

# Functional coordination of non-myocytes plays a key role in adult zebrafish heart regeneration

Hong Ma<sup>1,2,9,†</sup>, Ziqing Liu<sup>1,2,†</sup>, Yuchen Yang<sup>1,2,3,†</sup>, Dong Feng<sup>1,2</sup>, Yanhan Dong<sup>1,2</sup>, Tiffany A. Garbutt<sup>1,2</sup>, Zhiyuan Hu<sup>4</sup>, Li Wang<sup>1,2</sup>, Changfei Luan<sup>1,2</sup>, Cynthia D. Cooper<sup>5</sup>, Yun Li<sup>3,6,7</sup>, Joshua D. Welch<sup>8,\*\*\*</sup>, Li Qian<sup>1,2,\*\*</sup>, Jiandong Liu<sup>1,2,\*</sup>

1 McAllister Heart Institute, University of North Carolina, Chapel Hill, NC 27599, USA

2 Department of Pathology and Laboratory Medicine, University of North Carolina, Chapel Hill, NC 27599, USA

3 Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, USA

4 Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599, USA

5 School of Molecular Biosciences, Washington State University Vancouver, Vancouver, WA 98686, USA

6 Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599, USA

7 Department of Computer Science, University of North Carolina, Chapel Hill, NC 27599, USA

8 Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

9 Present address: Department of Cardiology, 2<sup>nd</sup> Affiliated Hospital, School of Medicine, Zhejiang University. Hangzhou 310009, China

\* Corresponding author. Tel: +1-919-962-0326; E-mail: [Jiandong\\_liu@med.unc.edu](mailto:Jiandong_liu@med.unc.edu)

\*\* Corresponding author. Tel: +1-919-962-0340; E-mail: [li\\_qian@med.unc.edu](mailto:li_qian@med.unc.edu)

\*\*\* Corresponding author. Tel: +1-734-615-0618; E-mail: [welchjd@umich.edu](mailto:welchjd@umich.edu)

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.15252/EMBR.202152901](https://doi.org/10.15252/EMBR.202152901)

This article is protected by copyright. All rights reserved

† These authors contribute equally to the work.

## Abstract

Cardiac regeneration occurs primarily through proliferation of existing cardiomyocytes, but also involves complex interactions between distinct cardiac cell types including non-cardiomyocytes (nonCMs). However, the subpopulations, distinguishing molecular features, cellular functions, and intercellular interactions of nonCMs in heart regeneration remain largely unexplored. Using the LIGER algorithm, we assemble an atlas of cell states from 61,977 individual nonCM scRNA-seq profiles isolated at multiple time-points during regeneration. This analysis reveals extensive nonCM cell diversity, including multiple macrophage (MC), fibroblast (FB) and endothelial cell (EC) subpopulations with unique spatiotemporal distributions and suggests an important role for MC in inducing the activated FB and EC subpopulations. Indeed, pharmacological perturbation of MC function compromises the induction of the unique FB and EC subpopulations. Furthermore, we developed computational algorithm Topologizer to map the topological relationships and dynamic transitions between functional states. We uncover dynamic transitions between MC functional states and identify factors involved in mRNA processing and transcriptional regulation associated with the transition. Together, our single-cell transcriptomic analysis of nonCMs during cardiac regeneration provides a blueprint for interrogating the molecular and cellular basis of this process.

**Running title:** nonCM coordination in heart regeneration

**Keywords:** Zebrafish, heart regeneration, non-myocytes, scRNA-Seq, Topologizer

## Introduction

Adult mammalian hearts exhibit limited regenerative capacity, and are therefore susceptible to massive and irreversible cardiomyocyte (CM) loss due to myocardial infarction (Laflamme & Murry, 2011). In contrast, adult zebrafish and neonatal mice can efficiently regenerate their injured hearts through activation of CM proliferation (Jopling *et al*, 2010; Kikuchi *et al*, 2010; Porrello *et al*, 2011; Poss *et al*, 2002; Tzahor & Poss, 2017). Thus, much of the basic research on

cardiac regeneration has focused on CMs, aiming to unravel cardiac renewal mechanisms for future development of therapeutic interventions to stimulate CM proliferation and regeneration in human patients (Bassat *et al*, 2017; D'Uva *et al*, 2015; Leach *et al*, 2017; Mahmoud *et al*, 2013; Morikawa *et al*, 2017; Nakada *et al*, 2017; Price *et al*, 2019; Tao *et al*, 2016; Wu *et al*, 2016). Yet the heart as a whole contains many other cell types including endothelial cells, fibroblasts, and a wide variety of immune cells. In particular, it is increasingly recognized that non-myocytes (nonCMs) play active roles in regulating CM behaviors (Kikuchi *et al*, 2011b; Klotz *et al*, 2015; Lai *et al*, 2017; Riley, 2012). Despite substantial advances in understanding genetic regulation of zebrafish heart regeneration (Gonzalez-Rosa *et al*, 2017), the cardiac nonCM composition and its dynamic changes in response to injury remains largely unexplored. A better understanding of how diverse cells compose zebrafish heart to maintain its homeostasis will shed lights on the mechanisms underlying its robust regenerative capacity and is required for the development of therapeutic strategies.

In this study, using the newly developed LIGER algorithm (Welch *et al*, 2019) that allows flexible modeling across highly diverse single-cell datasets, we analyzed the transcriptome dynamics of 61,977 individual nonCMs isolated at multiple time points during zebrafish heart regeneration. Through this analysis, we identified major nonCM cell types, including multiple macrophage, fibroblast and endothelial cell subpopulations with unique tempo-spatial distributions and highly cooperative interactions during the process of cardiac regeneration. Interestingly, perturbation of macrophage functional dynamics resulted in compromised interactions among nonCMs concomitant with reduced cardiomyocyte proliferation and defective cardiac regeneration. Furthermore, we developed a computational algorithm Topologizer and revealed the topological relationship of the cellular manifolds. Combining Topologizer and RNA velocity analyses, we uncovered dynamic transition between macrophage functional states and identified factors involved in mRNA processing and transcriptional regulation associated with the transition. Together, our single-cell transcriptomic analysis of nonCMs during cardiac regeneration provides a blueprint for interrogating the molecular and cellular basis of cardiac regeneration.

## **Results**

### **Single-cell transcriptome atlas of cardiac nonCMs in adult zebrafish heart**

The lack of detailed information on the cellular identities and cell states of the nonCMs associated with tissue regeneration is a major hurdle to precisely delineating the biological events underlying the regeneration process. To address this challenge, we first sought to generate a single-cell map of nonCMs in wildtype adult zebrafish ventricle. Following cell dissociation and low-speed centrifugation to remove CMs (Materials and Methods, and Fig 1A), we enriched nonCMs and performed single-cell RNA sequencing (scRNA-Seq) using the 10x Genomics Chromium platform (Fig 1A). In total, we obtained 7041 high-quality nonCMs that passed quality control and filtering criteria (Appendix Table S1). We then performed unsupervised dimensionality reduction and clustering, and identified 12 distinct cell clusters (Appendix Fig S1A). Cells from two independent experiments intermingled well, suggesting minimal batch effects (Appendix Fig S1B). Based on known marker gene expression, we found eight major nonCM cell types, including endothelial cells (ECs; *cdh5*, *kdrl*, *fli1a*, *flt1*) (Habeck *et al*, 2002; Larson *et al*, 2004; Lawson & Weinstein, 2002), epicardial cells/fibroblasts (FBs; *tcf21*, *fn1b*, *coll1a1*) (Ivey & Tallquist, 2016; Kikuchi *et al*, 2011a; Moore-Morris *et al*, 2014a; Snider *et al*, 2009; Wang *et al*, 2013), resident mesenchymal cells (Mes; *angptl7*, *rspo1*, *mgp*) (Costa *et al*, 2017; Gore *et al*, 2011), macrophages (MCs; *mpeg1.1*, *mfap4*, *c1qb*, *cd74a*) (Ellett *et al*, 2011; Spilsbury *et al*, 1995; Walton *et al*, 2015), neutrophils (Neutro; *lyz*, *mpx*) (Harvie & Huttenlocher, 2015; Walters *et al*, 2010), T/NK/B cells (T/NK/B; *sla2*, *irf4b*, *ccl36.1*, *cxcr4a*, *lck*, *nkl.2*, *zbtb32*, *cd79a*) (Athanasiadis *et al*, 2017; Carmona *et al*, 2017), erythrocytes (Eryth; *cahz*, *slc4a1a*) (Moore *et al*, 2018; Paw *et al*, 2003) and thrombocytes (Throm; *itga2b*, *gp1bb*) (Kato *et al*, 2004; Lin *et al*, 2005) (Fig 1B and C; Appendix Fig S1C).

Interestingly, the EC cells are the nonCM cell type that is grouped into distinct clusters. Because zebrafish hearts contain three types of highly specialized ECs—endocardial ECs (eECs), lymphatic ECs (IECs) and coronary ECs (cECs)—we performed a second round of analysis on nonCMs expressing the canonical EC marker genes *cdh5* and *kdrl* and identified three EC populations and mural cells based on the expression of marker genes—*gata5* for eECs (Nemer & Nemer, 2002), *lyve1a* and *prox1a* for IEC (Okuda *et al*, 2012; van Impel *et al*, 2014), *aplrra* for cEC (Cui *et al*, 2019) and *cd248a*, *acta2*, and *tagln* for mural cells (Bagley *et al*, 2008; Santoro *et al*, 2009) (Fig 1D and 1E; Appendix Fig S1D to F). The molecular signatures defining these three types of zebrafish ECs have not been fully explored. With the high resolution of our scRNA-seq data, we found that the transcriptome of cECs is more similar to that of their

associated mural cells—including both cells expressing pericyte markers and cells expressing smooth muscle cell markers—than to those of the eECs and IECs (Fig 1D; Correlation analysis in Appendix Fig S1H). Further differential gene expression analysis identified highly expressed and specific markers for each EC type (Fig 1F; Appendix Fig S1G). In zebrafish, vascularization of the ventricle is driven by angiogenesis of eECs (Harrison *et al*, 2015). However, due to limited numbers of cECs in zebrafish hearts and the lack of genetic tools to isolate and enrich this population, whether and how cECs differ from eECs at the molecular level is unclear.

Our single-cell study revealed that adult zebrafish heart had a similar cellular composition to that of adult mouse heart (Pinto *et al*, 2016) (Fig 1G). To identify new markers for each nonCM cell type, we performed differential gene expression analysis for each cell type and identified panels of highly expressed genes specific for each nonCM population (Fig 1H). Gene Ontology (GO) analysis demonstrated that each cell population was associated with distinct biological functions and supported the assignment of cell identities based on canonical markers (Fig 1C; Appendix Fig S2). Therefore, our results provide a new benchmarking dataset for defining zebrafish cardiac nonCM identities. The newly characterized markers promise to increase the feasibility and resolution of functional studies on zebrafish nonCM populations.

### **Mapping of coordinated responses of nonCMs during heart regeneration**

We next sought to resolve the composition and dynamics of nonCMs during cardiac regeneration. To this end, we performed scRNA-seq at multiple time points (2 days post injury (dpi), 7 dpi and 14 dpi) that correspond to major pathophysiological events post cardiac injury (Cao *et al*, 2016; Lai *et al.*, 2017; Poss *et al.*, 2002). We obtained transcriptomes of 20,124 nonCMs that passed quality control and filtering criteria from the three time points post injury (Appendix Table S1). These cells were then jointly analyzed with the nonCMs obtained from uninjured ventricles. Integrating scRNA-seq datasets containing a variety of cell types from multiple biological time points proved challenging: cells separate by a combination of dataset of origin and cell type, suggesting the existence of technical and biological differences (Appendix Fig S2A). We thus applied the recently published algorithm LIGER (Welch *et al.*, 2019) that delineates each cell by shared and dataset-specific features (metagenes) and allows for jointly defining shared cell populations even across multiple heterogeneous datasets. A further advantage of LIGER is interpretability—the ability to associate each factor (metagene set) with specific populations of

cells, which is unique among currently available integration analysis methods. The interpretability of the LIGER factorization allowed us to exclude technical (e.g. ribosomal, mitochondrial, and stress genes) and biological (e.g. cell cycle states) confounding factors during joint analysis of all cell types (Appendix Fig S2B to E). Using the aforementioned markers (Fig 1C and H), we assigned cell type identity to LIGER clusters and found that various nonCM types identified in the uninjured hearts were present post injury, albeit with varying frequencies (Fig 2A and B; Appendix Fig S2F to H). Overall, the higher alignment uniformity of mixing samples not only accurately preserved the cell-type architectures, but also enabled us to assemble an integrated atlas of cell states using datasets from multiple replicates and time points.

Changes in nonCM composition occurred most dramatically at 2 dpi, remained pronounced at 7dpi but became minimal by 14 dpi. Among all nonCM cell types, MCs showed the most significant frequency change (Fig 2B), suggesting an acute expansion of the MC population followed by gradual resolution of immune response as the heart regenerates. To comprehensively chart the behavior of MCs over time, we jointly analyzed MCs from all time points and discovered significant diversity within this population, including 5 distinct subpopulations (Fig 2C; Appendix Fig S3A and B). All MC subpopulations shared common MC marker genes such as *mfap4* (Walton *et al.*, 2015) and *mpeg1* (Ellett *et al.*, 2011), yet each subpopulation expressed distinct marker genes (Fig 2D; Appendix Fig S3C to E). Interestingly, while MC1 and MC2 appeared across all examined stages, MC3-5 became more apparent post cardiac injury at 2 and 7 dpi (Fig 2E; Appendix Fig S3F). Among all MCs, MC1 cells exhibited the highest level of *illb* expression and specifically expressed pro-inflammatory factors *tnfa* and *csf3b* with enriched GO terms related to inflammatory response and neutrophil chemotaxis (Fig 2D. Appendix Fig S3D and E). Largely due to an expansion of the number of MC1 cells, the overall expression of *tnfa*, *csf3b* and other pro-inflammatory cytokines *ccl35.1*, *ccl34.4*, and *cxcl11.1* was transiently upregulated at 2 and 7 dpi. (Fig 2E and F; Appendix Fig S3G). MC1 therefore represents the major MC population that mediates the critical acute phase of pro-inflammatory activation post cardiac injury. In contrast, MC2 cells highly expressed genes involved in antigen presentation such as *cd74a*, *cd74b* (Schroder, 2016) and *mhc2dab* (Wittamer *et al.*, 2011) and related to GO terms like antigen processing and presentation, suggestive of a role in immune surveillance (Epelman *et al.*, 2014) (Fig 2D; Appendix Fig S3D and E). MC2's relative frequency within the MC population decreased initially but subsequently increased over

time, and MC2 became the dominant MC subpopulation at 14 dpi (Fig 2E). Furthermore, we found that MC2 marker genes *cd74a* and *mhc2dab*, after being initially downregulated at 2 dpi, were continuously upregulated until reaching a peak level at 14 dpi (Appendix Fig S3H), suggesting a gradually enhanced activation of MC2 cells. MC3 cells highly expressed *cd9b*, which encodes a tetraspanin family protein that interacts with Fc $\gamma$  to activate phagocytosis (Huang *et al*, 2011; Kaji *et al*, 2001) (Fig 2E). Consistently, MC3 cells also highly expressed other genes involved in phagocytosis and proteolysis including *cd63* and cathepsins (*ctsc*, *ctsd* and *ctsla*) (Aderem, 2003; Pols & Klumperman, 2009) (Fig 2D; Appendix Fig S3I). Though barely present in the uninjured heart, MC3 emerged as a major cluster at 2 dpi and then gradually decreased its frequency (Fig 2E). The MC4 subpopulation showed cell cycle activity (Fig 2D; Appendix Fig S3D, E and J), likely representing a proliferating pool of cardiac MCs to replenish the MC pools post cardiac injury (Davies *et al*, 2013). The remaining minor cluster MC5 (0.5~3.9%) highly expressed granulin genes *grn1* and *grn2* and mostly existed at 2 dpi and 7 dpi (Fig 2D and E; Appendix Fig S3D, E and K), likely representing MCs actively engulfing and degrading cell debris (Altmann *et al*, 2016; Tsuruma *et al*, 2018; Yoo *et al*, 2019). In support of MC subpopulation clustering, fluorescence in situ hybridization (FISH) indicated that the marker genes for MC subpopulations were expressed in subsets of the *mpeg1.1* expressing MCs. Consistently, we also found that *tnfa* and *ctsc* marked non-overlapping MC populations (Fig 2G). In summary, MCs exist in multiple definable states that exhibit dynamic functional changes from homeostatic conditions to acute immune response till inflammation resolution (Fig 2E). This dynamic change in the number and composition of MCs may reflect differential requirements for temporally regulated functions of MC subpopulations in cardiac repair and regeneration.

