtz
G
NTR+ + Vehicle
NTR- + Mtz
NTR+ + Mtz
Fibrin
Collagen
100 μm
30 dpa
H
P = 0.0017
NS P = 0.0006
%
100
80
60
40
20
0
NTR+ + Veh
NTR- + Mtz
NTR+ + Mtz
robust
partial
blocked

Figure 8

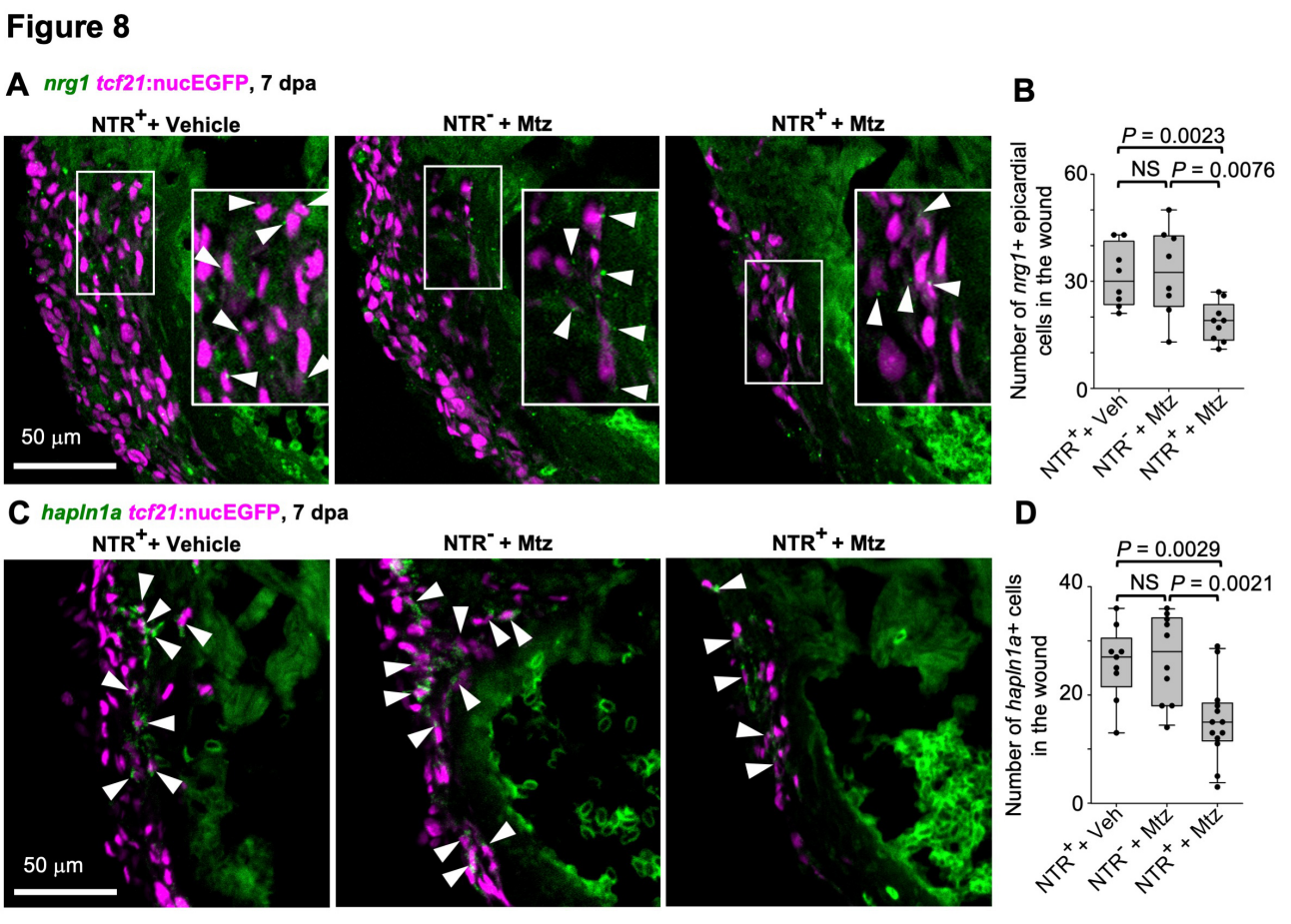
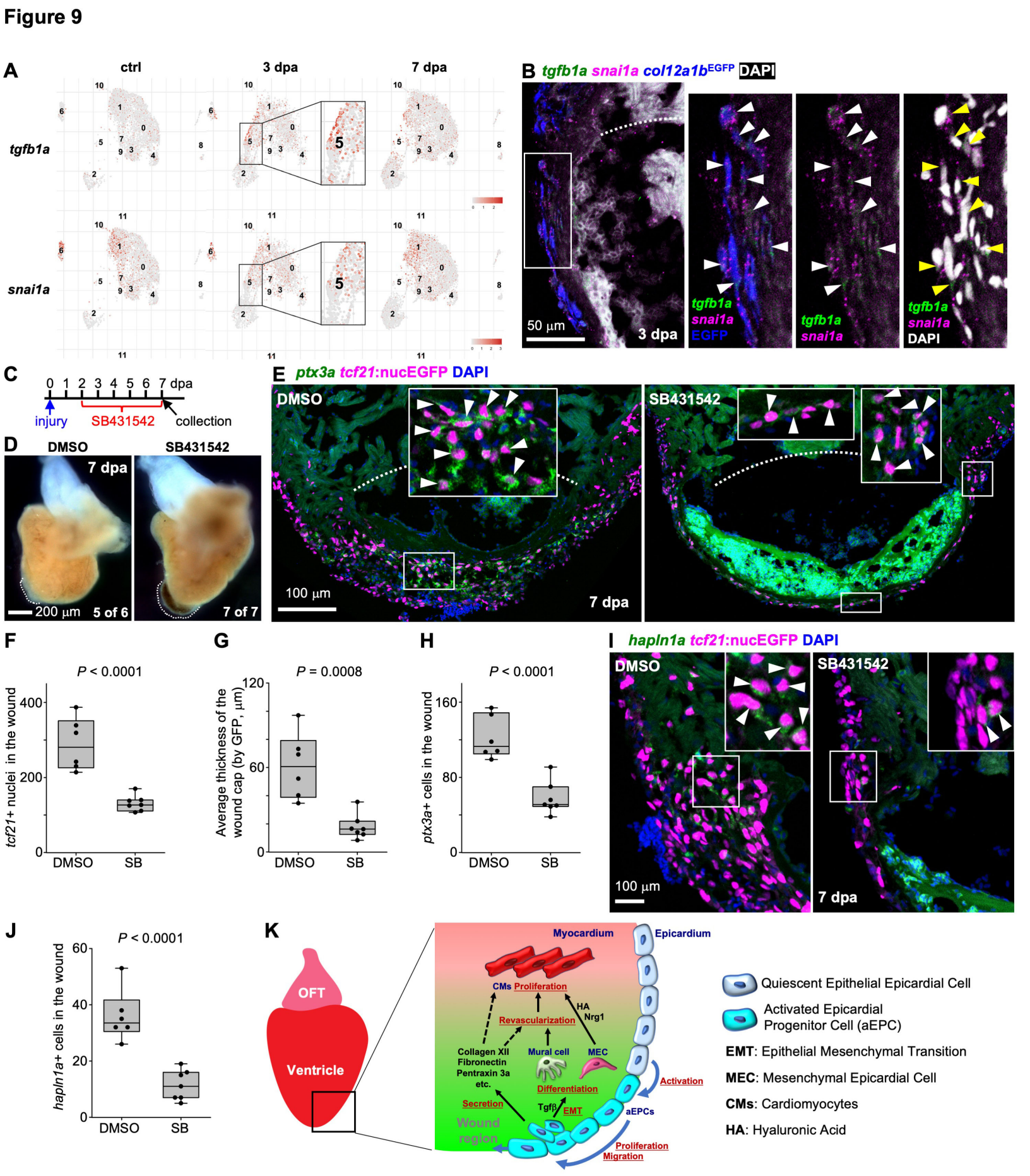