Next we determined whether and how other nonCM populations change their cellular composition in response to the temporal dynamics of MC activation and function. We first characterized FB, a cell type traditionally regarded as responsible for extracellular matrix (ECM) production. Unbiased clustering identified 4 FB subpopulations expressing the canonical fibroblast marker genes *tcf21*, *fn1b* and *coll1a1* (Ivey & Tallquist, 2016) (Fig 3A; Appendix Fig S4A to C). However, these four FB subpopulations clearly exhibited distinct ECM gene expression profiles (Fig 3B). Upon cardiac injury, fibroblasts became activated as evidenced by their transient upregulation of ECM genes expression (Appendix Fig S4D). FB1 and FB2 upregulated essentially the same ECM genes (i.e. *fn1b*, *dcn* and *sparc*), yet FB2 consistently

demonstrated higher expression level than FB1 (Appendix Fig S4E). FB3 was a unique cluster that drastically and acutely expanded in response to cardiac injury (Fig 3C; Appendix Fig S4F). *In silico* cell cycle assignment and expression of proliferation markers also suggest that FB3 is a highly proliferating FB subpopulation (Appendix Fig S4G and H). Compared to the other FB subtypes, FB3 cells transiently upregulated a unique set of ECM genes *coll2a1a*, *coll2a1b*, *postnb* and *fn1a* (aka, *fn1*) as well as *gstm.3*, which encodes the mu class glutathione S-transferase that functions to detoxify, among others, the products of oxidative stress (Glisic *et al*, 2015) (Fig 3D; Appendix Fig S4I and J). FB3 also transiently expressed the smooth muscle marker gene *tagln* (aka, *sm22*), suggesting a transformed phenotype (Fig 3E). FB3 likely corresponds to the *postnb*-positive “activated fibroblasts” that when ablated, led to reduced CM proliferation after cardiac injury (Sanchez-Iranzo *et al*, 2018; Wang *et al.*, 2013). Interestingly, double FISH for *pcf21* and FB3 specific marker *postnb* demonstrated that FB3 localization was restricted to the site of injury at 7 dpi (yellow, Fig 3F), while the *postnb* negative FB1 and FB2 cells were located along the periphery of the ventricle (green, Fig 3F; Appendix Fig S4K). Importantly, the appearance of FB3 coincided temporally and spatially with that of *tnfa*:GFP positive MC1 cells (Fig 2E and G; Fig 3F; Appendix Fig S4L).. Additionally, we found that, compared to their expression in other MC and FB subtypes, the ligands (*tnfa* and *tgfb1a*) and their respective receptors (*tnfrsf1a* and *tgfbr2a*) were more highly expressed in MC1 and FB3 subtypes compared to the other MC and FB subtypes, respectively (Fig 2D; Fig 3G; Appendix Fig S3L). Consistently, double FISH analysis showed a colocalization of *tnfrsf1a* and *postnb* in FB3, and a non-overlapping expression pattern of *tnfa* and *tnfrsf1a* at the site of injury (Fig 3H). Together, these data suggest potential intercellular communications between MCs and FBs that could play a role in stimulating phenotypic conversion of fibroblasts into a transformed state (FB3). Indeed, FB3 showed 897 upregulated genes post injury – significantly more than FB1 (411) and FB2 (238) – further supporting its transformed phenotype (Appendix Fig S4M). FB3 also highly expressed a unique set of transcription factors such as *fosl2*, *sap18*, *phb2b* and *prmt1*, and mRNA splicing factors (Appendix Fig S4N), providing candidate regulators for future research to understand the regulatory mechanisms underlying fibroblast transformed phenotypes. FB4 is a minor FB subpopulation highly expressed genes involved in Wnt signaling, likely corresponding to a recently reported murine FB subpopulation (Farbehi *et al*, 2019). Together, our analyses identified four fibroblast subpopulations with unique ECM gene signatures from



distinct anatomical sites that confer differential roles in maintaining cardiac structural integrity and modulating cell behaviors. We also identified a unique FB3 subtype that exhibited a transformed phenotype likely induced by its intercellular interactions with MCs.

eEC constitutes the largest nonCM population in adult zebrafish heart (Fig 1B and F) with 4 major subpopulations (Fig 3I; Appendix Fig S5A and B). eEC1 was the largest cluster, containing over 70% of eECs in both uninjured and injured hearts and likely representing the eECs that localized at a distance from the injury site. Following cardiac injury, eEC2 percentage in eEC transiently increased (Fig 3I), closely resembling MC1 and FB3. eEC2 highly expressed inflammatory chemokine genes such as *cxcl18b* (Fig 3J) (Torraca *et al*, 2017) and E selectin (*sele*, Fig 3J) (Silva *et al*, 2017), and transiently upregulated *nppc* (C-type natriuretic peptide, top marker of eEC2, Appendix Fig S5C and D), a recently identified regulator of angiogenesis and vascular remodeling in response to ischemic injury (Bubb *et al*, 2019; Munch *et al*, 2017).

More interestingly, eEC2 cells highly expressed transcription factors *atf3* and *fosl1a* (Fig 3J), which are stress response genes recently revealed to be responsible for eEC proliferation and wound closure post mouse aorta denudation injury (McDonald *et al*, 2018). Double FISH confirmed a colocalization of *fosl1a* and pan EC marker *cdh5* (Fig 3K). Interestingly, the *fosl1a* positive eEC cells were localized to the site of injury, similar to that of the MC1 and FB3 subpopulations at 7 dpi. We also found expression of *cxcl18b*, *sele*, *atf3* and *fosl1a* to be much higher at 2 dpi and 7 dpi compared to uninjured and 14 dpi eEC2 cells (Appendix Fig S5E), suggesting transient activation of this cell type (Munch *et al*, 2017; Sanchez-Iranzo *et al*, 2018). eEC3 did not show significant frequency change after injury (Fig 3I) and it highly expressed genes implicated in cardiac valve development and function such as *frzb*, *bmp6*, *wnt11r* and *tgfb1* (Kim *et al*, 2001; Norris *et al*, 2005; Person *et al*, 2005; Sinha *et al*, 2015; Su *et al*, 2019; Touma *et al*, 2017) (Appendix Fig S5C and F), suggesting its identity as valvular endothelial cells. eEC4 was a small cluster of proliferating cells (Fig 2I; Appendix Fig S5C, G and H). Altogether, our analysis indicates that cardiac eECs exhibit heterogeneity with a unique activated eEC2 subtype and involved in endocardial angiogenesis.

## **Perturbed nonCM functional dynamics and cooperative interaction in non-regenerating heart**

To study the biological significance of the highly cooperative interactions of major nonCM populations during the process of cardiac regeneration, we utilized a zebrafish *kit* (aka, *c-kit*) mutant to determine the role of the proposed functional states of the distinct nonCM subpopulations in heart regeneration. The zebrafish genome contains two *kit* paralogs: *kita* and *kitb*. The *kita* mutant allele *kita*<sup>w34b</sup> was previously uncovered and is predicted to be functionally null (Cooper *et al*, 2009). The *kitb* mutant allele, *kitb*<sup>sa15348</sup>, harbors a T to A substitution in exon 3 that results in an early truncation of the encoded protein upstream of the kinase domain (Kettleborough *et al*, 2013). We generated *kita*<sup>w34b</sup>; *kitb*<sup>sa15348</sup> double mutants (hereafter referred to *kit* mutants). The homozygous *kit* mutants survive to reproductive adulthood with no overt morphological and growth defects except for a reduction in overall pigmentation (Appendix Fig S6A). Complete loss of *kit* function in zebrafish also doesn't cause primordial germ cell development defects and macrocytic anemia (Parichy *et al*, 1999). Nevertheless, we did not observe any apparent differences in cardiac size and morphology between control and *kit* mutants (Fig 4A). We performed cardiac resection on control and *kit* mutant hearts of 4- to 6-month-old animals. At 14 dpi, the injury areas of both control and *kit* mutant hearts showed pronounced accumulation of fibrin and collagen deposits. At 30 dpi, while the control hearts were mostly devoid of fibrin and collagen deposits and had fully regenerated, the mutant hearts displayed substantial fibrin and collagen deposits and impaired myocardium regeneration (Fig 4A and B). Consistent with the impairment of myocardium regeneration, the CM proliferation index of the mutant hearts was dramatically lower than that of control hearts (Fig 4C and D). To further assess the proliferation defect of mutant CMs, we conducted transcriptome-wide bulk RNA-seq of control and mutant CMs (Appendix Fig S6B). Principal component analysis (PCA) demonstrated minimal batch effect between biological replicates (Appendix Fig S6C). Gene Set Enrichment Analysis (GSEA) indicated that proliferation-related gene sets including Myc targets and E2F targets were substantially downregulated in mutant CMs compared to control CMs post injury (Fig 4E; Appendix Fig S6D). Furthermore, mutant CMs significantly downregulated the expression of positive cell cycle regulators and upregulated the expression of negative cell cycle regulators (see Materials and Methods) compared to control CMs (Fig 4F). Together, these data demonstrate that loss of *kit* function compromised injury-induced CM proliferation and myocardium regeneration.

Since *kit*-expressing cells rarely give rise to CMs (Sultana *et al*, 2015; van Berlo *et al*, 2014), the impairment of mutant myocardium regeneration may arise from defects in nonCMs (Aurora *et al*, 2014). We thus performed scRNA-seq to examine transcriptomic alterations of nonCMs caused by loss of *kit* function. To assign cell type identity to mutant nonCM cells, we applied LIGER to jointly analyze datasets collected from *kit* mutant hearts at all four time-points (Fig 4G; Appendix Fig S6E to H). The mutant hearts contained all nonCM cell types found in wildtype hearts, albeit with different frequencies (Fig 4H). Interestingly, the mutant hearts had a significantly higher percentage of MCs before injury but exhibited less dramatic expansion of MCs at 2 dpi (1.8-fold increase) compared to wildtype hearts (4.6-fold increase). More importantly, at 7 dpi, MC frequency in mutant hearts restored to a preinjury level, while a 3-fold higher frequency was still observed in wildtype hearts (Fig 4H). We also observed a significant reduction in MC numbers in 7 dpi mutant hearts compared to the 7dpi wildtype hearts (Fig 4I).

To further investigate the effect of *kit* knockout on MCs (Bertrand *et al*, 2005; Kierdorf *et al*, 2013; Stremmel *et al*, 2018), we jointly analyzed MCs from wildtype and mutant hearts and identified 5 corresponding mutant MC subpopulations (abbreviated as KMC1-5) that expressed markers of MC1-5, respectively (Fig 5A; Appendix Fig S7A to C). To determine whether and how loss of *kit* impacted MC function in regeneration, we analyzed and compared the composition and transcriptome dynamics of mutant MCs to those of wildtype MCs (Fig 5B and C). GO analysis of differentially expressed genes (DEGs) between wildtype and mutant MC subpopulations indicated that mutant MC subpopulations exhibited increased expression of proteolysis- and translation-related genes but decreased expression of genes related to their immune functions, suggestive of impaired MC function (Fig 5B; Appendix Fig S7D). In uninjured hearts, consistent with GO analysis, KMC1 cells expressed lower levels of pro-inflammatory genes *il1b*, *ccl35.1* and *ccl34a.4* than MC1 cells (Fig 5D; Appendix Fig S7E and G, “un” for uninjured). Yet, the percentage of KMC1 doubled that of MC1 in wildtype hearts (Fig 5C, un). Consequently, mutant hearts expressed significantly higher total levels of the pro-inflammatory genes than wildtype controls (Fig 5E; Appendix Fig S7F and H, un). Interestingly, while the expression of pro-inflammatory genes in KMC1 remained slightly lower than that in MC1 at both 2 and 7 dpi, the expansion of KMC1 was less pronounced than that of MC1 (Fig 5C and D), leading to reduced total level of pro-inflammatory gene expression in the mutant hearts than wildtype controls (Fig 5E; Appendix Fig S7F and H). Surprisingly, at 14 dpi when

inflammation in wildtype hearts was already diminished, KMC1 more than doubled the frequency of MC1 (Fig 5C, orange bars), and exhibited significantly higher level of pro-inflammatory genes expression than MC1 (Fig 5D and E), suggesting incomplete resolution of inflammation.

In uninjured hearts, MC3 was a minor population, yet the percentage of its counterpart KMC3 in the mutant hearts was significantly higher. Similar to KMC1, expansion of KMC3 post injury at 2 and 7 dpi was much less pronounced than MC3 (Fig 5C). In line with GO analysis, KMC cells expressed slightly reduced level of phagocytosis-related genes such as *cd63* and *ctsd* than MC cells (Fig 5F; Appendix Fig S7D and I). As a result, the mutant hearts expressed lower overall levels of *ctsd* after injury (Fig 5G), suggesting compromised MC phagocytic function. The percentage of the proliferating KMC4 cells is much lower than MC4 at both 2 and 7 dpi, possibly contributing to impaired post-injury expansion of MC in the *kit* mutant (Fig 5C). Together, our data indicated that, though the mutant hearts were more inflammatory under uninjured conditions, they failed to mount or maintain a robust inflammatory response and exhibited compromised activation of phagocytic function following cardiac injury. Paradoxically, the mutant hearts also exhibited impaired resolution of inflammation when inflammation in wildtype hearts diminished. Failure to mount a strong inflammatory and phagocytic response and to resolve inflammation in a timely manner could all contribute to the compromised cardiac regeneration in the mutant hearts.

We further investigated whether impaired MC function in the mutant hearts led to abnormal FB or eEC activation. We characterized mutant FBs and identified 4 subpopulations that corresponded to FB1-4, respectively (abbreviated as KFB1-4, Fig 5H; Appendix Fig S8A to C). The percentage of KFB3, counterpart of the activated FB3 cells in wildtype hearts, was decreased in mutant hearts (Fig 5I; Appendix Fig S8D and E), consistent with reduced expression of proliferation markers in KFB3 (Appendix Fig S8F). Interestingly, KFB3 cells expressed higher levels of their marker genes such as *coll2a1a* and *fn1a* before cardiac injury (Fig 5J; Appendix Fig S8G, un), which was supported by significantly higher total level of *coll2a1a* and *fn1a* expression in the uninjured heart (Fig 5K; Appendix Fig S8G, un), possibly in response to the overall more inflammatory environment in the mutant heart. Nevertheless, *coll2a1a* and *fn1a* expression became substantially reduced in the mutant than wildtype hearts after injury, suggesting compromised FB3 transformation (Fig 5K). While our study indicated a

role of MCs in inducing FB3 phenotype, communication between MCs and FBs is not unidirectional. Following cardiac injury, *mif* was dramatically upregulated in FBs. *mif* encodes a cytokine to counteract the anti-inflammatory effect of glucocorticoids, thereby enhancing immune response in wildtype hearts (Fig 5L, dark green). However, *mif* expression was only marginally upregulated in KFBs (Fig 5L, orange). This deficiency in *mif* upregulation post injury was supported by quantification of whole heart expression (Fig 5M) and it likely contributed to the dampened inflammatory response of MC in the mutant hearts. Taken together, decreased numbers of transformed FB3 cells, reduced total level of ECM gene expression in FB3 cells, and dampened cytokine production from FBs could contribute to impaired regeneration observed in the mutant hearts.

LIGER analysis also identified 4 mutant eEC subpopulations that corresponded to eEC1-4, respectively (denoted as KeEC1-4, Fig 5N; Appendix Fig S9A to C). Yet, KeECs exhibited distinct transcriptome dynamics. Most notably, KeECs in the uninjured mutant hearts upregulated inflammation and stress response genes including *ifitm1* (Interferon-Induced Transmembrane Protein 1) and *junba* (Fig 5O; Appendix Fig S9D, un), which were only expressed in wildtype eEC post injury, suggesting that KeECs were inflammatory, activated and stressed in the mutant hearts. However, upregulation of these genes was significantly dampened in KeECs post injury, suggesting suppressed activation of KeECs (Fig 5O). KeEC2 did not expand like their counterparts in wildtype after cardiac injury (Fig 5N), and total expression of *ifitm1* and *junba* in mutant vs. wildtype hearts supported the deficient activation of KeECs (Fig 5P; Appendix Fig S9E). More interestingly, KeEC2 expressed much lower levels of genes implicated in chemotaxis, angiogenesis, and vascular regeneration such as *cxcl18b*, *nppc* and *atf3* after injury compared to wildtype eEC2, which was also supported by total expression of these genes in mutant vs. wildtype hearts (Appendix Fig S9F-K). Taken together, the mutant eECs seemed to be already inflammatory and stressed before the injury. But paradoxically, they failed to activate important genes for leukocyte attraction, stress response, angiogenesis, and vascular remodeling after cardiac injury, which may contribute to reduced MC infiltration and impaired regeneration of the mutant hearts. To further experimentally determine the effect of impaired MC function, we depleted MCs by clodronate liposome treatment or inhibited the pro-inflammatory response using the anti-inflammatory corticosteroid dexamethasone (Kyritsis *et al.*, 2012; Lai *et al.*, 2017). Both treatments resulted in a significant reduction in the number of *tcf21:nucGFP*-

positive FBs at the site of injury at 7 dpi, similar to what observed in the *kit* mutant. In contrast, the number of *tcf21:nucGFP*-positive FBs on the periphery of the ventricular wall was not affected (Fig 5Q and R).

### **Topologizer reveals nonCM cell state transition during heart regeneration**

We characterized the behaviors of multiple molecularly distinct subpopulations for the three major nonCM populations, namely MC, FB and eEC, during cardiac regeneration. Each cell within these subpopulations represents a transcriptomic snapshot of the dynamic events during the process of cardiac regeneration. Thus, we reasoned that these asynchronous gene expression dynamics provide an opportunity to investigate the cell-state transitions between different functional subtypes as cardiac regeneration proceeds. To accomplish this, we developed a novel approach, Topologizer, which leverages the mathematics of algebraic topology to characterize the “shape” of the cellular manifold during heart regeneration. Building on the pioneering work of Rizvi et al. (Rizvi *et al*, 2017), we applied the Mapper, an algorithm that reconstructs a graph-based representation theoretically guaranteed to converge to the underlying topology of a point cloud (Fig 6A, see Materials and Methods for details). Our approach interfaces directly with LIGER, allowing us to define cell trajectories even in the presence of batch effects or biological variation across time points. Additionally, the graph-based representations built by Topologizer summarize large and complex single-cell datasets, allowing easy comprehension of multiple branching and looping events during sequential processes. Furthermore, we devised a novel method for overlaying dynamic RNA velocity information on the topological representation, which enables interpretation of the directionality of cell transitions. The five MC subpopulations identified during cardiac regeneration exhibited distinct functions and temporal distributions, raising the possibility that MC cells may undergo cell-state transition between functionally distinct subtypes. On the topological structure produced by Topologizer, we found that the MC2 subtype was located at the central position, connecting to the other two major MC subpopulations: MC1 and MC3. However, no obvious connections were observed among the latter three subtypes. This topological structure suggests that, if MCs switch their functional states, the state transition would occur primarily between MC2 and any of the other three MC subpopulations (Fig 6B).

Although the topological structure can provide a static relationship between different subtypes, it does not reveal the direction or rate of the change of cell states during the dynamic processes. We therefore integrated our topological representations with RNA velocity (La Manno *et al.*, 2018) to infer the time derivative of gene expression along the temporal trajectory. RNA velocity can infer future states of single cells along dynamic processes by quantifying the relative abundance of un-spliced and spliced mRNAs. The combined analyses allowed us to investigate the transitions between the nodes from different subtypes, as well as within the same subtypes, thereby providing deep insights into the cellular dynamics during heart regeneration. The RNA velocity vector field overlaid on the Topologizer structure revealed potential cell-state transitions from MC2 to MC1 subtype prior to cardiac injury (Fig 6C), suggesting that, under homeostatic condition, macrophages could switch their functional states to respond to physiological changes or environmental challenges. After cardiac injury, MC2 to MC1 transition was initially enhanced as evidenced by the increased length and number of the velocity vectors at 2 and 7 dpi (Fig 6C and D; Wilcoxon test,  $p$ -value  $< 0.05$  for both 2 and 7 dpi compared to those before injured), but these transitions became significantly dampened at 14 dpi (Fig 6C and D; Wilcoxon test,  $p$ -value = 0.87 between uninjured and 14 dpi), further supporting an initially enhanced and subsequent declined pro-inflammatory activity after cardiac injury. Since the rate of MC2 to MC1 transition exhibited dynamic temporal changes, we computed the overall transcriptional dissimilarity between these two MC subtypes. Consistent with RNA velocity analysis, we found that the molecular features of these 2 subtypes were more similar at both 2 and 7 dpi compared to prior to cardiac injury, but they subsequently became molecularly more divergent again at 14 dpi (Fig 6E).

In zebrafish, the kidney marrow is the equivalent of the mammalian bone marrow. Upon injury, the myeloid-derived monocytes from kidney marrow rapidly infiltrate the injured tissue, where they differentiate and expand to generate local macrophage population (Lai *et al.*, 2017). To further determine the relationship between the MC subtypes, we performed integrative analysis across the scRNA-seq datasets of our MC subtypes and myeloid cells from kidney marrow (Tang *et al.*, 2017) (Fig 6F). Interestingly, we found that, at the transcriptome level, the kidney marrow myeloid cells more closely resembled MC2 and MC4, and were more distantly related to the inflammatory MC1 cells. As MC subtype transition occurred from MC2 to MC1,

the integrative analysis suggests that kidney marrow-derived monocytes may differentiate into MC2 subtype first before transiting into MC1 subtype.

To understand the molecular mechanisms underlying the transition from MC2 to MC1, we used nonparametric regression to identify three major clusters of genes showing changes during this MC functional state transition. Cluster 1 shows a trend of initial upregulation followed by downregulation of gene expression and is enriched in GO terms related to mRNA splicing, response to oxidative stress, and ubiquitin-dependent ERAD pathway (Fig 6G and H). Consistent with this observation, previous studies demonstrated that multiple mRNA splicing factors coordinate pro-inflammatory responses in macrophages (Janssen *et al*, 2020; Ostareck & Ostareck-Lederer, 2019). The initially downregulated but subsequently upregulated genes in cluster 2 are enriched in GO terms related to negative regulation of transcription and chromatin organization (Fig 6G and H). Cluster 2 includes genes encoding histone methylation enzymes *Ash1l* (*ash1l*) and *Setdb1* (*setdb1a*), which were found to suppress the expression of pro-inflammatory cytokines (Hachiya *et al*, 2016; Xia *et al*, 2013). The genes involved in inflammatory and immune responses (cluster 3) exhibited an initially upregulated level of expression (Fig 6G and H). As transcriptional regulation, chromatin remodeling and regulation of mRNA splicing are viewed as major regulatory mechanisms of cell fate/state determination, our results indicate that MC functional state transition during cardiac regeneration is governed by orchestrated interplay between chromatin landscape modification, and transcriptional and posttranscriptional regulation.

The topological structure for FBs is clearly different from that of MCs (Fig 6I). The three major FB subtypes, namely FB1-3, were positioned next to each other with FB2 abutting FB1 on one side and FB3 on the other (Fig 6I). Interestingly, RNA velocity analysis indicated that, in sharp contrast to MCs, the FBs did not exhibit clear cell-state transitions among their subtypes (Fig 6J), suggesting they arose from fibroblasts (or epicardial cells) upon injury by distinct molecular mechanisms. This observation is consistent with the finding that FB3 was a unique cluster that was strongly activated following cardiac injury. The eEC subpopulations assumed yet another topological arrangement on the Topologizer structure. eEC1 was found to be located at a central position, connecting to all other eEC subtypes (Fig 6K). RNA velocity analysis on Topologizer structure revealed cell-state transition from eEC1 to eEC2 that was enhanced following cardiac injury (Fig 6L).



## Discussion

Our study provides the first comprehensive *in vivo* characterization of the molecular features and cellular composition of the highly heterogeneous cardiac nonCM populations during the process of cardiac regeneration. Combining single-cell analysis and *in situ* hybridization, we identified multiple novel subpopulations for major nonCM cell types that exhibited distinct tempo-spatial dynamics. Further analysis revealed cooperative interactions among nonCM subtypes, including an important role for the MC in inducing the formation of the transformed fibroblast subtype and activated endocardial endothelial subpopulation. In this study, we also found that the loss of *kit* function compromised transformation and activation of FB and EC subtypes. The mutant hearts also demonstrated reduced cardiomyocyte proliferation and myocardium regeneration. Interestingly, we found that both *kita* and *kitb* are not expressed in the embryonic hearts. Nor are they expressed in the injured or uninjured adult hearts. Instead, RT-PCR analysis revealed high level of *kita* and *kitb* expression in the kidney marrow. The expression of *kit* genes in the kidney marrow is consistent with the monocyte defects reported by Parichy *et al.*, for the single *kita* mutant. These data suggest that *kit* mutant cardiac regeneration defects most likely resulted from the immune defect. Nevertheless, a more detailed characterization of *kita* and *kitb* expression pattern and *kit* mutant might be required to fully determine the exact cause of observed cardiac nonCM and regeneration defects. In this study, we also developed Topologizer, a novel computational approach to uncover the topological structure of nonCMs during heart regeneration. By overlaying RNA velocity information on the topological structure, we identified transitions among macrophage functional states, including a transition to a pro-inflammatory state that occurs more frequently post cardiac injury.

The nonCMs have been increasingly recognized as active participants of cardiac function with important signaling roles that modulate cardiomyocyte behaviors. In spite of substantial progress towards our understanding of nonCM biology (Moore-Morris *et al.*, 2014a; Moore-Morris *et al.*, 2014b; Pinto *et al.*, 2016), a holistic and unbiased picture of nonCM cell types and their molecular features is missing, largely due to limitations of the traditional technologies such as lineage marker expression and population-based omics. In this study, we performed scRNA-seq transcriptome profiling of the nonCMs during heart regeneration. Using the latest developed algorithm LIGER (Welch *et al.*, 2019), we successfully defined shared cell populations across

multiple heterogeneous datasets at multiple time points during cardiac regeneration. Through this analysis, we found that macrophages exhibited the most significant frequency change at the early stage of regeneration. Macrophages are regarded as major professional phagocytic cells (Kantari *et al*, 2008). They function as the first line of defense against pathogen (infection) or damage (sterile inflammation). Classically, macrophages are divided into pro-inflammatory M1 macrophages and anti-inflammatory M2 macrophages (Das *et al*, 2015). The M1 macrophages highly express pro-inflammatory cytokines and exhibit strong microbicidal activity. In contrast, the M2 macrophages are characterized by immune modulation and tissue remodeling function. Nevertheless, macrophages are heterogeneous population characterized by their high diversity and plasticity. In our study, we identified five macrophage subpopulations with distinct molecular signatures and functions. The MC1 subpopulation highly expressed inflammatory cytokine, such as *illb* and *tnfa*, probably resembling the M1 macrophage. The MC2 macrophage, on the other hand, highly expressed genes involved in antigen presentation and is predominating macrophage subpopulation at 14 dpi, likely representing the M2 macrophages (Roszer, 2015). This M1/M2 dichotomy is now being increasingly recognized as an oversimplified classification. Macrophages could actually exist along the M1/M2 continuum. Indeed, we found that as heart regeneration proceeded, the molecular features of MC1 and MC2 became more similar upon cardiac injury. We developed a novel algorithm Topologizer to map the topological relationships of nonCM subpopulations. By overlaying RNA velocity information on the topological structure generated by our novel algorithm Topogolizer, we were able to identify functional state transition from macrophage subtype MC2 to MC1. Further nonparametric regression analysis allowed us to identify splicing and epigenetic factors as the potential regulators of this interesting macrophage functional state transition, suggesting important role of epigenetic mechanisms and posttranscriptional regulation in this transition. Further studies are needed to determine precisely the function of these potential regulators. In contrast to MCs, it is interesting to note that cell state transition doesn't occur between FB subpopulations, indicative of cell type specific cell behaviors.

Our scRNA-seq profiling of nonCMs also allowed us to identify multiple subpopulations for the MCs, FBs and eECs. We observed that the relative nonCM composition changed substantially during cardiac regeneration, suggesting that nonCMs respond dynamically to the ever-changing environment as post-injured hearts proceed through the wound healing and

regeneration processes in an orderly and orchestrated fashion. Yet, we can not rule out the possibility that the observed changes in nonCM cellular composition may be attributable to some extent to the injury-induced variations in digestion susceptibility of the cardiac tissue. More importantly, we found that the appearance of the FB and eEC subpopulations FB3 and eEC2 coincided temporally with that of macrophage MC1 subpopulation, suggesting potential interaction among these nonCM subpopulations. Indeed, we found that genetic and pharmacological perturbation of macrophage functional dynamics compromised interactions among these nonCM subpopulations, indicating a critical role of the inflammatory macrophages in coordinating concerted actions among nonCMs to support and facilitate cardiac regeneration. In sum, this single-cell transcriptomics study provides detailed datasets that reveal nonCMs molecular features and their concerted interactions underlying cardiac regeneration and rigorous analytical scRNA-seq pipelines for predicting nonCM cell behavior changes.

## Materials and Methods

### Zebrafish strains

Zebrafish were raised and maintained under standard laboratory conditions (Westerfield, 2000). All animal husbandry and experiments were conducted in accordance with Institutional Animal Care and Use Committee approved protocol. The zebrafish mutant and transgenic lines used in this study were as follows: *kita*<sup>w34b</sup> (Cooper *et al.*, 2009), *kitb*<sup>sa15348</sup> (Kettleborough *et al.*, 2013), *Tg(tc21:nucEGFP)*<sup>pd41</sup> (Wang *et al.*, 2011), *TgBAC(tnfa:GFP)*<sup>pd1028</sup> (Marjoram *et al.*, 2015).