A nrg1 tcf21:nucEGFP, 7 dpa
NTR+ + Vehicle
NTR- + Mtz
NTR+ + Mtz
50 μm
B
P = 0.0023
NS
P = 0.0076
Number of nrg1+ epicardial cells in the wound
60
30
0
NTR+ + Veh
NTR- + Mtz
NTR+ + Mtz
C hapln1a tcf21:nucEGFP, 7 dpa
NTR+ + Vehicle
NTR- + Mtz
NTR+ + Mtz
50 μm
D
P = 0.0029
NS
P = 0.0021
Number of hapln1a+ cells in the wound
40
20
0
NTR+ + Veh
NTR- + Mtz
NTR+ + Mtz

Figure 9


A
ctrl
3 dpa
7 dpa
tgfb1a
snai1a
B tgfb1a snai1a col12a1b^EGFP DAPI
50 μm
3 dpa
tgfb1a
snai1a
EGFP
tgfb1a
snai1a
tgfb1a
snai1a
DAPI
C
0 1 2 3 4 5 6 7 dpa
injury
SB431542
collection
D
DMSO
SB431542
7 dpa
200 μm
5 of 6
7 of 7
E ptx3a tcf21:nucEGFP DAPI
DMSO
SB431542
100 μm
7 dpa
F
P < 0.0001
tcf21+ nuclei in the wound
DMSO
SB
G
P = 0.0008
Average thickness of the wound cap (by GFP, μm)
DMSO
SB
H
P < 0.0001
ptx3a+ cells in the wound
DMSO
SB
I hapln1a tcf21:nucEGFP DAPI
DMSO
SB431542
100 μm
7 dpa
J
P < 0.0001
hapln1a+ cells in the wound
DMSO
SB
K
OFT
Ventricle
Myocardium
Epicardium
CMs Proliferation
HA
Nrg1
Revascularization
Collagen XII
Fibronectin
Pentraxin 3a
etc.
Mural cell
MEC
Differentiation
Activation
Secretion
Tgfβ
EMT
aEPCs
Wound region
Proliferation
Migration
Quiescent Epithelial Epicardial Cell
Activated Epicardial Progenitor Cell (aEPC)
EMT: Epithelial Mesenchymal Transition
MEC: Mesenchymal Epicardial Cell
CMs: Cardiomyocytes
HA: Hyaluronic Acid