### Genotyping

Tail clips were lysed at 95°C for 10 minutes in the buffer containing 10 mM TRIS pH 8.0, 50 mM KCl, 0.3% Tween-20, and were then incubated with 1 µg/ml Proteinase K at 55°C for 3 hours (Samsa *et al.*, 2016) followed by deactivation of proteinase K 95°C for 10 minutes. The *kita*<sup>w34b</sup> mutation causes genomic deletion that removes the entire exon 10 and was genotyped in individual samples by PCR amplification (primers: 5'-CCTCAGTTTGAAAATCCCTGCACCA-3' and 5'-ACCCAAAATGGGTTGCGCTAGAAC-3'). The *kitb*<sup>sa15348</sup> mutation was scored in individual samples by digesting the PCR product (primers: 5'-TTGAGGGCTGCTACTTCTGCGC-3' and 5'-CCACTCACTCACCTAGCCGCACAACCAAGCT-3') with Hind III.

### **Apex resection**

Zebrafish were anesthetized by immersion with 0.04% tricaine and immobilized in a dampened foam with ventral side up. A small incision was made between the gills to expose the ventricle. About 20% of ventricular apex was resected using iridectomy scissors as previously described (Poss *et al.*, 2002). After apex resection, fish were returned to a recovery tank with fresh system water. Fish were randomized into surgery or sham group. All procedures and subsequent histological analyses were performed in a blinded fashion.

### **Clodronate liposome and drug treatment**

Clodronate liposomes (5  $\mu$ l at a concentration of 10 mg/ml, Cat# F70101C-A, FormuMax Scientific, Inc.), Dexamethasone (Bollaerts *et al.*, 2019) (3  $\mu$ l at a concentration of 1.65 mg/ml in sterile 0.68% saline; Cat# D1756, Sigma-Aldrich) or PBS was injected intrathoracically one day before apex resection into individual anaesthetized zebrafish using a Picospritzer microinjector (Parker) as described previously (Bevan *et al.*, 2019; Bise & Jazwinska, 2019). The injection was repeated one more time at 3 dpi.

### **Histology**

Adult zebrafish hearts were removed, washed with ice-cold PBS followed by overnight fixation with 4% PFA at 4°C. The cardiac Samples were then embedded in paraffin or OCT for sectioning. Acid fuchsin-orange G (AFOG) staining was performed to assess the ventricular injuries and connective tissue deposition as previously described (Poss *et al.*, 2002). To determine cardiomyocyte proliferation indices, cardiac sections were stained with antibodies against Mef2 (cat# 55609, Anaspec) and PCNA (cat# P8825, Sigma) to assess the proportion of Mef2/PCNA double positive cells in Mef2 positive cells (Wang *et al.*, 2011). Dylight 594-conjugated isolectin B4 (IB4-594, Vector Laboratories, Burlingame, CA) was used to label macrophages (Lai *et al.*, 2017). Anti-Myosin heavy chain (F59, Developmental Studies Hybridoma Bank) was used to label cardiomyocytes (Wang *et al.*, 2011).

### **In situ hybridization**

In situ hybridization (ISH) was performed on 10  $\mu\text{m}$  cardiac sections using digoxigenin-labeled antisense probes as previously described (Poss *et al.*, 2002). All probes were generated using T7 RNA polymerase from zebrafish cDNA. All primers used are listed in Appendix Table S2.

### **Fluorescence *in situ* hybridization**

Briefly, after fixation in 10% neutral buffered formalin overnight, the cardiac tissues were embedded in OCT compound, frozen, and cut at 10  $\mu\text{m}$  using a cryostat. RNAscope *in situ* hybridization was performed according to the manufacturer's protocol using the Multiplex Fluorescent V2 Assay (Advanced Cell Diagnostics, Hayward, CA, #323110). Modifications to the protocol were as follows. Target retrieval was performed for 5 minutes and pre-treatment was done for 15 minutes using Protease III (Advanced Cell Diagnostics, Hayward, CA, #322337). For detection, the TSA Plus Cyanine 3 System (Perkin Elmer, Waltham, MA, #NEL744001KT) and the TSA Plus Cyanine 5 System (Perkin Elmer, Waltham, MA, #NEL745001KT) were used. Finally, slides were coverslipped using Fluor-Gel II with DAPI (Electron Microscopy Sciences, Hatfield, PA, #17985-50).

### **Expression analysis**

Real-time PCR was performed in triplicate using the QuantStudio 6 real-time PCR system and the Power SYBR Green Master Mix (ThermoFischer Scientific). All primers were validated by high resolution melt analysis and gel electrophoresis, and are listed in Appendix Table S2. For quantification, we used the  $\Delta\Delta\text{CT}$  method whereby raw CT values were normalized to *ef1a* as a house-keeping gene. Fold-change was calculated as  $2^{(-\Delta\Delta\text{CT})}$ .

### **Imaging**

For confocal images, cardiac sections were imaged using a Zeiss LSM 700 Confocal Laser Scanning Microscope. AFOG stained cardiac tissues were imaged using an EVOS FL cell imaging system, or a Leica DM IRB microscope. Quantification of cardiomyocyte proliferation was performed as previously described (Kikuchi *et al.*, 2010), by assessing almost all Mef2<sup>+</sup>/PCNA<sup>+</sup> cardiomyocytes near the injury in three ventricular sections. The proliferation index was calculated by averaging the percentages of proliferating Mef2<sup>+</sup> and PCNA<sup>+</sup> cardiomyocytes from three ventricular sections as previously described (Kikuchi *et al.*, 2010).

### **Isolation of primary cells from adult zebrafish ventricles**

Primary cells from zebrafish ventricles subjected to control operation or apex resection were isolated as previously described (Sander *et al*, 2013). Briefly, zebrafish were euthanized by immersion in an ice-cold water. Ventricles were excised and placed in the ice-cold perfusion buffer (1x PBS containing 10 mM HEPES, 30 mM taurine, 5.5 mM glucose and 10 mM BDM). After gently tore apart, tissues were incubated in digestion buffer (5 mg/mL collagenase II, 5 mg/mL collagenase IV and 12.5  $\mu$ M CaCl<sub>2</sub> in perfusion buffer) for 2 hours at 32°C with gently flicking every 10 minutes. Cells were completely disaggregated by pipetting up and down and filtered through a 200- $\mu$ m mesh. Cardiomyocytes and non-cardiomyocytes were separated by 3 times of centrifugation at 200g for 5 minutes at 4°C. The pellet contains cardiomyocytes and the supernatant contains non-cardiomyocytes.

### **Bulk RNA sequencing (RNA-seq) and analysis**

Total RNAs of isolated adult zebrafish cardiomyocytes from eight pooled ventricles were extracted using QIAGEN RNeasy Micro kit as per manufacturer's instructions. RNAs were assessed using Agilent RNA Analysis ScreenTape. RNAs with Integrity Number (RIN) more than 8 were processed to prepare the Illumina library using the TrueSeq Standard mRNA Sample Preparation Kit according to manufacturer's instructions. Agilent BioAnalyzer was applied to assess the quality of the libraries to ensure an average fragment size of ~280 bp, and quantify the amount of the libraries. The normalized libraries were pooled and sequenced with 50 bp paired-end using Illumina HiSeq 4000. The alignment was performed using BMAP/38.12 (Bushnell, 2014) against the reference zebrafish genome UCSC\_DanRer10. Overall mapping rates were above 92% for all samples (Appendix Table S3). Gene counts were obtained using featureCounts (Liao *et al*, 2014). The raw counts were normalized to sequencing depth using counts per million (CPM) (Shi *et al*, 2015). Genes were considered as differentially expressed (DEGs) if Benjamini-Hochberg adjusted P-value was < 0.05 as raw counts were analyzed by DESeq2 (Love *et al*, 2014). The DEGs between wildtype and *kit* mutant CMs were then analyzed by Gene Set Enrichment Analysis (GSEA) (Subramanian *et al*, 2005). The list of positive and negative regulators of cell cycle was derived from previous studies (Lu *et al*, 2019; Malumbres,

2014; Yu, 2007). The RNA-seq data was deposited in the Gene Expression Omnibus (GEO) database with accession number GSE145979.

## **Single-cell RNA sequencing**

### ***Data collection***

As described above, isolated non-cardiomyocytes from zebrafish ventricles subjected to control operation or apex resection at indicated time points were loaded into the Chromium controller (10X Genomics) and processed with Chromium Single Cell Reagent Kit v2 (10X Genomics). Following droplet generation and barcoding, cDNA was synthesized and amplified with 12 cycles of PCR as per the manufacturer's instruction. The cDNA was further processed to construct Illumina sequencing libraries and sequenced using NextSeq 500/550 kit v2.5 according to the manufacturer's instructions. In order to collect at least 4000 cells at each time point, a total of 14 samples from two genotypes (wildtype and *kit* mutant) and four time points (uninjured—day 0, post-injury—2, 7, and 14 dpi) were collected and sequenced in two lanes. The scRNA-seq data was deposited in the Gene Expression Omnibus (GEO) database with accession number GSE145980.

### ***Pre-processing***

Raw sequencing data was demultiplexed, debarcoded, mapped to UCSC\_DanRer10 genome, and counted by UNC TGL bioinformatic staff following the Cell Ranger/2.0.2 pipeline. The Cell Ranger output of collapsed UMI counts were then processed and analyzed in R. A total of 38,181 cells from wild-type and 49,133 cells from *kit* mutants were captured. To ensure the quality of our data and analysis, a series of quality control steps have then been performed. First, raw UMI counts were imported into Seurat/2.3.4 (Butler *et al*, 2018; Macosko *et al*, 2015) and low-quality cells expressing  $\leq 200$  unique genes were filtered out. Next, data from each sample was inspected carefully with Seurat by clustering and marker identification using default settings. Clusters expressing high levels of *ckma*, *tnnt2* and *nppa* were identified as residual cardiomyocytes and removed. Furthermore, clusters expressing high levels of canonical markers of any two of non-cardiomyocyte cell types were considered as doublets and also excluded from further analyses. The observed frequencies of doublets were consistent with the expected frequencies of the two nonCM cell types involved. After removal of low-quality cells, doublets

and residual cardiomyocytes, a total of 61,977 nonCMs were analyzed in this study, including 27,165 wildtype and 34,812 *kit* mutant cells. A summary of sequencing metrics of each sample post-QC was listed in Appendix Table S1. Median genes detected per cell in each sample ranged from 615 to 1431 in our samples, which are relatively high considering the small size of zebrafish cells, suggesting the high quality of our data.

### ***Data analyses with LIGER***

In order to integrate multiple single-cell datasets generated along heart regeneration, a recently developed algorithm LIGER/0.3.1 was employed (Welch *et al.*, 2019), because of its outstanding performance in batch effect correction (Stuart & Satija, 2019). For instance, when analyzing wildtype samples from all time points, LIGER clustered cells by cell type (Fig 2) not by time points or batches (Appendix Fig S2B and C). But clustering results of the CCA method in Seurat v2 were more driven by time points instead (Appendix Fig 2A). We performed integration analysis of all the wildtype samples (Fig 2), as well as all *kit* mutant samples (Fig 4), from all four time points using LIGER. Briefly speaking, first, post-QC doublets-removed digit gene expression (dge) data of nonCMs from each sample were normalized (the *normalize* function of LIGER). Highly variable genes (HVG) with variance > 0.1 were selected for each sample and union was taken (the *selectGenes* function). After scaling (the *scaleNotCenter* function), factorization was performed with the *optimizeALS* function. The number of factors *k* (number of estimated cell types in the dataset) was determined to be 45 by running the *suggestK* function for both wildtype and *kit* mutant analyses. Alignment was then performed using the *quantileAlignSNF* function with a resolution of 2. After alignment, genes (markers) in each factor were inspected in the word clouds plot generated from the *plotWordClouds* function. Nonspecific factors (present in more than a few clusters) entirely composed of mitochondrial, ribosomal, or stress genes were identified as technical factors and thus excluded by omitting them in *dims.use* during rerunning of the *quantileAlignSNF* function. Biological variation sometimes also appeared as nonspecific factors obscuring cell type assignment such as cell cycle stage and immune response. In such cases, these corresponding factors were also excluded. When combining all time points of wildtype samples, doublets expressing high levels of canonical markers of two cell types appeared again, especially for abundant cell types, and were removed for downstream analysis. Then, dimension reduction was performed for visualization



with the *runTSNE* function. To further confirm that our clustering is not affected by technical confounding factors, we overlaid number of genes (nGene) and UMIs (nUMI), and percentage of mitochondrial reads to tSNE plots of each analysis we performed. We consistently observed no impact of these widely-used cell quality measurements on clustering results (Appendix Fig S2D), confirming the high quality of our data and the rigorosity of LIGER. For joint analysis of all the wildtype samples or *kit* mutant samples, cell type identity was assigned to each cluster based on their expression pattern of canonical markers. It is hard to determine whether the erythrocytes and thrombocytes in our dataset are local or from circulation. Therefore, they were only included in the tSNE plots but excluded from any other downstream analysis. Genes differentially expressed in each cell type were then detected using the *FindAllMarkers* function in Seurat (Appendix Table S2). Ribosomal structural and mitochondrial genes were excluded from the gene lists. GO analysis was performed for the filtered gene lists using DAVID/6.8 (Huang da *et al*, 2009). Cell cycle assignment was performed with the *CellCycleScoring* function in Seurat using zebrafish cell cycle genes from a previous study (Lu *et al.*, 2019).

A second round of integration and clustering analysis was performed for each of the three cell types we are interested in (MC, FB, and EC) with LIGER. Due to less heterogeneity in the datasets, a smaller  $k$  between 14 and 25 and a smaller resolution between 0.5 and 1 were selected for factorization, alignment and clustering. Technical factors were identified and excluded as above while all biological factors such as cell cycle status were preserved during alignment. For each cell type, clusters of high similarity, which have high Pearson correlation coefficients and small numbers of cluster-specific markers, were merged into one subpopulation. Differential gene expression, GO enrichment analysis and cell cycle assignment were performed as described above.

### ***RNA velocity***

We calculated RNA velocity using the *velocyto.R* package (<https://github.com/velocyto-team/velocyto.R>) following the instruction of the package. Gene-relative velocity was estimated using a gamma fit based on extreme quantiles.

### ***Trajectory analysis with Topologizer***

We developed a novel approach, Topologizer, which leverages the mathematics of algebraic topology to characterize the “shape” of the cellular manifold during heart regeneration

(<https://github.com/welch-lab/topologizer>). Building on the pioneering work of Rizvi et al. (Rizvi *et al.*, 2017), we applied the Mapper (Singh *et al.*, 2007), an algorithm that reconstructs a graph-based representation theoretically guaranteed to converge to the underlying topology of a point cloud (Fig 5A). However, unlike Rizvi et al., our approach interfaces directly with LIGER, allowing us to define cell trajectories even in the presence of batch effects or biological variation across time points. We first use LIGER to jointly factorize multiple datasets (such as from multiple time points). Next, we construct a function mapping from the LIGER factors to 2 or 3 dimensions (a filter function, in the language of topology). To do this, we calculate diffusion components, then perform Uniform Manifold Approximation and Projection (UMAP) to map to two or three dimensions. UMAP is a powerful nonlinear dimensionality reduction technique built on the mathematics of algebraic geometry, and is designed to find topology-preserving embeddings (McInnes & Healy, 2018). Then, we tile the range of the filter function (to construct a “cover”) using overlapping hypercubes (squares for 2D, cubes for 3D). We next cluster the cells within each hypercube using hierarchical clustering on the original iNMF factors. Clusters sharing at least one cell are then connected, yielding a graph that captures the topology of the high-dimensional data (a “simplicial complex”). We then visualize this graph in 2D using an interactive force-directed layout that iteratively updates when the user moves the nodes using the mouse. We use the functions in the *kmapper* python package to carry out these steps. Our approach (based on the Mapper (Singh *et al.*, 2007)) has a strong theoretical foundation—the Mapper has been shown to converge to the true topology of a point cloud (Carriere & Oudot, 2018) as the sampling density increases. Consequently, Topologizer can capture any topological features of cells undergoing a dynamic transition, including extreme points, branches, and loops/holes.

We also devised a novel method for overlaying dynamic RNA velocity information on the topological representation, which enables interpretation of the directionality of cell transitions. To do this, we first calculated the velocity vectors for each individual cell, then averaged the vectors within each node of the topological representation (corresponding to a cluster of cells within a hypercube). We then used the correlations among the averaged velocities to project them onto the topological representation (*show.velocity.on.embedding.cor* function from *velocityto.R* package with default parameters).

To investigate the genes showing dynamic changes along the topological representation, we implemented an interactive visualization that allows users to select nodes of interest from the topological representation. We then fit a principal curve through the UMAP coordinates of the selected cells, project the cells onto the curve, and calculated distance along the curve from a selected starting node. Using this distance, we fit a generalized additive model as previously described (Liu *et al.*, 2017; Trapnell *et al.*, 2014) to identify genes with trends that significant differ from the null hypothesis of a horizontal line. We then clustered these genes to identify kinetic trends using *k*-medoid clustering, as previously described (Liu *et al.*, 2017; Trapnell *et al.*, 2014).

### **Statistics**

All data were presented as mean  $\pm$  SEM. Prism was used for the statistical analysis. For the difference between two groups, a two-tailed independent sample t test was used to compare the mean. For the difference among three or more groups, a one-way or two-way ANOVA was used to compare the mean. A Least Significant Difference (LSD) test or the Tukey's multiple comparison test was further applied for post hoc analysis to detect the pairwise difference while adjusting for multiplicity. P-value  $< 0.05$  was considered statistically significant.

### **Data availability**

RNA-sequencing data are deposited in GEO under accession number GSE145982 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE145982>). .

### **Acknowledgements**

We thank Drs. Kenneth Poss and Michel Bagnat for providing zebrafish stocks. We also thank the University of North Carolina (UNC) Microscopy Service Laboratory for confocal microscope use, UNC Translational Genomics Laboratory Core, High Throughput Sequencing Facility Core for technical and bioinformatics support, Histology Research Core Facility for RNAscope *in situ* hybridization and the UNC Zebrafish Aquaculture Core Facility for fish care and microscope use. We thank members of the Qian laboratory and the Liu laboratory for helpful discussions and critical reviews of the manuscript. The Liu lab is supported by NIH/NHLBI R01 grants R01 HL139976 and R01 HL139880, and American Heart Association (AHA) Established

Investigator Award 20EIA35320128. This work is also supported by AHA 18TPA34180058, 20EIA35310348 and NIH R35HL155656 to L. Q., and NIH grant (R01 HG010883) and Chan-Zuckerberg Initiative grant 2018-183155 to J.D.W. This work is also partially supported by NIH grants R01 GM105785 to Y.L.

### **Author contributions**

H.M. and J.L. conceived and designed the study. H.M. designed and performed scRNA-seq and prepared samples for bulk RNA-seq. L.W. helped performing scRNA-seq. H. M.. characterized *kit* mutant phenotypes. D. F. performed in situ hybridization. Y. D. performed qRT-PCR and drug treatment experiments. T.A.G. performed immunohistochemistry. Z. H. processed the initial scRNA-seq data. H.M. Z.L. Y.Y. J.W. performed data analysis and modeling. C.L. performed data analysis. C.D.C provided critical reagents. J.W. developed the Topologizer approach. Z.L., Y.Y., H.M., J.W., L.Q. and J.L. wrote the manuscript with extensive input from all authors. Y.L. supervised the data analysis. L.Q. and J.L. provided funding and overall supervision.

### **Conflict of interest**

The authors declare no competing financial interest.

### **References**

- Aderem A (2003) Phagocytosis and the inflammatory response. *J Infect Dis* 187 Suppl 2: S340-345
- Altmann C, Vasic V, Hardt S, Heidler J, Haussler A, Wittig I, Schmidt MHH, Tegeder I (2016) Progranulin promotes peripheral nerve regeneration and reinnervation: role of notch signaling. *Molecular neurodegeneration* 11: 69
- Athnasiadis EI, Botthof JG, Andres H, Ferreira L, Lio P, Cvejic A (2017) Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat Commun* 8: 2045
- Aurora AB, Porrello ER, Tan W, Mahmoud AI, Hill JA, Bassel-Duby R, Sadek HA, Olson EN (2014) Macrophages are required for neonatal heart regeneration. *The Journal of clinical investigation* 124: 1382-1392

Bagley RG, Honma N, Weber W, Boutin P, Rouleau C, Shankara S, Kataoka S, Ishida I, Roberts BL, Teicher BA (2008) Endosialin/TEM 1/CD248 is a pericyte marker of embryonic and tumor neovascularization. *Microvascular research* 76: 180-188

Bassat E, Mutlak YE, Genzelinakh A, Shadrin IY, Baruch Umansky K, Yifa O, Kain D, Rajchman D, Leach J, Riabov Bassat D *et al* (2017) The extracellular matrix protein agrin promotes heart regeneration in mice. *Nature* 547: 179-184

Bertrand JY, Jalil A, Klaine M, Jung S, Cumano A, Godin I (2005) Three pathways to mature macrophages in the early mouse yolk sac. *Blood* 106: 3004-3011

Bevan L, Lim ZW, Venkatesh B, Riley PR, Martin P, Richardson RJ (2019) Specific macrophage populations promote both cardiac scar deposition and subsequent resolution in adult zebrafish. *Cardiovasc Res* 116: 1357-1371.

Bise T, Jazwinska A (2019) Intrathoracic Injection for the Study of Adult Zebrafish Heart. *J Vis Exp* 14: doi: 10.3791/59724.

Bollaerts I, Van Houcke J, Beckers A, Lemmens K, Vanhunsel S, De Groef L, Moons L (2019) Prior Exposure to Immunosuppressors Sensitizes Retinal Microglia and Accelerates Optic Nerve Regeneration in Zebrafish. *Mediators Inflamm* 2019: 6135795

Bubb KJ, Aubdool AA, Moyes AJ, Lewis S, Drayton JP, Tang O, Mehta V, Zachary IC, Abraham DJ, Tsui J *et al* (2019) Endothelial C-Type Natriuretic Peptide Is a Critical Regulator of Angiogenesis and Vascular Remodeling. *Circulation* 139: 1612-1628

Bushnell B (2014) BBMap: A Fast, Accurate, Splice-Aware Aligner. *Report Number: LBNL-7065E*

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature biotechnology* 36: 411-420

Cao J, Navis A, Cox BD, Dickson AL, Gemberling M, Karra R, Bagnat M, Poss KD (2016) Single epicardial cell transcriptome sequencing identifies Caveolin 1 as an essential factor in zebrafish heart regeneration. *Development* 143: 232-243

Carmona SJ, Teichmann SA, Ferreira L, Macaulay IC, Stubbington MJ, Cvejic A, Gfeller D (2017) Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. *Genome research* 27: 451-461

Carriere M, Oudot SY (2018) Structure and Stability of the 1-Dimensional Mapper. *Foundations of Computational Mathematics* 18: 1333-11396

Cooper CD, Linbo TH, Raible DW (2009) Kit and foxd3 genetically interact to regulate melanophore survival in zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* 238: 875-886

Costa RA, Cardoso JC, Power DM (2017) Evolution of the angiopoietin-like gene family in teleosts and their role in skin regeneration. *BMC evolutionary biology* 17: 14

Cui Y, Zheng Y, Liu X, Yan L, Fan X, Yong J, Hu Y, Dong J, Li Q, Wu X *et al* (2019) Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell reports* 26: 1934-1950 e1935

D'Uva G, Aharonov A, Lauriola M, Kain D, Yahalom-Ronen Y, Carvalho S, Weisinger K, Bassat E, Rajchman D, Yifa O *et al* (2015) ERBB2 triggers mammalian heart regeneration by promoting cardiomyocyte dedifferentiation and proliferation. *Nature cell biology* 17: 627-638

Das A, Sinha M, Datta S, Abas M, Chaffee S, Sen CK, Roy S (2015) Monocyte and macrophage plasticity in tissue repair and regeneration. *Am J Pathol* 185: 2596-2606

Davies LC, Rosas M, Jenkins SJ, Liao CT, Scurr MJ, Brombacher F, Fraser DJ, Allen JE, Jones SA, Taylor PR (2013) Distinct bone marrow-derived and tissue-resident macrophage lineages proliferate at key stages during inflammation. *Nat Commun* 4: 1886

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

Epelman S, Lavine KJ, Beaudin AE, Sojka DK, Carrero JA, Calderon B, Brija T, Gautier EL, Ivanov S, Satpathy AT *et al* (2014) Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity* 40: 91-104

Farbehi N, Patrick R, Dorison A, Xaymardan M, Janbandhu V, Wystub-Lis K, Ho JW, Nordon RE, Harvey RP (2019) Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury. *Elife* 8

Glisic B, Mihaljevic I, Popovic M, Zaja R, Loncar J, Fent K, Kovacevic R, Smital T (2015) Characterization of glutathione-S-transferases in zebrafish (*Danio rerio*). *Aquatic toxicology* 158: 50-62

Gonzalez-Rosa JM, Burns CE, Burns CG (2017) Zebrafish heart regeneration: 15 years of discoveries. *Regeneration (Oxf)* 4: 105-123

Gore AV, Swift MR, Cha YR, Lo B, McKinney MC, Li W, Castranova D, Davis A, Mukoyama YS, Weinstein BM (2011) Rspo1/Wnt signaling promotes angiogenesis via Vegfc/Vegfr3. *Development* 138: 4875-4886

Habeck H, Odenthal J, Walderich B, Maischein H, Schulte-Merker S, Tübingen screen c (2002) Analysis of a zebrafish VEGF receptor mutant reveals specific disruption of angiogenesis. *Curr Biol* 12: 1405-1412

Hachiya R, Shiihashi T, Shirakawa I, Iwasaki Y, Matsumura Y, Oishi Y, Nakayama Y, Miyamoto Y, Manabe I, Ochi K *et al* (2016) The H3K9 methyltransferase Setdb1 regulates TLR4-mediated inflammatory responses in macrophages. *Scientific reports* 6: 28845

Harrison MR, Bussmann J, Huang Y, Zhao L, Osorio A, Burns CG, Burns CE, Sucov HM, Siekmann AF, Lien CL (2015) Chemokine-guided angiogenesis directs coronary vasculature formation in zebrafish. *Developmental cell* 33: 442-454

Harvie EA, Huttenlocher A (2015) Neutrophils in host defense: new insights from zebrafish. *J Leukoc Biol* 98: 523-537

Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44-57

Huang W, Febbraio M, Silverstein RL (2011) CD9 tetraspanin interacts with CD36 on the surface of macrophages: a possible regulatory influence on uptake of oxidized low density lipoprotein. *PLoS One* 6: e29092

Ivey MJ, Tallquist MD (2016) Defining the Cardiac Fibroblast. *Circulation journal : official journal of the Japanese Circulation Society* 80: 2269-2276

Janssen WJ, Danhorn T, Harris C, Mould KJ, Lee FF, Hedin BR, D'Alessandro A, Leach SM, Alper S (2020) Inflammation-Induced Alternative Pre-mRNA Splicing in Mouse Alveolar Macrophages. *G3* 10: 555-567

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

Kaji K, Takeshita S, Miyake K, Takai T, Kudo A (2001) Functional association of CD9 with the Fc gamma receptors in macrophages. *J Immunol* 166: 3256-3265

Kantari C, Pederzoli-Ribeil M, Witko-Sarsat V (2008) The role of neutrophils and monocytes in innate immunity. *Contrib Microbiol* 15: 118-146

Kato K, Martinez C, Russell S, Nurden P, Nurden A, Fiering S, Ware J (2004) Genetic deletion of mouse platelet glycoprotein Ibbeta produces a Bernard-Soulier phenotype with increased alpha-granule size. *Blood* 104: 2339-2344

Kettleborough RN, Busch-Nentwich EM, Harvey SA, Dooley CM, de Bruijn E, van Eeden F, Sealy I, White RJ, Herd C, Nijman IJ *et al* (2013) A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature* 496: 494-497

Kierdorf K, Erny D, Goldmann T, Sander V, Schulz C, Perdiguero EG, Wieghofer P, Heinrich A, Riemke P, Holscher C *et al* (2013) Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat Neurosci* 16: 273-280

Kikuchi K, Gupta V, Wang J, Holdway JE, Wills AA, Fang Y, Poss KD (2011a) tcf21+ epicardial cells adopt non-myocardial fates during zebrafish heart development and regeneration. *Development* 138: 2895-2902

Kikuchi K, Holdway JE, Major RJ, Blum N, Dahn RD, Begemann G, Poss KD (2011b) Retinoic acid production by endocardium and epicardium is an injury response essential for zebrafish heart regeneration. *Developmental cell* 20: 397-404

Kikuchi K, Holdway JE, Werdich AA, Anderson RM, Fang Y, Egnaczyk GF, Evans T, Macrae CA, Stainier DY, Poss KD (2010) Primary contribution to zebrafish heart regeneration by gata4(+) cardiomyocytes. *Nature* 464: 601-605

Kim RY, Robertson EJ, Solloway MJ (2001) Bmp6 and Bmp7 are required for cushion formation and septation in the developing mouse heart. *Dev Biol* 235: 449-466

Klotz L, Norman S, Vieira JM, Masters M, Rohling M, Dube KN, Bollini S, Matsuzaki F, Carr CA, Riley PR (2015) Cardiac lymphatics are heterogeneous in origin and respond to injury. *Nature* 522: 62-67

Kyritsis N, Kizil C, Zocher S, Kroehne V, Kaslin J, Freudenreich D, Iltzsche A, Brand M (2012) Acute inflammation initiates the regenerative response in the adult zebrafish brain. *Science* 338: 1353-1356

La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastrioti ME, Lonnerberg P, Furlan A *et al* (2018) RNA velocity of single cells. *Nature* 560: 494-498



Laflamme MA, Murry CE (2011) Heart regeneration. *Nature* 473: 326-335

Lai SL, Marin-Juez R, Moura PL, Kuenne C, Lai JKH, Tseke AT, Guenther S, Looso M, Stainier DY (2017) Reciprocal analyses in zebrafish and medaka reveal that harnessing the immune response promotes cardiac regeneration. *Elife* 6

Larson JD, Wadman SA, Chen E, Kerley L, Clark KJ, Eide M, Lippert S, Nasevicius A, Ekker SC, Hackett PB *et al* (2004) Expression of VE-cadherin in zebrafish embryos: a new tool to evaluate vascular development. *Developmental dynamics : an official publication of the American Association of Anatomists* 231: 204-213

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

Leach JP, Heallen T, Zhang M, Rahmani M, Morikawa Y, Hill MC, Segura A, Willerson JT, Martin JF (2017) Hippo pathway deficiency reverses systolic heart failure after infarction. *Nature* 550: 260-264

Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923-930

Lin HF, Traver D, Zhu H, Dooley K, Paw BH, Zon LI, Handin RI (2005) Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. *Blood* 106: 3803-3810

Liu Z, Wang L, Welch JD, Ma H, Zhou Y, Vaseghi HR, Yu S, Wall JB, Alimohamadi S, Zheng M *et al* (2017) Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *Nature* 551: 100-104

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550

Lu CJ, Fan XY, Guo YF, Cheng ZC, Dong J, Chen JZ, Li LY, Wang MW, Wu ZK, Wang F *et al* (2019) Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development. *Journal of molecular cell biology* 11: 435-447

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM *et al* (2015) Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161: 1202-1214

Mahmoud AI, Kocabas F, Muralidhar SA, Kimura W, Koura AS, Thet S, Porrello ER, Sadek HA (2013) Meis1 regulates postnatal cardiomyocyte cell cycle arrest. *Nature* 497: 249-253

Malumbres M (2014) Cyclin-dependent kinases. *Genome Biol* 15: 122

Marjoram L, Alvers A, Deerhake ME, Bagwell J, Mankiewicz J, Cocchiaro JL, Beerman RW, Willer J, Sumigray KD, Katsanis N *et al* (2015) Epigenetic control of intestinal barrier function and inflammation in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America* 112: 2770-2775

McDonald AI, Shirali AS, Aragon R, Ma F, Hernandez G, Vaughn DA, Mack JJ, Lim TY, Sunshine H, Zhao P *et al* (2018) Endothelial Regeneration of Large Vessels Is a Biphasic Process Driven by Local Cells with Distinct Proliferative Capacities. *Cell Stem Cell* 23: 210-225 e216

McInnes L, Healy J (2018) UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* 1802.03426 [stat.ML]

Moore-Morris T, Guimaraes-Camboa N, Banerjee I, Zambon AC, Kisseleva T, Velayoudon A, Stallcup WB, Gu Y, Dalton ND, Cedenilla M *et al* (2014a) Resident fibroblast lineages mediate pressure overload-induced cardiac fibrosis. *The Journal of clinical investigation* 124: 2921-2934

Moore-Morris T, Tallquist MD, Evans SM (2014b) Sorting out where fibroblasts come from. *Circulation research* 115: 602-604

Moore C, Richens JL, Hough Y, Ucanok D, Malla S, Sang F, Chen Y, Elworthy S, Wilkinson RN, Gering M (2018) Gfi1a and Gfi1b set the pace for primitive erythroblast differentiation from hemangioblasts in the zebrafish embryo. *Blood advances* 2: 2589-2606

Morikawa Y, Heallen T, Leach J, Xiao Y, Martin JF (2017) Dystrophin-glycoprotein complex sequesters Yap to inhibit cardiomyocyte proliferation. *Nature* 547: 227-231

Munch J, Grivas D, Gonzalez-Rajal A, Torregrosa-Carrion R, de la Pompa JL (2017) Notch signalling restricts inflammation and serpine1 expression in the dynamic endocardium of the regenerating zebrafish heart. *Development* 144: 1425-1440

Nakada Y, Canseco DC, Thet S, Abdisalaam S, Asaithamby A, Santos CX, Shah AM, Zhang H, Faber JE, Kinter MT *et al* (2017) Hypoxia induces heart regeneration in adult mice. *Nature* 541: 222-227

Nemer G, Nemer M (2002) Cooperative interaction between GATA5 and NF-ATc regulates endothelial-endocardial differentiation of cardiogenic cells. *Development* 129: 4045-4055

Norris RA, Kern CB, Wessels A, Wirrig EE, Markwald RR, Mjaatvedt CH (2005) Detection of betaig-H3, a TGFbeta induced gene, during cardiac development and its complementary pattern with periostin. *Anatomy and embryology* 210: 13-23

Okuda KS, Astin JW, Misa JP, Flores MV, Crosier KE, Crosier PS (2012) *lyve1* expression reveals novel lymphatic vessels and new mechanisms for lymphatic vessel development in zebrafish. *Development* 139: 2381-2391

Ostareck DH, Ostareck-Lederer A (2019) RNA-Binding Proteins in the Control of LPS-Induced Macrophage Response. *Frontiers in genetics* 10: 31

Parichy DM, Rawls JF, Pratt SJ, Whitfield TT, Johnson SL (1999) Zebrafish *sparse* corresponds to an orthologue of *c-kit* and is required for the morphogenesis of a subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. *Development* 126: 3425-3436

Paw BH, Davidson AJ, Zhou Y, Li R, Pratt SJ, Lee C, Trede NS, Brownlie A, Donovan A, Liao EC *et al* (2003) Cell-specific mitotic defect and dyserythropoiesis associated with erythroid band 3 deficiency. *Nature genetics* 34: 59-64

Person AD, Garriock RJ, Krieg PA, Runyan RB, Klewer SE (2005) *Frzb* modulates Wnt-9a-mediated beta-catenin signaling during avian atrioventricular cardiac cushion development. *Dev Biol* 278: 35-48

Pinto AR, Ilinykh A, Ivey MJ, Kuwabara JT, D'Antoni ML, Debuque R, Chandran A, Wang L, Arora K, Rosenthal NA *et al* (2016) Revisiting Cardiac Cellular Composition. *Circulation research* 118: 400-409

Pols MS, Klumperman J (2009) Trafficking and function of the tetraspanin CD63. *Exp Cell Res* 315: 1584-1592

Porrello ER, Mahmoud AI, Simpson E, Hill JA, Richardson JA, Olson EN, Sadek HA (2011) Transient regenerative potential of the neonatal mouse heart. *Science* 331: 1078-1080

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

Price EL, Vieira JM, Riley PR (2019) Model organisms at the heart of regeneration. *Disease models & mechanisms* 12

Riley PR (2012) An epicardial floor plan for building and rebuilding the mammalian heart. *Current topics in developmental biology* 100: 233-251

Rizvi AH, Camara PG, Kandror EK, Roberts TJ, Schieren I, Maniatis T, Rabadan R (2017) Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nature biotechnology* 35: 551-560

- Roszer T (2015) Understanding the Mysterious M2 Macrophage through Activation Markers and Effector Mechanisms. *Mediators Inflamm* 2015: 816460
- Samsa LA, Ito CE, Brown DR, Qian L, Liu J (2016) IgG-Containing Isoforms of Neuregulin-1 Are Dispensable for Cardiac Trabeculation in Zebrafish. *PLoS One* 11: e0166734
- Sanchez-Iranzo H, Galardi-Castilla M, Sanz-Morejon A, Gonzalez-Rosa JM, Costa R, Ernst A, Sainz de Aja J, Langa X, Mercader N (2018) Transient fibrosis resolves via fibroblast inactivation in the regenerating zebrafish heart. *Proceedings of the National Academy of Sciences of the United States of America* 115: 4188-4193
- Sander V, Sune G, Jopling C, Morera C, Izpisua Belmonte JC (2013) Isolation and in vitro culture of primary cardiomyocytes from adult zebrafish hearts. *Nat Protoc* 8: 800-809
- Santoro MM, Pesce G, Stainier DY (2009) Characterization of vascular mural cells during zebrafish development. *Mechanisms of development* 126: 638-649
- Schroder B (2016) The multifaceted roles of the invariant chain CD74--More than just a chaperone. *Biochim Biophys Acta* 1863: 1269-1281
- Shi W, Liao Y, Willis SN, Taubenheim N, Inouye M, Tarlinton DM, Smyth GK, Hodgkin PD, Nutt SL, Corcoran LM (2015) Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. *Nat Immunol* 16: 663-673
- Silva M, Videira PA, Sackstein R (2017) E-Selectin Ligands in the Human Mononuclear Phagocyte System: Implications for Infection, Inflammation, and Immunotherapy. *Frontiers in immunology* 8: 1878
- Singh G, Mémoli F, Carlsson G (2007) Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition. *Symposium on Point Based Graphics* 7: 91-100
- Sinha T, Lin L, Li D, Davis J, Evans S, Wynshaw-Boris A, Wang J (2015) Mapping the dynamic expression of Wnt11 and the lineage contribution of Wnt11-expressing cells during early mouse development. *Dev Biol* 398: 177-192
- Snider P, Standley KN, Wang J, Azhar M, Doetschman T, Conway SJ (2009) Origin of cardiac fibroblasts and the role of periostin. *Circulation research* 105: 934-947
- Spilsbury K, O'Mara MA, Wu WM, Rowe PB, Symonds G, Takayama Y (1995) Isolation of a novel macrophage-specific gene by differential cDNA analysis. *Blood* 85: 1620-1629

Stremmel C, Schuchert R, Wagner F, Thaler R, Weinberger T, Pick R, Mass E, Ishikawa-Ankerhold HC, Margraf A, Hutter S *et al* (2018) Yolk sac macrophage progenitors traffic to the embryo during defined stages of development. *Nat Commun* 9: 75

Stuart T, Satija R (2019) Integrative single-cell analysis. *Nature reviews Genetics* 20: 257-272

Su SA, Xie Y, Zhang Y, Xi Y, Cheng J, Xiang M (2019) Essential roles of EphrinB2 in mammalian heart: from development to diseases. *Cell communication and signaling : CCS* 17: 29

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 102: 15545-15550

Sultana N, Zhang L, Yan J, Chen J, Cai W, Razzaque S, Jeong D, Sheng W, Bu L, Xu M *et al* (2015) Resident c-kit(+) cells in the heart are not cardiac stem cells. *Nat Commun* 6: 8701

Tang Q, Iyer S, Lobbardi R, Moore JC, Chen H, Lareau C, Hebert C, Shaw ML, Neftel C, Suva ML *et al* (2017) Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. *The Journal of experimental medicine* 214: 2875-2887

Tao G, Kahr PC, Morikawa Y, Zhang M, Rahmani M, Heallen TR, Li L, Sun Z, Olson EN, Amendt BA *et al* (2016) Pitx2 promotes heart repair by activating the antioxidant response after cardiac injury. *Nature* 534: 119-123

Torraca V, Otto NA, Tavakoli-Tameh A, Meijer AH (2017) The inflammatory chemokine Cxcl18b exerts neutrophil-specific chemotaxis via the promiscuous chemokine receptor Cxcr2 in zebrafish. *Developmental and comparative immunology* 67: 57-65

Touma M, Kang X, Gao F, Zhao Y, Cass AA, Biniwale R, Xiao X, Eghbali M, Coppola G, Reemtsen B *et al* (2017) Wnt11 regulates cardiac chamber development and disease during perinatal maturation. *JCI insight* 2

Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature biotechnology* 32: 381-386

- Tsuruma K, Saito Y, Okuyoshi H, Yamaguchi A, Shimazawa M, Goldman D, Hara H (2018) Granulin 1 Promotes Retinal Regeneration in Zebrafish. *Investigative ophthalmology & visual science* 59: 6057-6066
- Tzahor E, Poss KD (2017) Cardiac regeneration strategies: Staying young at heart. *Science* 356: 1035-1039
- van Berlo JH, Kanisicak O, Maillet M, Vagnozzi RJ, Karch J, Lin SC, Middleton RC, Marban E, Molkentin JD (2014) c-kit<sup>+</sup> cells minimally contribute cardiomyocytes to the heart. *Nature* 509: 337-341
- van Impel A, Zhao Z, Hermkens DM, Roukens MG, Fischer JC, Peterson-Maduro J, Duckers H, Ober EA, Ingham PW, Schulte-Merker S (2014) Divergence of zebrafish and mouse lymphatic cell fate specification pathways. *Development* 141: 1228-1238
- Walters KB, Green JM, Surfus JC, Yoo SK, Huttenlocher A (2010) Live imaging of neutrophil motility in a zebrafish model of WHIM syndrome. *Blood* 116: 2803-2811
- Walton EM, Cronan MR, Beerman RW, Tobin DM (2015) The Macrophage-Specific Promoter mfap4 Allows Live, Long-Term Analysis of Macrophage Behavior during Mycobacterial Infection in Zebrafish. *PLoS One* 10: e0138949
- Wang J, Karra R, Dickson AL, Poss KD (2013) Fibronectin is deposited by injury-activated epicardial cells and is necessary for zebrafish heart regeneration. *Dev Biol* 382: 427-435
- Wang J, Panakova D, Kikuchi K, Holdway JE, Gemberling M, Burris JS, Singh SP, Dickson AL, Lin YF, Sabeih MK *et al* (2011) The regenerative capacity of zebrafish reverses cardiac failure caused by genetic cardiomyocyte depletion. *Development* 138: 3421-3430
- Welch JD, Kozareva V, Ferreira A, Vanderburg C, Martin C, Macosko EZ (2019) Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell* 177: 1873-1887 e1817
- Westerfield M (2000) *The Zebrafish Book: A Guide for The Laboratory Use of Zebrafish (Danio rerio)*. University of Oregon Press, Eugene, OR
- Wittamer V, Bertrand JY, Gutschow PW, Traver D (2011) Characterization of the mononuclear phagocyte system in zebrafish. *Blood* 117: 7126-7135
- Wu CC, Kruse F, Vasudevarao MD, Junker JP, Zebrowski DC, Fischer K, Noel ES, Grun D, Berezikov E, Engel FB *et al* (2016) Spatially Resolved Genome-wide Transcriptional Profiling

Identifies BMP Signaling as Essential Regulator of Zebrafish Cardiomyocyte Regeneration. *Developmental cell* 36: 36-49

Xia M, Liu J, Wu X, Liu S, Li G, Han C, Song L, Li Z, Wang Q, Wang J *et al* (2013) Histone methyltransferase Ash1l suppresses interleukin-6 production and inflammatory autoimmune diseases by inducing the ubiquitin-editing enzyme A20. *Immunity* 39: 470-481

Yoo W, Lee J, Noh KH, Lee S, Jung D, Kabir MH, Park D, Lee C, Kwon KS, Kim JS *et al* (2019) Progranulin attenuates liver fibrosis by downregulating the inflammatory response. *Cell death & disease* 10: 758

Yu H (2007) Chk1: a double agent in cell cycle checkpoints. *Developmental cell* 12: 167-168

## Figure Legends

**Figure 1. Single-cell transcriptomics reveals heterogeneity of zebrafish cardiac non-myocytes.** **A**, Experimental workflow of nonCM isolation from zebrafish hearts and scRNA-seq (10x Genomics). **B**, scRNA-seq data of adult zebrafish cardiac nonCM visualized on tSNE and colored by cell types. EC, endothelial cells; FB, fibroblasts; MC, macrophages; Mes, resident mesenchymal cells; T/NK/B, T/NK/B cells; Neutro, neutrophils; Eryth, erythrocytes; Throm, thrombocytes. **C**, Violin plots showing expression of canonical markers for each cell type. GO analysis (DAVID) of upregulated genes in each population was performed and representative GO terms were listed on the right. **D-F**: NonCMs expressing canonical EC markers in panel b were zoom-in analyzed with LIGER. **D**, Cells visualized on tSNE and colored by cell types. eEC, endocardial EC; lEC, lymphatic EC; cEC, coronary EC; Mural, mural cells. **E** Pie chart showing contribution of each cell type. **F** Expression of newly identified markers of each EC subpopulation shown on tSNE. **G**, Pie chart showing nonCM composition (erythrocytes and thrombocytes excluded). **H**, Dotplot showing expression of top 8 positive markers identified for each nonCM population.

## Figure 2. Transcriptome dynamics of macrophages during zebrafish heart regeneration.

**A, B**, Joint analysis of nonCM scRNA-seq data from uninjured hearts and hearts at 2, 7, and 14 dpi with LIGER. **A**, All nonCM visualized on tSNE. **B**, top, NonCM from each time point visualized on the tSNE embedding in (A). Data was down-sampled to the same cell number at each time point. **B**, bottom, Pie charts showing contribution of different cell types at each time

point. **C-F**, Zoom-in analysis of macrophages identified in (A). **C**, tSNE plot colored by MC subpopulations. **D**, Expression levels of representative markers of each MC subpopulation color-coded and mapped to tSNE embeddings in (C) (top), and corresponding representative GO terms for each subpopulation (bottom). **E**, Dynamics of MC subpopulations during heart regeneration as shown by the proportion of each subpopulation in total nonCM or total MC. **F**, Expression of *tnfa* and *csf3b* in MC at each time point. **G**, Fluorescent *in situ* hybridization for *cd74a*, *tnfa*, and *ctsc* at 7 dpi, respectively. *mpeg1* was used to label all MC cells. The blue boxed region is highlighted in the zoom-in image to the right. Scale bar = 50  $\mu$ m. A stands for atrium, V stand for ventricle and OFT stands for outflow tract. White dashed lines outline the heart and the yellow dashed lines indicate approximate resection plane.

**Figure 3. Transcriptome dynamics of fibroblasts and endothelial cells during zebrafish heart regeneration.**

**A-C**, Zoom-in analysis of fibroblasts identified in Fig 2. **A**, tSNE plot colored by FB subpopulations. **B**, Dot plot showing expression of ECM genes in FB subpopulations. **C**, Bar plot showing relative proportion of FB subpopulations in FB at each time point. **D-E**, Expression of *fn1a*, *coll2a1a* (D) and *tagln* (E) in FB at each time point. **F**, *In situ* hybridization showing temporal spatial expression patterns of *postnb*. The white boxed region is shown in zoom-in images at the bottom. Scale bar = 25  $\mu$ m. White dashed lines outline the heart and the yellow dashed lines indicate approximate resection plane. A stands for atrium, V stand for ventricle and OFT stands for outflow tract. **G**, Expression of *tnfrsf1a* and *tgfbr2a* in FB at each time point. **H**, Fluorescent *in situ* hybridization for *postnb*, *tnfa*, and *tnfrsf1a* in the injury area at 7 dpf, respectively. The white boxed regions are shown in their respective zoom-in images to the right. White dashed lines outline cardiac apex and the yellow dashed lines indicate approximate resection plane. Scale bar = 25  $\mu$ m. **I**, tSNE plot colored by eEC subpopulations. Bar plot showing relative proportion of eEC subpopulations in eEC at each time point. **J**, Violin plots showing expression of *cxcl18b*, *sele*, *atf3*, and *fosl1a* in each eEC subpopulation. In panels E and I, “un” stands for uninjured. **K**, Fluorescent *in situ* hybridization for *fosl1a* at 7 dpi. The white boxed region is shown in a higher magnification image to the right for the injury area. *cdh5* was used to label all EC cells and the red boxed region is shown in a higher magnification image to highlight a pan-EC *cdh5* expression in green. White dashed lines outline the heart and the yellow



dashed lines indicate approximate resection plane. Scale bar = 25  $\mu\text{m}$ . A stands for atrium, V stand for ventricle and OFT stands for outflow tract.

**Figure 4. Loss of *kit* function impairs heart regeneration.**

**A**, Wildtype or *kit* mutant stained with AFOG at 14 days and 30 dpi. Sham-operated zebrafish hearts serve as control. Scale bar = 100  $\mu\text{m}$ . **B**, Quantification of scar area at 30 dpi. N = 10 hearts. Data are presented as Mean  $\pm$  SEM. P value calculated with two-tailed students' t test. \*\* P < 0.01. **C**, PCNA/Mef2 double staining showing the proliferating CMs in wildtype and mutant hearts at 7 dpi. The white boxed regions are shown in zoom-in images to their right. Blue dashed lines indicate approximate resection plane. White arrows point to the PCNA/Mef2 double positive nuclei. Scale bar = 50  $\mu\text{m}$ . **D**, Quantification of % PCNA+ CMs (Mef2+) in C. N = 10 hearts. Data are presented as Mean  $\pm$  SEM. P value calculated with two-tailed students' t test. \*\* P < 0.01. **E-F**, Analysis of bulk RNA-seq data of freshly isolated CMs at different time points during heart regeneration. **E**, Two representative gene sets from GSEA analysis, which are enriched in DEGs between mutant and wildtype CMs at 2 dpi. **F**, Heatmap showing expression levels of cell cycle regulators in CMs from wildtype and *kit* mutant hearts. **G-H**, Joint analysis of *kit* mutant nonCM scRNA-seq data from 0, 2, 7, and 14 dpi with LIGER. **G**, All nonCM visualized on tSNE. **H**, Bar plots showing comparison of nonCM cell type contribution between wildtype and *kit* mutant at each time point. **I**, Immunostaining for MC marker IB4 in WT (left) and mutant hearts (middle) at 7 dpi. Quantification of IB4+ cell number per unit area (right). The white boxed regions are shown in their respective zoom-in images to the right. White dashed lines indicate approximate resection plane. Scale bar = 50  $\mu\text{m}$ . N = 6 hearts. Data are presented as Mean  $\pm$  SEM. P value is calculated with two-tailed students' t test. \* P < 0.05. In panels f and h, "un" stands for uninjured.

**Figure 5. Loss of *kit* function altered transcriptome dynamics of cardiac nonCM during zebrafish heart regeneration.**

**A-G**, Joint analysis of WT and *kit* mutant MC. **A**, tSNE plot colored by WT and *kit* mutant MC subpopulations. **B**, Representative GO terms of genes upregulated in MC1 (WT, red bars) or KMC1 (*kit* mutant, blue bars) from the uninjured hearts, respectively. Dotted lines indicate P= 0.05. **C**, Bar plots showing proportion of each WT and *kit* mutant MC subpopulation in total

nonCM at each time point. **D, F**, Violin plots showing the expression of *illb* in MC1 (**D**), and *ctsd* expression in MC (**F**). **E, G**, Expression of *illb* (**E**) and *ctsd* (**G**) in WT and *kit* mutant hearts determined by qRT-PCR. **H-N**, Joint analysis of WT and *kit* mutant FB. **H**, tSNE plot colored by WT and *kit* mutant FB subpopulations. **I**, Bar plots showing proportion of each WT and *kit* mutant FB subpopulation in total nonCM at each time point. **J, L**, Violin plots showing expression of *coll2a1a* in FB3 (**J**) and *mif* expression in FB (**L**). **K, M**, Expression of *coll2a1a* (**K**) and *mif* (**M**) in WT and *kit* mutant hearts determined by qRT-PCR. **N**, Bar plots showing relative proportion of each WT and *kit* mutant eEC subpopulation in eEC at each time point. **O**, Violin plots showing expression of *junba* in eEC. **P**, Expression of *junba* in WT and *kit* mutant hearts determined by qRT-PCR. **Q**, Representative image of hearts from 7 dpi *tcf21:nucGFP* transgenic fish treated by PBS liposome, Clodronate liposome and Dex, respectively. The red boxes mark the peripheral area, and the white boxes mark the injury area and are shown in their respective zoom-in images to the right.. White dashed lines indicate approximate resection plane. Scale bar = 50  $\mu$ m. **R**, Quantification of the number of *tcf21:nucGFP* positive cells in the boxed areas in U. N = 5 hearts. Data are presented as Mean  $\pm$  SEM. Two-tailed students' t test. \*\*\* P < 0.001. Data information: **E, G, K, M, P**. Mean  $\pm$  SEM shown. N = 3. P value calculated with two-tailed students' t test. \* P < 0.05, \*\* P < 0.01, \*\*\* P < 0.001. In panels (C-G, I-P), "un" stands for uninjured. Clod lipos stands for clodronate liposomes. Dex for dexamethasone.

**Figure 6.** Combined Topologizer and RNA velocity analyses identified molecular regulatory mechanism associated with MC functional state transition during heart regeneration.

**A**, Schematic of Topologizer. **B**, Topological structure of WT MC subpopulations. **C**, Vector field of RNA velocity projected onto the Topologizer structure for MCs collected from uninjured WT hearts or WT hearts at 2, 7, and 14 dpi. Arrows indicate the direction and "speed" of the velocity at each node. **D**, Quantification of arrow length in C. P value is calculated with one-tailed Wilcoxon Rank Sum test. \* P < 0.05. **E**, Transcriptional similarity between MC1 and MC2 subtypes from the uninjured (un) hearts or hearts at 2, 7 and 14 dpi, respectively. **F**, Hierarchical tree showing the relationship between the MC subtypes and the myeloid cells from kidney marrow. **G**. Three gene clusters identified associated with MC2 to MC1 transition. **H**, Feature significant gene ontology (GO) terms (adjusted *p*-value < 0.05) with representative genes. The number of genes is shown in parentheses. Dotted lines indicate P= 0.05. **I**, Topological structure

of WT FB subpopulations. **J**, Vector field of RNA velocity projected onto the Topologizer structure for FBs collected from uninjured WT hearts or WT hearts at 2, 7, and 14 dpi. Arrows indicate the direction and “speed” of the velocity at each node. **K**, Topological structure of WT eEC subpopulations. **L**, Vector field of RNA velocity projected onto the Topologizer structure for eECs collected from uninjured WT hearts or WT hearts at 2, 7, and 14 dpi. Arrows indicate the direction and “speed” of the velocity at each node.

Author Manuscript

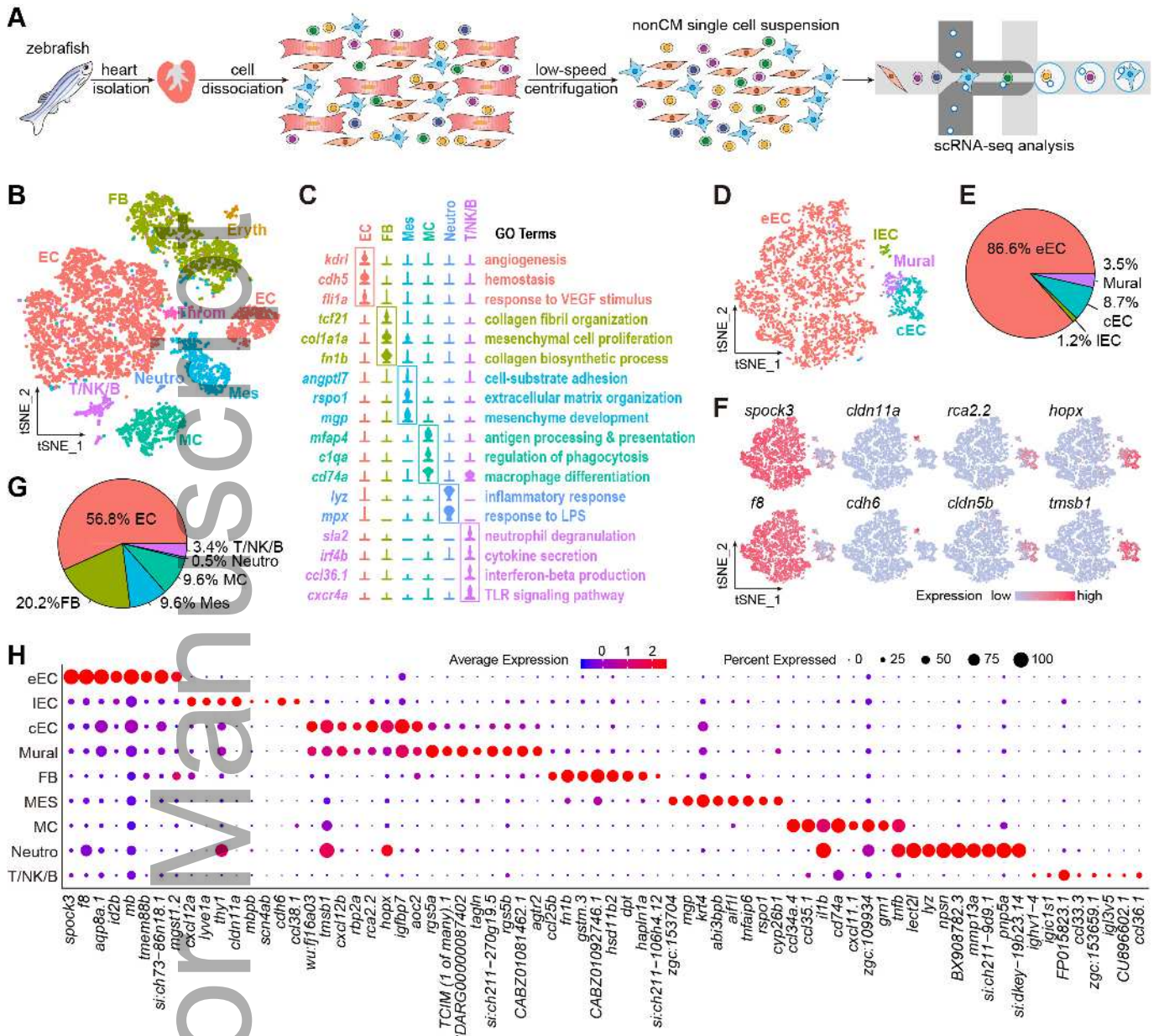
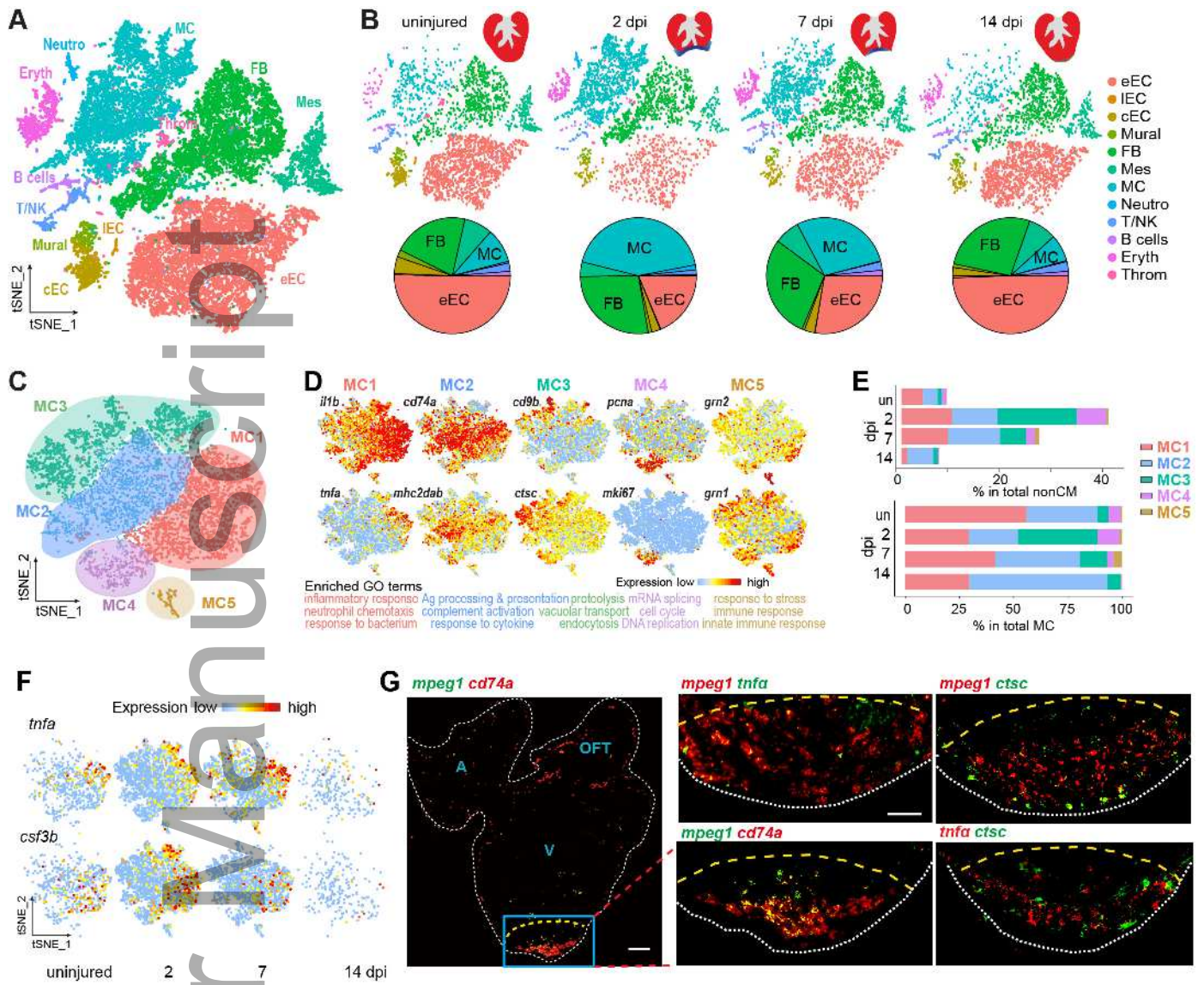


Figure 1

embr\_202152901\_f1.jpg



**Figure 2**

embr\_202152901\_f2.jpg

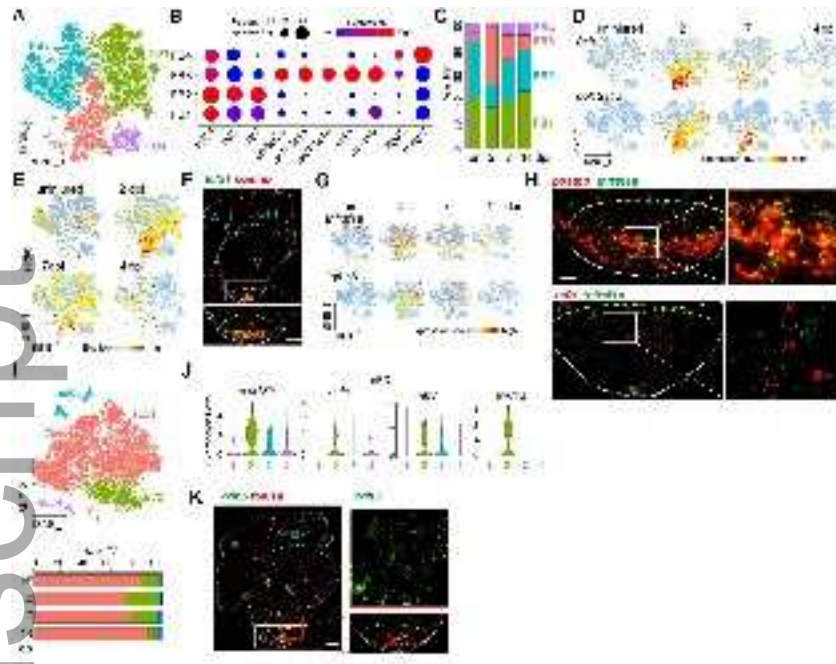


Figure 3

embr\_202152901\_f3.jpg

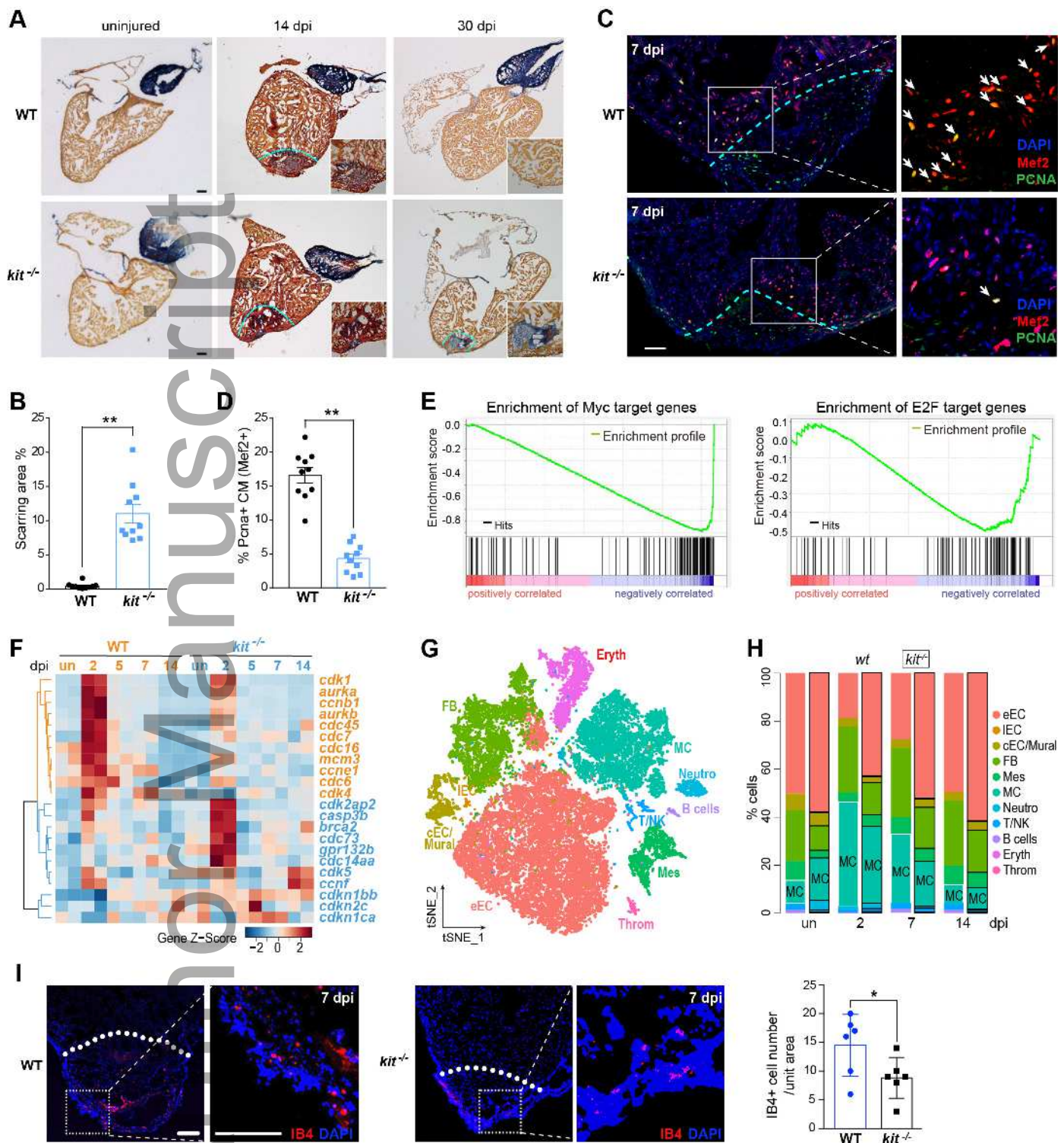


Figure 4

embr\_202152901\_f4.jpg

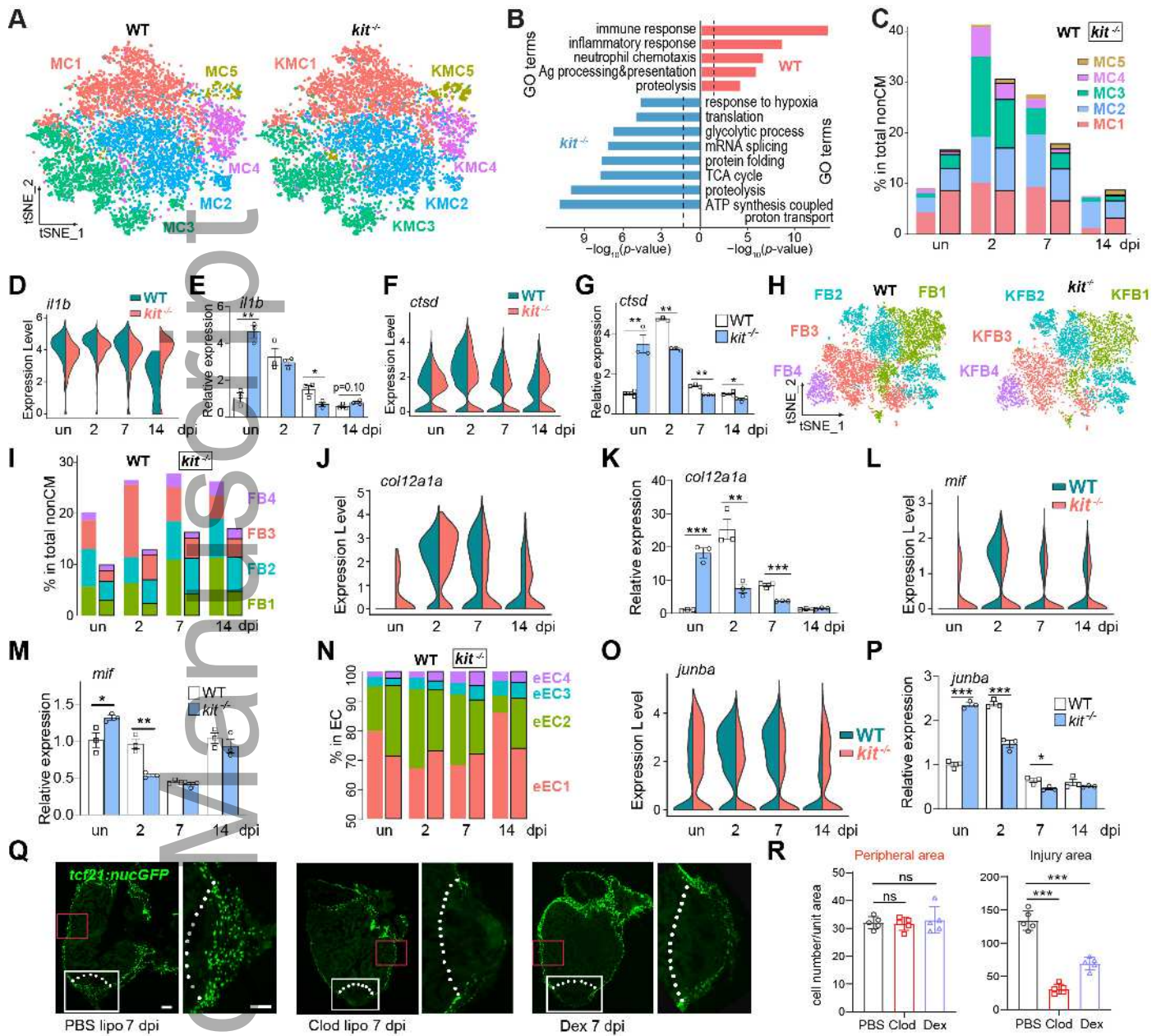


Figure 5

embr\_202152901\_f5.jpg



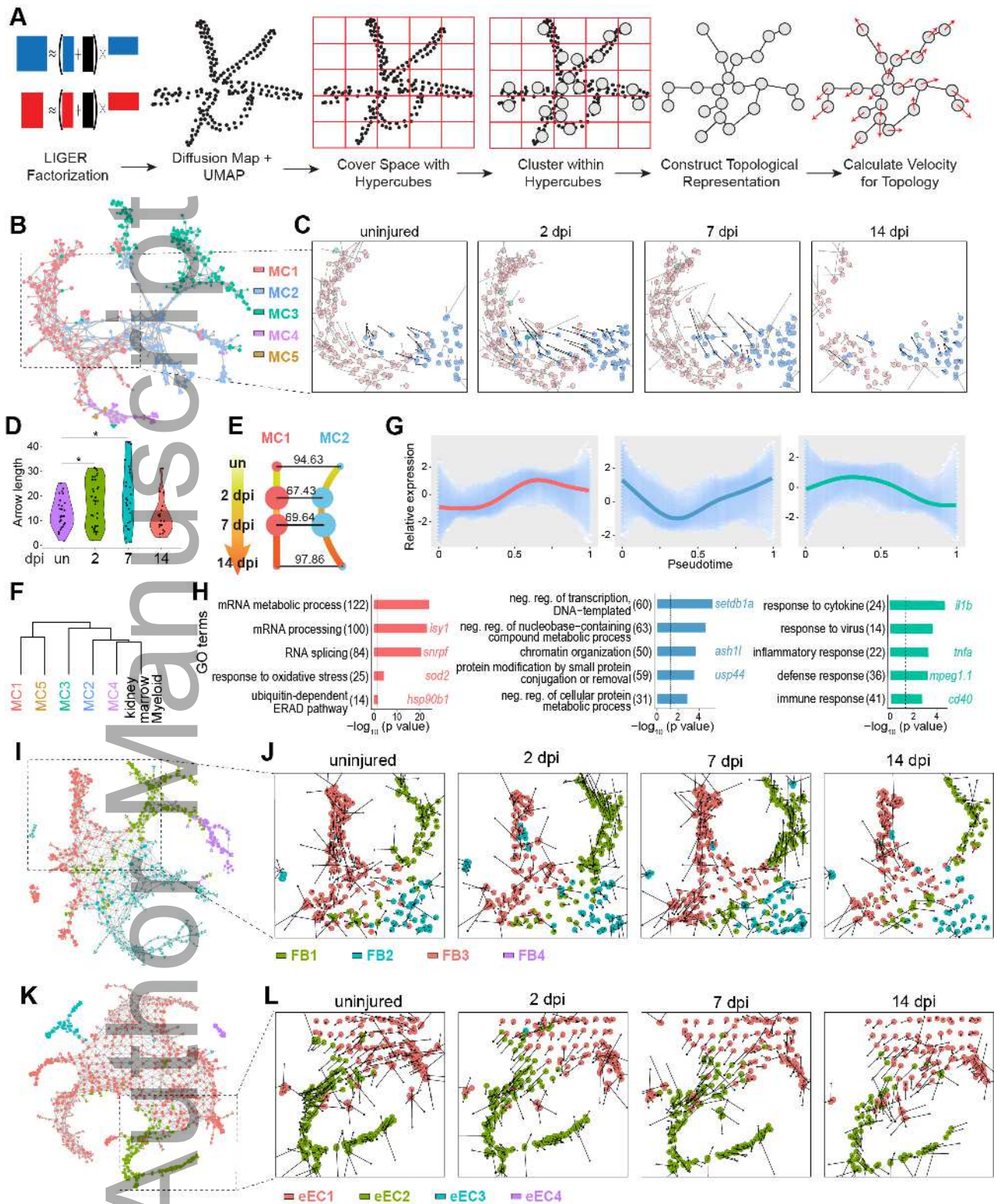


Figure 6

embr\_202152901\_f6.jpg